Title
Fixed and Random Effects in Panel Data Using Structural Equations Models

Authors
Bollen, Kenneth A.
Brand, Jennie E.

Publication Date
2008

California Center for Population Research
University of California - Los Angeles

Fixed and Random Effects in Panel Data Using Structural Equations Models

Kenneth A. Bollen
Jennie E. Brand

PWP-CCPR-2008-003

January 2008

California Center for Population Research
On-Line Working Paper Series

FIXED AND RANDOM EFFECTS IN PANEL DATA USING STRUCTURAL EQUATON MODELS*

Kenneth A. Bollen
H.W. Odum Institute for Research in Social Science
Department of Sociology & Statistics
University of North Carolina, Chapel Hill

Jennie E. Brand
Department of Sociology
California Center for Population Research
University of California, Los Angeles

Word Count: 12,961

*Kenneth Bollen gratefully acknowledges support from NSF SES 0617276 from NIDA 1-RO1-DA13148-01 & DA013148-05A2. Earlier versions of this paper were presented at the Carolina Population Center at the University of North Carolina at Chapel Hill, the Methods of Analysis Program in Social Sciences (MAPSS) Colloquium at Stanford University, the Social Science Research Institute at Duke University, and the 2007 North Carolina Chapter of the American Statistical Association conference at the University of North Carolina, Greensboro. We thank Guang Guo for helpful comments. Direct correspondence to: Kenneth A. Bollen, CB 3210 Hamilton, Department of Sociology, University of North Carolina, Chapel Hill, NC 27599-3210 (bollen@unc.edu).

Running Head: FIXED AND RANDOM EFFECTS MODELS

Contact Information:

Kenneth A. Bollen
CB 3210 Hamilton
Department of Sociology
University of North Carolina – Chapel Hill
Chapel Hill, NC 27599-3210
(919) 843-5990
bollen@unc.edu

Jennie E. Brand
264 Haines Hall – Box 951551
Department of Sociology
University of California – Los Angeles
Los Angeles, CA 90095-1551
(310) 266-0826
brand@soc.ucla.edu

KENNETH A. BOLLEN

Kenneth A. Bollen is Director of the Odum Institute for Research in Social Science and the Immerwahr Distinguished Professor at the University of North Carolina. In 2000 he received the *Lazarsfeld Award for Methodological Contributions in Sociology*. The *ISI* named him among the *World's Most Cited Authors* in the Social Sciences. In 2005 three of his articles were recognized as among the most cited articles in the history of the *ASR*. He has published *Latent Curve Models* (with P. Curran, 2006, Wiley), *Structural Equation Models with Latent Variables* (1989, Wiley) and over 100 papers. Bollen's primary research areas are in structural equation models, population studies, and democratization.

JENNIE E. BRAND

Jennie E. Brand is Assistant Professor of Sociology at the University of California – Los Angeles and a faculty affiliate at the California Center for Population Research. Her research focuses on the relationship between social and human capital, educational attainment, job conditions, and socioeconomic attainment and well-being over the life course. She is also interested in the application and innovation of statistical models for panel data. Current research projects include evaluation of heterogeneity in the effects of early schooling and college completion on socioeconomic and family outcomes, and the effects of employment insecurity, job loss, and job conditions on health.

FIXED AND RANDOM EFFECTS IN PANEL DATA USING STRUCTURAL EQUATON MODELS

ABSTRACT

Applications of classic fixed and random effects models for panel data are common in sociology and in ASR. A primary advantage of these models is the ability to control for time-invariant omitted variables that may bias observed relationships. This paper shows how to incorporate fixed and random effects models into structural equation models (SEMs) and how to extend the standard models to a wide variety of more flexible models. For instance, a researcher can test whether a covariate's impact on the repeated measure stays the same across all waves of data; test whether the error variances should be allowed to vary over time; include lagged covariates or lagged dependent variables; estimate the covariance of the latent time-invariant variables with the observed time-varying covariates; and include observed time-invariant variables in a fixed effects model. The paper explains how to take advantage of the estimation, testing, and fit assessment capabilities that are readily available for SEMs and how to reveal flaws not visible with typical assessment techniques. The paper is oriented towards applied researchers with most technical details given in the appendix and footnotes.

FIXED AND RANDOM EFFECTS IN PANEL DATA USING STRUCTURAL EQUATON MODELS

INTRODUCTION

Longitudinal data are more available today than ever before. The National Longitudinal Study of Youth (NLSY), National Longitudinal Study of Adolescent Health (Add Health), Panel Study of Income Dynamics (PSID), and National Education Longitudinal Study (NELS) are just a few of the more frequently analyzed panel data sets in sociology. The relative advantages of longitudinal data compared to cross-sectional are well-known (Halaby 2004) and panel data are permitting more sophisticated analyses than were available before.

In sociology two common analysis tools for such data are referred to as the random and fixed effects models (Allison 1994; Guo and Hipp 2004). Indeed, in this journal alone a number of recent papers have made use of fixed and random effects models (Alderson 1999; Alderson and Nielsen 1999; Conley and Bennet 2000; Mouw 2000; Budig and England 2001; Wheaton and Clarke 2003; Yakubovich 2005; Beckfield 2006; Matsueda, Kreager, and Huizinga 2006). A major attraction of these models is that they provide a means to control for all time-invariant unmeasured (or latent) variables that influence the dependent variable whether these time-invariant variables are known or unknown to the researcher. Given the likely presence of such omitted variables, this is a major advantage of these models. As the term "time-invariant" suggests, these are variables that do not change over the period of observation while "time-varying" describes variables that change over time. The major difference of random effects models from the fixed effects model is that in the former the omitted time-invariant variables are assumed to be uncorrelated with the included time-varying covariates while in the latter they are allowed to correlate (Mundlak 1978). The random effects model has the advantage of greater efficiency relative to the fixed effects model leading to smaller standard errors and higher statistical power to detect effects (Hsiao 2003). A Hausman test enables researchers to distinguish between the random and fixed effects model (Hausman, 1978). Statistical software for random and fixed effects models is readily available (e.g., xtreg in Stata and Proc GLM, Proc Mixed in SAS).

Despite the many desirable features of the random and fixed effects models for longitudinal data there are a number of limitations of the standard implementations that are not fully appreciated by users. First, these models have implicit restrictions that are rarely tested but that if wrong, could bias the estimated effects. For instance, both the random and fixed effects models assume that the coefficients of the same covariate remain equal across all waves of data. If individuals pass through major transition points during the time period of the study (e.g., actively employed to retired or adolescence to adult), this assumption of stable effects for each wave could be invalid. So in an analysis of say, income's effect on conservatism, our model forces the impact of income on conservatism to remain the same across all waves of data. Another implicit restriction is that the unexplained variance stays the same over time. This means that despite the changes that might be occurring in an individual's life, the error variance is not permitted to differ. Another constraint is that the latent time-invariant variables must either correlate with all covariates as in the fixed effects model or be uncorrelated with all covariates as in the random effects model. Even with strong prior evidence that some correlations are zero and some are not, the researcher must choose all correlated or all uncorrelated when using these models. Incorrectly assuming that the latent time-invariant variable is uncorrelated with the observed covariates is likely to bias our estimates of effects.

Yet another implicit constraint is that autoregressive relations with lagged dependent variables are set to zero. In some areas, there may be reason to suspect that prior values of the dependent variable influence current values even net of other variables. Hsiao (2003) and Halaby (2004) provide good summaries of the complications that emerge with the usual random and fixed effects estimators in models with lagged endogenous variables and the corrections needed might discourage researchers from exploring these possibilities. A closely related alternative is that there are lagged effects of a covariate on the dependent variable, yet these lagged variables are rarely considered.[1]

If false, any of these preceding implicit restrictions could lead to biased and inconsistent estimators of the coefficients in the model. Furthermore in the fixed and random effects models, the observed covariates in the models are not permitted to influence the latent time-invariant variable. Given the zero correlation restriction between the covariates and latent time-invariant variable in random effects models, this is hardly surprising. The fixed effects model does allow correlation, but it does not permit direct effects between the covariates and the time-invariant variables.

Problems with these implicit restrictions go undetected because the usual fixed and random effects models lack tests of overall fit of the hypothesized model to the saturated model. As we will explain, the traditional fixed and random effects models are overidentified models. That is, the model structure implies that the means and covariance matrix of the observed variables can be predicted by the parameters of the model and there is more information in the means and covariance matrix than the minimum required to estimate parameters. This excess information allows overidentification tests that provide information on the goodness of fit of the model relative to fitting a saturated model. The overidentification test is evidence relevant to the validity of the model for the data. The Hausman test just compares the fixed and random effects models. It is possible that neither is adequate and this would not be known from the usual fixed and random effects output.

An advantage of the fixed effects model over the random effects model is that it permits the latent time-invariant variable to correlate with the time-varying covariates. However, the implementation of the fixed effects model does not report the magnitude of the correlations, something that could provide a better understanding of the latent time-invariant variable. Another disadvantage of the typical implementation of the fixed effects estimator is that it does not permit the researcher to include time-invariant observed variables such as sex, race, place of birth, etc. This is a disadvantage to the degree that a researcher is interested in the impact of these time-invariant observed variables.

Our purpose is to show how researchers can easily overcome these limitations by incorporating the fixed and random effects models into a structural equation modeling framework. More specifically, we will demonstrate that the traditional fixed and random effects models can be formulated as structural equation models (SEMs) and with the appropriate estimator will reproduce the estimates obtained with standard procedures.[2]

More importantly, we will create SEMs that overcome all of the preceding limitations that we described above for the fixed and random effects models. That is, we will show how to test whether coefficients or error variances are equal at each wave; we will illustrate how estimates of the covariance of the latent time-invariant variable and the covariates is estimable; we will develop models that permit lagged dependent variables or lagged covariates and create new models that permit time-invariant observed variables in fixed effects like models; we will provide alternatives to the Hausman test for comparing fixed and random effects models and will give a test of whether either of these models is adequate. We will illustrate our results using a model and data from an excellent empirical application of traditional fixed effects estimators that was published in this journal to demonstrate that we can gain new insights with this approach even in a well-executed study.

We emphasize that this paper is oriented to readers who are familiar with traditional fixed and random effects estimators. Our citations above to recent *ASR* publications reveal their presence in a broad range of fields in sociology and suggest that these techniques have broad appeal. Though new results are provided, we do *not* intend our results for specialists in quantitative methods though we hope that such readers will gain something from reading this paper. In fact, the models we describe can be readily implemented with any of the numerous SEM software programs (e.g., LISREL, AMOS, Mplus, EQS, etc.) that are widely available. Because of our intended audience, full technical details are not provided in the text, but are reserved for footnotes, the appendix, and in the cited works.

TRADITIONAL FIXED AND RANDOM EFFECTS MODELS

In this section we will provide a brief overview of the traditional fixed and random effects models for pooled cross-sectional and time-series data that are commonly used. More detailed accounts are in Allison (2005), Baltagi (2001), Greene (1997), Halaby (2004), Hsiao (2003), and Wooldrige (2002). Our overview will emphasize restrictions that are often implicit in other presentations in preparation for the more general models that we present in a later section. We begin with a general model that can represent either the random or fixed effects models depending on the restrictions placed on it. Then we discuss the random effects model, the fixed effects model, and the Hausman test that distinguishes between them.

General Model

Consider the following equation

$$y_{it} = \mathbf{B}_{y\mathbf{x}t}\mathbf{x}_{it} + \mathbf{B}_{y\mathbf{z}t}\mathbf{z}_i + \eta_i + \varepsilon_{it} \qquad (1)$$

where y_{it} is the value of the dependent variable for the i th case in the sample at the t th time period, $\mathbf{x}_{it}$ is the vector of time-varying covariates for the i th case at the t th time period, $\mathbf{B}_{y\mathbf{x}t}$ is the row vector of coefficients that give the impact of $\mathbf{x}_{it}$ on y_{it} at time t, $\mathbf{z}_i$ is the vector of observed time-invariant covariates for the i th case with $\mathbf{B}_{y\mathbf{z}t}$ its row vector of coefficients at time t, η_i is a scalar of all other latent time-invariant variables that influence y_{it}, and ε_{it} is the random disturbance for the i th case at the t th time period with $E(\varepsilon_{it}) = 0$ and $E(\varepsilon_{it}^2) = \sigma_{\varepsilon_t}^2$. It also is assumed that ε_{it} is uncorrelated with

$\mathbf{x}_{it}$, $\mathbf{z}_i$, and η_i. As an example, y_{it} might be the infant mortality rate in county i at time t, $\mathbf{x}_{it}$ might consist of time-varying variables such as unemployment rate, physicians per capita, medical expenditures per capita, etc. all for county i at time t, $\mathbf{z}_i$ might be time-invariant variables such as region and founding date of county, and η_i would contain all other time-invariant variables that influence infant mortality, but that are not explicitly measured in the model.

There are several things to notice about this equation. One is that i always indexes the cases in the sample while t indexes the wave or time period. If either subscript is missing from a variable or coefficient, then the variable or coefficient does not change either over individual or over time. For instance, $\mathbf{z}_i$ and η_i have no t subscript, but do have an i subscript. This means that these variables vary across different individuals, but do not change over time for that individual[3] and are time-invariant variables. Common examples of time-invariant observed variables ($\mathbf{z}_i$) for individuals would be characteristics such as race, sex, and place of birth. The time-invariant *latent* or *unobserved* variables that are in η_i represent the collection of all time-invariant variables that influence y_{it}, but that are not explicitly measured and included in the model. Implicitly, this variable has a coefficient of "1" that does not change over time. Intelligence or stable personality characteristics are examples of two types of variables that might influence an outcome of interest, but not be explicitly measured, are stable, and would therefore be part of η_i. In a similar fashion, the absence of an $\imath$ subscript means the coefficient does not change over individual. If a time period or wave of data were distinct, then we might include a dummy variable, say D_t, that is the same value across all individuals but could differ over time, though we do not do this in the above equation. In the above model, $\mathbf{B}_{y\mathbf{x}t}$ and $\mathbf{B}_{y\mathbf{z}t}$ are assumed to be the same values over individuals, but the subscript of t permits these to differ over time. This is different than the usual presentation of these models where the effects are assumed constant over time and the t subscript is missing. The subscript of t on $\sigma^2_{\varepsilon_t}$ signifies that the error variance for the equation can differ depending on the wave of data.

Equation (1) is more general than the traditional fixed and random effects model. However, we can use it to get to these models by introducing constraints. We do so starting with the random effects model.

Random Effects Model

The random effects model is

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \mathbf{B}_{y\mathbf{z}}\mathbf{z}_i + \eta_i + \varepsilon_{it} \qquad (2)$$

where all variables are defined as above. At first glance this appears the same as equation (general model), but there are differences. First, the random effects model assumes that the effects of $\mathbf{x}_{it}$ and $\mathbf{z}_i$ on y_{it} do not change over time. This is why $\mathbf{B}_{y\mathbf{x}}$ (no t subscript) replaces $\mathbf{B}_{y\mathbf{x}t}$ and $\mathbf{B}_{y\mathbf{z}}$ replaces $\mathbf{B}_{y\mathbf{z}t}$. Second, the random effects model assumes that η_i is a random latent variable that is uncorrelated with ε_{it}, $\mathbf{x}_{it}$ and $\mathbf{z}_i$. Considering the type of unmeasured time-invariant variables that might appear in η_i (e.g., intelligence, personality characteristics), it is a strong assumption to force these to be uncorrelated with all other explanatory variables in the analysis. If this assumption or the assumption of the same coefficients over time is incorrectly imposed, either can bias the estimated effects that we find. Another assumption of the random effects model is that the error variance does not change over time ($\sigma^2_{\varepsilon_t} = \sigma^2_{\varepsilon}$). Violation of this assumption can lead to inaccuracy of the estimates.

One common estimator of the coefficients and variances of the error and time-invariant latent variable is the feasible Generalized Least Square (GLS) estimator (e.g., Hsiao 2003:35-38; Wooldridge 2002:257-65). In this approach, an estimate of the variances of η_i and of ε_{it} are used to form the "weight matrix" for GLS estimation.[4] If the preceding assumptions hold, then this procedure has desirable large sample properties such as consistency, asymptotic unbiasedness, and readily available significance tests. It also provides an estimate and test of whether there are latent time-invariant variables (H_o : $\sigma^2_{\eta} = 0$) such that zero variance of ☼$_i$ implies their absence.[5] A maximum likelihood estimator of the random effects model also is possible under the assumption that η_i and ε_{it} come from a normal distribution (Hsiao 2003:39-41). One or both of these estimators are available in statistical software such as SAS (e.g., TSCSREG) or STATA (xtreg). However, the use of these procedures presupposes that the assumptions of the random effects model hold, with the assumption that the latent time-invariant variable is uncorrelated with all observed covariates being the most problematic assumption. The fixed effects model removes this restriction.

Fixed Effects Model

The equation for the fixed effects model is

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \eta_i + \varepsilon_{it} \tag{3}$$

where the symbols are defined as above. The most obvious difference between the fixed effects model and the random effects one is the absence of the $\mathbf{B}_{y\mathbf{z}}\mathbf{z}_i$ term. These are the time-invariant *observed* variables and their coefficients. The traditional fixed effects model does not include these variables, but rather folds them into η_i, the latent time-invariant variable term. The reason is that the fixed effects model allows η_i to correlate with $\mathbf{x}_{it}$ and if we were to also include time-invariant *observed* variables ($\mathbf{z}_i$), these would be perfectly collinear with η_i and we could not get separate estimates of the effects of η_i and $\mathbf{z}_i$. Hence, we allow η_i to include $\mathbf{z}_i$ as well as latent time-invariant variables. Though losing the ability to estimate the impact of time-invariant variables such as race, sex, etc. is a disadvantage of the technique, we still are controlling for their effects by including η_i in the model. In addition, we have a more realistic assumption that the η_i variable correlates with the time-varying covariates in $\mathbf{x}_{it}$. In the hypothetical example of infant mortality rates in counties, we could not include the time-invariant variables of region and founding date explicitly in the model. But these and all other time-invariant variables would be part of η_i and hence controlled. If a researcher is not explicitly interested in the specific effects of the time-invariant variables, then this is not a serious disadvantage of the fixed effects model since the potentially confounding effects of all time-invariant variables would be controlled. In addition, we allow these time-invariant variables to correlate with the time varying variables such as unemployment, physicians per capita, and so on.

Another subtle difference of the fixed effects versus the random effects model is that the η_i variable is a fixed unknown constant for each case. That is, the collection of time-invariant variables takes a single unchanged value rather than being one realization of a random variable in the fixed effects model. In contrast, the random effects model treats η_i as a latent *random* variable. In practice, this distinction between η_i as a constant that differs by case or as a random variable is largely inconsequential (Mundlak 1978). A far more important difference is that the fixed effects model allows η_i to correlate with $\mathbf{x}_{it}$ whereas the random effects model forces this correlation to zero. Though on this count, the fixed effects model is less restrictive than the random effects models, both models contain often unrealistic restrictive assumptions. For instance, both models assume that the coefficients of the time-varying variables ($\mathbf{B}_{y\mathbf{x}}$) do not change over time, meaning that these variables have the same effect in each wave of data and the η_i variables have the same effect (implicit coefficient of 1) on y_{it} for each time period. The traditional fixed effects like the random effects models also assumes that the error variance is constant over time ($\sigma^2_{\varepsilon_t} = \sigma^2_{\varepsilon}$).

Two common estimators for the fixed effects model are the least squares dummy variable (LSDV) estimator and the mean transformed data approach. In LSDV a separate dummy variable is coded for each case and entered into the model to control for η_i (see, e.g., Hsiao 2003). Though this is tedious and sometimes impractical for large samples, it can work well for small samples and the coefficients for the dummy variables provide an estimate of the η_i for each case. In large samples researchers are more likely to run their model on a transformation that removes the fixed effects from the model. In this case, ($y_{it} - \bar{y}_{i\cdot}$) is regressed on ($\mathbf{x}_{it} - \bar{\mathbf{x}}_{i\cdot}$) where $\bar{y}_{i\cdot}$ is the mean of the dependent variable for the i th case across all waves of data and $\bar{\mathbf{x}}_{i\cdot}$ is the mean of the time-varying variables for the i th case across all waves. In this approach, the latent time-invariant variable can be estimated as $\eta_i = (\bar{y}_{i\cdot} - \mathbf{B}_{y\mathbf{x}}\bar{\mathbf{x}}_{i\cdot})$ [see, e.g., Wooldridge 2002:265-272].

Hausman Test

Though they share similar features, the random and fixed effects models differ in important ways. Most importantly, the random effects model assumption that the latent time-invariant variables (η_i) are uncorrelated with the time-varying ($\mathbf{x}_{it}$) and time-invariant ($\mathbf{z}_i$) variables is a strong assumption. If true, then the random effects model has the advantage of permitting time-invariant observed variables ($\mathbf{z}_i$) as well as time-invariant latent variables (η_i) in the model. It also leads to a more efficient estimator than the fixed effects model in that the standard errors of the coefficients will tend to be smaller enabling the detection of smaller effects. However, if the assumption of η_i being uncorrelated with $\mathbf{x}_{it}$ and $\mathbf{z}_i$ is incorrect, then application of the random effects model will bias coefficients and possibly undermine a researcher's results.

The Hausman Test is a way of determining the plausibility of the fixed versus random effects model (Greene, 1997). Formally, the test is

$$T_{Hausman} = \left[\widehat{\mathbf{B}}_{y\mathbf{x}FE} - \widehat{\mathbf{B}}_{y\mathbf{x}RE}\right]' \left[avar\left(\widehat{\mathbf{B}}_{y\mathbf{x}FE}\right) - avar\left(\widehat{\mathbf{B}}_{y\mathbf{x}RE}\right)\right]^{-1} \left[\widehat{\mathbf{B}}_{y\mathbf{x}FE} - \widehat{\mathbf{B}}_{y\mathbf{x}RE}\right] \quad (4)$$

where $\hat{\mathbf{B}}_{yxFE}$ are the coefficient estimates of the time-varying covariates from the fixed effects model and $\hat{\mathbf{B}}_{yxRE}$ are the corresponding estimated coefficients from the random effects model. The $avar\left(\hat{\mathbf{B}}_{yxFE}\right)$ is the estimate of the asymptotic (large sample) variances and covariances of the $\hat{\mathbf{B}}_{yxFE}$ estimated coefficients and $avar\left(\hat{\mathbf{B}}_{yxRE}\right)$ is the analogous quantity for the estimate of $\hat{\mathbf{B}}_{yxRE}$. Intuitively, the coefficients from the fixed and random effects model should converge to the same parameter values if the random effects model is true. That is, the fixed effects model might needlessly allow the latent time-invariant variables to correlate with the time-varying variables, but other than leading to more variance of the coefficients, the estimator remains a consistent estimator of the true coefficients. This implies that $\hat{\mathbf{B}}_{yxFE}$ and $\hat{\mathbf{B}}_{yxRE}$ should only differ within sampling fluctuations. The central term in brackets on the right-hand side of equation (Hausman test) is an estimate of the sampling variance of this difference in coefficients. The test statistic, $T_{Hausman}$, follows a chi-square distribution in large samples with degrees of freedom equal to the number of coefficients for the time-varying variables. The null hypothesis is that the random effects model is true so that these coefficients are equal in the population. The alternative hypothesis is that at least one of these coefficients differ and hence the fixed effects model is more plausible. In practice, it is quite common to reject the null hypothesis of the random effects model in favor of the fixed effects model.

Empirical Example

We illustrate the classic fixed and random effects models by examining the wage penalty for motherhood using data from the National Longitudinal Survey of Youth (NLSY). The NLSY is a national probability sample of 12,686 young men and women who were 14 to 22 years old when they were first interviewed in 1979; blacks and

Hispanics are oversampled. These individuals were interviewed annually through 1994 and biannually thereafter. We begin by replicating the results from an earlier study by Budig and England (2001), published *ASR*, and examine data from the 1982-93 waves of the NLSY. Budig and England were interested in whether the relationship between number of children and women's earnings is spurious or causal, and use fixed effects models to address this question. This study builds on a still earlier study, also published in *ASR*, Waldfogel (1997), which also uses traditional fixed effects estimators to examine the wage penalty for motherhood. Budig and England's study is an excellent empirical application of traditional fixed effects estimators. Nevertheless, we will show how we gain new insights using our approach.

We limit our sample to women employed part-time or full-time during at least two of the years from 1982-93, to replicate Budig and England's sample selection. Out of a total of 6,283 women in the 1979 NLSY, we have a final sample size of 5,285 women.[6] The dependent variable is log hourly wages in the respondent's current job, where person-years whose hourly wages appear to be outliers (i.e., less than $1 or above $200 per hour) are eliminated. The main independent variable of interest is the total number of children that a respondent reported by the interview date.[7] Our baseline model, Model 1, includes only number of children as a covariate. In Model 2, we control for marital status using dichotomous measures to indicate married and divorced (including separated and widowed), where never married is the reference category. In Model 3, we further control for measures of human capital including: years of education, years of full-time and part-time work experience, years of full-time and part-time job seniority, and the total number of breaks in employment.[8] We also control for whether or not the respondent is currently in school. Budig and England also include a fourth model with a range of job characteristics. However, they find that these additional variables do little to change their estimates of the wage penalty for motherhood. We therefore only replicate the first three models.

[TABLE 1 ABOUT HERE]

For all their models, Budig and England conduct a Hausman test to assess whether random effects models were adequate, and in all cases, they find that the test indicated a need for fixed effects models. Therefore, they do not present the estimates for random effects models, only for fixed effects models and OLS models. In Table 1, we present our estimates of the effect of number of children on log earnings for Models 1-3 using Stata SE 9.0 xtreg for both fixed effects and random effects models. These results generally replicate Budig and England's fixed effects findings [(reported in Table 2 of Budig and England (2001)], with slight differences in Model 3.[9] We find a 7 percent wage penalty per child that decreases to about a 4 percent penalty with controls for human capital variables. Random effects estimates for Models 1 and 2 are larger, indicating about an 8 percent penalty per child, but the estimate decreases to less than 4 percent with controls for human capital variables. We also conduct Hausman tests for Models 1-3 to determine whether random effects models performed adequately; Model 1 has a $T_{Hausman}$ = 39.66 with 4 degrees of freedom, Model 2 a $T_{Hausman}$ = 68.56 with 6 degrees of freedom, and Model 3 a $T_{Hausman}$ = 349.37 with 13 degrees of freedom, each indicating that the fixed effects were preferred to the random effects models.

These results represent the traditional random and fixed effects models and software. We now turn to how we can replicate these results and create alternative models by moving to SEMs.

TRADITIONAL AND ALTERNATIVE MODELS AS SEMS

Statistical software that has procedures to handle fixed and random effects models generally provide specialized procedures that are customized to these models. In addition, there are some extensions to these models that overcome some restrictions that we have mentioned such as allowing the error variance to vary over time, including lagged dependent variables, and permitting autocorrelated disturbances. However, there does not appear to be any software that permits all these options and there are other limitations of the traditional models that are not addressed. In this section, we demonstrate that generic structural equation models (SEMs) software can incorporate these traditional models and that doing so renders advantages. SEMs refer to a general multiequation model that permits latent and observed variables, multiple measures of latent variables, and takes account of measurement error when estimating relationships (see e.g., Bollen 1989; Arminger and Browne 1995). SEMs are sufficiently general so that they can incorporate the traditional and nontraditional random and fixed effects models as special cases. It is this aspect that is of primary interest to us.

Random Effects Model

The traditional random effects model equation is

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \mathbf{B}_{y\mathbf{z}}\mathbf{z}_i + \eta_i + \varepsilon_{it} \qquad (5)$$

where the definitions and assumptions were given above. Figure 1 is a path diagram representation of a random effects model that is kept simple with a single time-varying variable (x s) for four waves of data and a single time-invariant variable (z_1). A path diagram is a graph that represents a multiequation system and its assumptions. By convention, boxes represent observed variables, ovals represent latent variables, single-headed straight arrows represent the direct effect of the variable at the base of the arrow on the variable at the head of the arrow, and two-headed arrows such as those connecting the x s and z_1 stand for possible associations between the connected variables where that association is taken account of, but not explained within the model.[10] To simplify the notation the i subscript is excluded for the variables. It is noteworthy that the latent time-invariant variable (η) is part of the model, but it is shown to be uncorrelated with the time-varying variables ($\mathbf{x}_t$) and the time-invariant variable (z_1) since there are no two-headed arrows linking it to the observed variables. The direct impact of the latent time-invariant variable (η) on the repeated measures (y s) is equal to 1 as is implicit in the equation for the random effects model.

[FIGURE 1 ABOUT HERE]

We could remove many of the restrictions that are implicit and explicit in this model (e.g., equal coefficients over time, equal error variances). The modifications are so similar to those for the fixed effects model that we momentarily postpone discussing them until we present the SEM approach to the fixed effects model in the next section.

Fixed Effects Model

The traditional fixed effects model equation is

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \eta_i + \varepsilon_{it} \tag{6}$$

where all symbols and assumptions are as previously. This model also is straightforward to treat as a SEM. Figure 2 is a path diagram representation of a fixed effects model that is kept simple with a single time-varying variable. Easily visible within the diagram is the covariance of the time-varying x_{1t} and η that is part of the fixed effects specification. But one difference from the usual implementation of fixed effects models is that the covariances of the time-varying variables with η is available. This can provide the researcher a better sense of the properties of these latent time-invariant variables and their pattern of associations. The equality of the coefficients from x_{1t} to y_t is shown by using the same coefficient for each path as is the coefficient of 1 from η to y_t. But within this SEM framework it is straightforward to permit more flexible structures. In our empirical application, the repeated measure (y_t) would be wages, the number of children is a time-varying covariate (x_{1t}) that we initially assume to have the same impact on wages during each period, and all omitted time-invariant variables (e.g., intelligence, motivation, other stable personality traits are combined in η with this latent variable permitted to correlate with x_{1t}.

[FIGURE 2 ABOUT HERE]

Loosening equality restrictions

It is straightforward in a SEM to construct a fixed effects model where the time-varying variables $\mathbf{x}_{it}$ are permitted to have different effects at different times so that we have $\mathbf{B}_{y\mathbf{x}t}$ instead of $\mathbf{B}_{y\mathbf{x}}$ leading to $y_{it} = \mathbf{B}_{y\mathbf{x}t}\mathbf{x}_{it} + \eta_i + \varepsilon_{it}$. Furthermore, we can easily allow the error variances to differ over time and estimate $\sigma^2_{\varepsilon_t}$ where t distinguishes the error variances at the different waves. It is even possible to allow the latent time-invariant variables to have different effects on y_{it} by adding a coefficient, say λ_t to η_i

leading to

$$y_{it} = \mathbf{B}_{y\mathbf{x}t}\mathbf{x}_{it} + \lambda_t \eta_i + \varepsilon_{it}. \tag{7}$$

At least one of the λ_t should be set to a one to scale the latent time-invariant variables. For instance, the impact of personality characteristics or intelligence as latent time-invariant variables might change on an outcome such as depression as a person ages. This could be tested by including λ_t as a coefficient for η_i. Analogous modifications are available for the random effects model.

Lagged effects

Lagged values of the time-varying variables are possible by adding those lagged values to the vector, $\mathbf{x}_{it}$. However, this will lead to the loss of the first wave of data since the lagged value of the time-varying covariate is not available and thus cannot be included. The fixed effects equation could represent this model, though there would be one fewer wave of data. The path diagram would be modified to show these lagged effects.

Lagged endogenous variables for autoregressive effects are also straightforward to include. Here the new equation would be

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \rho y_{it-1} + \eta_i + \varepsilon_{it} \tag{8}$$

where ρ would be the autoregressive effect of y_{it-1} on y_{it}. Here too we would lose one wave of data due to the use of a lagged variable. Furthermore, the first wave y_{i1} should be treated as predetermined and correlated with the time-varying ($\mathbf{x}_{it}$) and latent time-invariant variables (η_i). An added complication to check with lagged endogenous variables is whether there is an autoregressive disturbance. This is particularly problematic if present with a lagged dependent variable since it creates a correlation between the disturbance and explanatory variable. In the SEM approach we can treat this by adding an autoregressive relation between the disturbance term. Figure 3 is a modified version of Figure 2 that includes the lagged endogenous variable. We could further modify this to include an autoregressive disturbance provided we allow an association between ε_{i2} and y_{i1} that would be created by the autocorrelation. As we explain below, using a SEM model we can perform tests of autoregressive dependent variables (or autoregressive disturbances). Similar changes apply to the random effects model.

[FIGURE 3 ABOUT HERE]

Incorporating observed time-invariant variables

The traditional fixed effects model in which the latent time-invariant variable has a constant impact on y over time does not permit time-invariant observed variables ($\mathbf{z}_i$). However, there are special situations where we can introduce time-invariant observed variables easily in the SEM formulation. One such case is if a researcher hypothesizes that the observed time-invariant variable of interest is uncorrelated with the latent time-invariant variable (Allison and Bollen, 1997). Though this can be a difficult assumption to satisfy, keep in mind that it is a weaker assumption than that made with the traditional random effects model. Recall that the random effects model assumes that the latent time-invariant variable (η_i) is uncorrelated with *both* the time-varying variables ($\mathbf{x}_{it}$) and the time-invariant variables ($\mathbf{z}_i$). The hybrid model we are presenting allows η_i to correlate with $\mathbf{x}_{it}$ but sets η_i to be uncorrelated with $\mathbf{z}_i$. Furthermore, we do not need to force all the time-invariant variables to be uncorrelated with η_i, but can allow just one or a few to be uncorrelated. The other correlated observed time-invariant variables are dropped from the model and allowed to be part of the η_i variable. The SEM approach then allows for a hybrid model that stands between the fixed and random effects model with regard to the treatment of observed time-invariant variables.

Suppose that we are unwilling to assume that any of the observed time-invariant variables are uncorrelated with the latent time-invariant variables, but we still have a substantive interest in the impact of sex, race, or other observed time-invariant variables. There is another possibility. We can add an equation to the model where η_i is the dependent variable and the time-invariant observed variables are its predictors. The equations for this model would be

$$
\begin{aligned}
y_{it} &= \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \eta_i + \varepsilon_{it} \\
\eta_i &= \mathbf{B}_{\eta\mathbf{z}}\mathbf{z}_i + \varepsilon_{\eta i}
\end{aligned}
\qquad (9)
$$

where $\mathbf{B}_{\eta\mathbf{z}}$ gives the coefficients of $\mathbf{z}_i$ on η_i and $\varepsilon_{\eta i}$ is a disturbance with a mean of zero and uncorrelated with the covariates. Figure 4 is a path diagram of this model.

FIGURE 4 ABOUT HERE]

This η_i equation permits the time-invariant observed variables to have indirect effects on y_{it} through η_i.[11] In addition, examination of the R^2 of this equation would reveal to the reader how well the observed time-invariant variables predict the latent η_i. A high value suggests that the observed time-invariant are closely associated with the latent ones.

Estimation

A first step in estimating these models in a SEM is preparing the data set. Panel data commonly appears in one of two forms. One is the long form where observations of the same individual are stacked on top of each other. Each row of the data set in say a sample of individuals over several years would be a "person-year." Creating this data structure leads to a "long" data set. The wide form of data has each row referring to a different individual. The variables give the variable values for a particular individual in a particular wave of the data. So if an individual is observed for five years of income, then five income variables are created for that person, one for each year. Because new variables are created to refer to each time wave, this leads to a "wide" data set and hence the term the wide form. Some statistical software have routines that enable easy movement between the long and wide form of panel data (e.g., in Stata, reshape) and this simplifies data preparation. The wide form is most suitable for the SEM approach that we describe.

Though there are a variety of estimators for any SEM, the default and dominant estimator for continuous dependent variables is the maximum likelihood estimator (MLE). The MLE is derived under the assumption that the disturbances (ε_{it}) come from a multivariate normal distribution (Jöreskog 1973; Bollen 1989:126-28). Under these conditions, the coefficients and parameter estimates of the model have the desirable properties of MLE.[12] The appendix gives a more formal presentation of the model and MLE fitting function for SEMs. There is much work in the SEM literature that examines the robustness of the ML estimator to this assumption and it finds conditions where normality is not required for accurate significance tests (e.g., Satorra, 1990).

Furthermore, there are other readily available estimators for SEMs that either do not require normal disturbances or that correct for nonnormality (e.g., Bollen and Stine 1990, 1993; Satorra and Bentler 1994). All of this literature suggests that this distributional assumption is not critical in that there are several options to pursue if it is violated.

Missing data

Attrition or other sources of missing values on variables in panel analysis is common. It is useful to distinguish between data missing completely at random (MCAR) and missing at random (MAR) [Rubin and Little 1987; Schafer 2000]. MCAR suggests that the missing data values are a simple random sample of all data values. MAR is less restrictive. MAR assumes that the probability of an observation being missing can depend on the observed data, but it cannot depend on the missing data. In a SEM there are two options for treating data that are MCAR or MAR. One is the direct MLE approach that allows the variables available for a case to differ across individuals and that estimates the parameters with all of the nonmissing variable information (Arbuckle 1996). The second option is multiple imputation where multiple data sets are imputed, estimated, and their estimates combined. We apply the direct MLE in our application.

Tests of Model Fit

SEMs have tests of overall model fit that we can use to assess the fit of the fixed and random effects as well as the hybrid models that we have described here. To understand these tests, consider the null hypotheses of

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{\mu}(\boldsymbol{\theta}) \\ \boldsymbol{\Sigma} &= \boldsymbol{\Sigma}(\boldsymbol{\theta}) \end{aligned} \qquad (10)$$

where μ and Σ are the means and covariance matrix of the observed variables and $\mu(\theta)$ and $\Sigma(\theta)$ are the model-implied means and covariance matrix of the observed variables. The θ that is part of the model-implied means and covariance consists of the free parameters (e.g., coefficients) of a model. Each model that we specify will have a set of parameters to estimate. In addition, each model specification implies a particular form of $\mu(\theta)$ and $\Sigma(\theta)$ that predicts the means and covariance matrix. See the Appendix for these implied moment matrices for our models. If the model is valid, then having the parameter values will exactly reproduce the means and covariance matrix, that is, μ and Σ, of the observed variables. If the model is incorrect, then $\mu(\theta)$ and $\Sigma(\theta)$ will not exactly reproduce μ and Σ, even in the population. In light of this, the null hypothesis in equation (moment Ho) is a test of the validity of the model. Rejection suggests that the model is incorrect while failure to reject suggests consistency of the model with the data.

The MLE provides a readily available test statistic, say T, that is a likelihood ratio test that asymptotically follows a chi-square distribution with degrees of freedom of $df = \left(\frac{1}{2}\right)P(P+3) - t$ where P is the number of observed variables and t is the number of free parameters estimated in the model. The null hypothesis of this likelihood ratio test is $H_o : \mu = \mu(\theta) \ \& \ \Sigma = \Sigma(\theta)$. This corresponds to the implied moment hypotheses described earlier.

Statistical power raises a complication in the use of the likelihood ratio test. Large samples of several hundred or more cases generally have considerable statistical power to detect even minor misspecifications. In practice, this means that nearly all models will be rejected in a sufficiently large sample and this might be due to errors in specification that most would consider trivial. A direct way of approaching this is to estimate the statistical power of a model and sample (e.g., Satorra and Saris 1985; Matsueda and Bielby 1986).

Another way of assessing fit is with alternative measures of fit that have emerged in the SEM literature. In general it is good practice to report several fit indices along with the chi-square test statistic (T), degrees of freedom, and p-value. Several that we have found useful are reported below:

Baseline fit indices

One family of fit indices in SEMs are baseline fit indices. They are called baseline fit indices because they are measures of fit that compare the fit of the researcher's hypothesized model to a very restrictive alternative model called the baseline model. The logic behind the baseline fit indices is that a researcher's hypothesized model should have clearly superior fit to the baseline model if the specified model is plausible. The fit indices provide a measure of relative fit of the hypothesized to the baseline model.

Other than the chi-square test, the Tucker and Lewis (1973) Index is perhaps the oldest index used in SEM. It is calculated as

$$TLI = \frac{T_b / df_b - T_m / df_m}{T_b / df_b - 1} \tag{11}$$

where T_b $[= (N-1)F_b]$ and T_m $[= (N-1)F_m]$ are chi-square test statistics for a baseline model and the hypothesized model, respectively, and df_b and df_m are the corresponding degrees of freedom. The baseline model is typically a model where all observed variables are uncorrelated and their variances and means are freely estimated. The TLI compares the fit of the hypothesized model to the baseline model while taking account of the degrees of freedom of each model.

The Incremental Fit Index (IFI) [Bollen 1989] is another baseline fit index. Its definition is

$$IFI = \frac{T_b - T_m}{T_b - df_m} \tag{12}$$

where terms are defined the same as described for the TLI . The Relative Noncentrality Fit Index (RNI) [McDonald and Marsh 1990; Bentler 1990] is

$$RNI = \frac{(T_b - df_b) - (T_m - df_m)}{T_b - df_m} \tag{13}$$

Ideal fit for all three of these indices is a value of 1.0. Models with indices below 0.90 are typically deemed unacceptable. Although the exact distributions of these fit indices are unknown, simulation evidence suggests that their means of sampling distributions are stable across sample size except when the model fit is very poor.

Stand-alone fit indices

As the name suggests, stand-alone fit indices do not involve a comparison of a model's fit to a baseline model. Two useful stand-alone fit indices are the Root Mean Square Error of Approximation ($RMSEA$) [Steiger and Lind 1980) and the Bayesian Information Criterion (BIC) [Schwarz 1978]. The formula for the $RMSEA$ is

$$RMSEA = \sqrt{\frac{T_m - df_m}{(N-1)df_m}} \tag{14}$$

An attractive feature of the $RMSEA$ is that it has a confidence interval so that researchers have a better idea of its sampling distribution (Browne and Cudeck 1993). Common standards for the $RMSEA$ are that values less than 0.05 signify a reasonable fit whereas values that exceed 0.10 suggest poor fit. However, recent evidence suggests that these standards might work for large samples, but would be inappropriate for small samples (e.g., $N < 150$).

Schwartz (1978) proposed the BIC as a way to approximate the Bayes factor in comparing statistical models. Raftery (1993 2005) has discussed the use of the BIC in SEM. A useful formula for the BIC is

$$BIC = T_m - df_m \ln(N) \qquad (15)$$

where T_m is the chi-square test statistic, df_m are its degrees of freedom, $\ln(.)$ is the natural log, and N is the sample size. In this form, BIC compares the hypothesized model to the saturated model and negative values favor the hypothesized model whereas positive values favor the saturated model. In general, the lower the BIC, the better is the fit of a model. Jeffrey's (1961) suggests guidelines for interpreting the magnitude of the BIC. We use Raftery's (2005) modification of these so that absolute values of BIC of $0-2$ are weak differences, $2-6$ are positive evidence, $6-10$ are strong evidence, and >10 are very strong evidence.

Comparisons of Models

Another feature of the chi-square likelihood ratio test is that we also can compare nested models. Nested models occur when the parameters of one model can be obtained by placing restrictions on a more general model that encompasses it. Many of the models we describe have this feature. For instance, we can test the traditional random versus fixed effects models by first estimating the random effects model where the latent time-invariant variable is uncorrelated with the covariates and then estimating a separate model identical to this except that it allows the latent time-invariant variable to correlate with the covariates. These models are nested in that the random effects model is a restrictive form of the fixed effects model because the random effects model forces the correlations of the latent time-invariant variable to zero. The chi-square and degrees of freedom of the fixed effects model are subtracted from the corresponding chi-square and degrees of freedom of the random effects model and the resulting test statistic has an asymptotic chi-square with degrees of freedom equal to the difference in degrees of freedom between these models.[13] A nonsignificant chi-square is evidence in support of the random effects model whereas a significant chi-square supports the fixed effects model with covariates and latent time-invariant variable correlated.

This same chi-square difference test allows us to test other nested structures such as the traditional fixed effect model where the coefficient for the same variable at different points of time is equal versus an identical model where these coefficients are allowed to differ over time. Or we could test whether the error variance is equal or unequal over time.

The fit indices described above are also a tool to compare different model structures. We already have mentioned how the BIC with the lowest value indicates the best fit. Differences in these other fit indices might also provide useful information. Though in our experience, the differences in these other fit indices might be more difficult to interpret than the BIC.

Wage Penalty Empirical Example

Previously, we introduced our empirical example, i.e. the wage penalty of motherhood based on the study by Budig and England (2001) using data from the NLSY, and obtained estimates using the traditional fixed and random effects model framework estimated in Stata. In this section, we demonstrate how we can obtain the same coefficients as the traditional fixed and random effects models in the SEM framework; discuss overall fit measures from SEM, unavailable in traditional implementations, which suggest room for improvement of the traditional models; and provide estimates of the wage penalty of motherhood based upon alternative model specifications using the SEM approach that provide better empirical fit to the data.

Random and fixed effects models as SEMs

To estimate models in the SEM framework, we use data in wide format. We estimate all models using Mplus 4.0. We apply the direct MLE (Arbuckle 1996) that estimates the parameters with all of the nonmissing variable information, leaving us with a sample of 5,285 cases. Table 2 provides estimates for several model specifications in the SEM framework. Description of the variables and series of models was provided in section 2.5. The first two columns correspond to the standard random and fixed effects models, but estimated

[Table 2 About Here]

in the SEM framework. Comparing the coefficients to those in Table 1, we see that the estimates are virtually identical. We would expect this since we have programmed the SEM formulations to match the traditional versions of these models. Hence we can replicate the results of the traditional models using SEM. However, the SEM results provide additional information by way of the measures of model fit.

The model fit statistics that we include are the Likelihood Ratio (LR) test statistic (T_m), degrees of freedom (df), IFI, RNI, RMSEA, and BIC. We described the calculation of these fit indices above. In Table 2, the chi-square LR test statistic that compares the hypothesized fixed or random effects model to the saturated model leads to a highly statistically significant result, suggesting that these hypothesized models do not exactly reproduce the means and covariance matrix of the observed variables. With over 5,000 cases in this sample, the LR chi-square test has considerable statistical power to detect even small departures of the model from the data. In light of this statistical power, it is not surprising that the null hypotheses are rejected for these models ($p < 0.001$).

The supplemental fit indices provide an additional means by which to assess model fit. Both the random and fixed effects models for the model with no controls (Model 1) and the model that controls for marital status (Model 2) have values of IFI and RNI exceeding 0.90, a common cutoff value. However, values of RMSEA are often greater than 0.05 and values of BIC are positive, indicating problems with model fit for Models 1 and 2. In contrast, Model 3 that further controls for human capital variables have values of RNI and IFI close to 1 and values less than 0.05 of RMSEA and large negative values of BIC, all indicating good model fit. Thus, the fit statistics from the SEM results support the choice of Model 3 over Model 1 or 2 whether we use the random or fixed effects versions.

As explained previously, the random effects model is a restricted form of the fixed effects model where in the former, the latent time-invariant variables (η_i) are uncorrelated with all other covariates. Correlations are allowed in the fixed effects model. Given this nesting, we can form a LR chi-square difference test to compare the fixed and random effects model by subtracting T_m test statistics for the random vs. fixed effects model, taking a difference in their respective degrees of freedom, and comparing the results to a chi-square distribution. A statistically significant chi-square is evidence that favors the fixed effects model while a nonsignificant chi-square favors the random effects model. Performing these chi-square difference tests consistently leads to a statistically significant result lending support to the fixed effects versions of Models 1, 2, and 3 in Table 2. These results are consistent with the Hausman test in favoring the fixed effects model. However, the large sample size combined with the large degrees of freedom for these models complicates the picture in that the statistical power of all these tests is high and does not tell us the magnitude of the differences. The RNI and IFI (i.e. the baseline fit indices) mostly differ in the third decimal place and values as close as these are generally treated as essentially equivalent. The RMSEA has slightly larger differences in the pairwise comparisons with a tendency to favor the random effects models. The greatest separation in the pairwise comparisons for Models 1, 2, and 3 occur for the BIC and the BIC favors the random effects versions of the models.

These results are interesting in that they imply that the random and fixed effects models are closer in fit than the Hausman test or the LR chi-square tests suggested. One reason is that the random effects models have considerably more degrees of freedom than the fixed effects model since the random effects model is forcing to zero all of the covariances of the latent time-invariant variable with the time-varying observed covariates. The BIC gives considerable weight to the degrees of freedom of the model and the greater degrees of freedom contributes to making the BIC more favorable towards the random effects model. A second related reason for the random effects models appearing more competitive is that the magnitudes of the estimated covariances between the time-varying covariates and the latent variable are not always large in the fixed effects model. Table 3 provides a sample of the estimated covariances between the full set of time-varying covariates and η_i in the fixed effects version of Model 3. We provide only the first and last years' covariances and correlations for simplicity, but note that we actually estimate all 12 years of covariances. These results show that some of the correlations (covariances)

[Table 3 About Here]

of the latent time-invariant variable (η_i) and the time-varying xs are not statistically significantly different from zero even though the significance tests are based on an N greater than 5000. Interestingly, Number of children, the key explanatory variable, is essentially uncorrelated with η_i as is Divorced, and Part-time seniority. The other statistically significant correlations are often modest in magnitude (e.g., Married and η_i correlate less than 0.1). Thus our results suggest a more nuanced picture of the association between η_i and the xs than suggested by the random or fixed effects model. The latent time-invariant variable has a statistically significant association with some xs, but not with others. Even the statistically significant ones are modest in magnitude (e.g., ≤ 0.2). So the modest and sometimes not statistically significant covariances of η_i and the xs seems to be the reason that the random and fixed effects versions of the model have fits that are so close as gauged by our fit indices.

As we have noted, these fit indices and the estimates of the covariances of the latent time-invariant variables and the observed time-varying covariates are not available with traditional random and fixed effects statistical routines; the Hausman test indicated, as is often the case, that the fixed effects models are unambiguously superior to the random effects models. Our results present a more nuanced view of their relative fit, while at the same time the LR chi-square tests suggests that neither model exactly fits the data.

Alternative models as SEMs

The results in the prior section provide clear evidence of the superiority of Model 3 compared to Models 1 and 2. While Model 3 fits the data fairly well, the LR chi-square test suggests the potential for improvement in fit. In this section, we explore other specifications of Model 3 to illustrate the SEM approach. As we discussed in section 3.1.1, an alternative SEM specification is to allow the time-varying variables to have different effects at different times. If we had specific hypotheses concerning which variables are the most likely to vary over time, then we could estimate Model 3 freeing only those coefficients and compare the fit of this new model to the fixed and random effects versions of the same model where the coefficients of the same variable are set equal over time. The traditional fixed and random effects models would be nested in the models where the coefficients are free to vary over time. A LR chi-square difference test along with a comparison of fit indices could be made to determine relative fit.

In our case, we do not have specific hypotheses on which variable's coefficients might differ over time. Therefore, we compare the traditional fixed and random effects versions of Model 3 to versions where *all* the coefficients of the time-varying variables are free to differ over time. Table 4 contains the fit statistics. Starting with the LR test we find a chi-square difference for the nested fixed effects models of 1014.21 (=7597.55-

6583.34) with degrees of freedom (*df*) of 110 (=1386-1276) and a *p-value* <0.001 and for the random effects models we have a chi-square difference of 995.58 (=8118.40-7122.82) with *df* of 110 (=1506-1396) and a *p-value* <0.001. The chi-square difference tests are both statistically significant, supporting the model where at least some of the coefficients for the time-varying variables vary by wave. However, considering the large N and accompanying statistical power of the test it is worthwhile to examine the other fit statistics as well. The IFI, RNI, and RMSEA suggest only slight differences between models that do and do not allow the coefficients to vary over time. . The BIC comparisons support the conclusion that we prefer the models that allow the coefficients to vary over time. The difference in BICs between the fixed effects models is 71.22 and between the random effects models is 52.59. Taken together the evidence tends to favor the models that allows the coefficients to vary over time.

[Table 4 About Here]

We next allow the error variances to vary over time. The chi-square difference for the nested fixed effects models is 109.31 with 11 degrees of freedom and a *p-value* <0.001 and for the random effects models we have a chi-square difference of 108.23 with 11 *df* and a *p-value* <0.001. The chi-square difference tests are again both statistically significant, supporting the model where the error variances are allowed to vary over time. The IFI, RNI, and RMSEA suggest only slight differences between models that do and do not allow the error variances to vary over time. The BIC comparisons support the conclusion that we prefer the models that allow the coefficients to vary over time. Once again, most of the fit indices tend to favor the models that allow the coefficients and error variances to vary over time.

Figure 5 plots the effects of number of children on log wages for Model 3, respectively. The x-axis indicates year and the y-axis indicates the child effect on women's wages. The y-axis is in reverse order such that higher values indicate a larger wage penalty for motherhood. We plot the results for several model specifications. To provide a benchmark comparison, we first plot the effect estimates for the conventional random and fixed effects models, i.e. those that constrain the child coefficients to be equal over time. We also plot the effect estimates for models that allow all coefficients and error variances to vary over time, i.e. those models that produce distinct estimates corresponding to each year. In contrast to models that constrain coefficients to be equal over time, models that allow variation over time show smaller estimated effects of motherhood on wages and ones that oscillate over time.[14]

Lagged endogenous variable models

The last subsection suggested that the models that allow the coefficients and error variances to vary over time were the best fitting models. However, even these models might be improved. One possibility we have not considered is having the lagged value of wages as a determinant of current wages. Substantively, including such an effect makes sense in that there is a certain inertia in wages where last year's wages are likely to be a good predictor of this year's. Though raises typically occur, there is a high degree of stability in relative wages across individuals from year to year. As we described in section 3.1.2, lagged endogenous variables for autoregressive effects are also straightforward to include. Our appendix provides a more formal presentation of the SEM setup and assumptions for estimating such a model. We lose one wave of data by specifying such a model; thus, our hypothesized models will be compared to different saturated and baseline models than those above. Table 4 provides fit statistics for the best fitting versions of Model 3 from the last subsection, but adding a lagged endogenous variable.

We estimate a series of specifications for the lagged endogenous variable models, parallel to those we estimate above, freeing constraints to determine whether our model fit improves by loosening these restrictions. First, we allow the lagged wage and all other coefficients to vary over time. Then we allow the error variances to be freed. We plot the values for the specifications where the child coefficient is freed in Figure 5. As before, the effect of number of children on women's wages appears to oscillate over time.

The overall fit statistics of the model in Table 4 enable us to compare the models. The large sample size and accompanying high statistical power lead all the LR chi-square tests to be statistically significant as expected. However, the other fits statistics indicate that all the models with lagged endogenous variables fit very well. For instance, the RNI and IFI are consistently close to 1.00 and the RMSEA is considerably lower than the usual cutoff of 0.05. The BIC always takes large negative values supporting the selection of any of these models over the saturated model.

The LR chi-square nested test of the fixed effects vs. the random effects model consistently finds a *statistically* significant difference in fit that favors the fixed effects versions. But as we saw above in the models without the lagged endogenous variable, most of the covariances of the latent time-invariant variable with the observed covariates are *substantively* near zero. This helps to explain why the BIC consistently chooses the random effects over the fixed effects models. In fact, for these lagged endogenous variable models the BIC most favors the random effects model with equality constraints on the coefficients and error variances.

Taken together we reach the following conclusions on overall fit of the models. First, the lagged endogenous variable models have very good fit.[15] Second, the IFI, RNI, and RMSEA do not reveal large differences among the different versions of these models. Third, the BIC suggests that the most parsimonious of these models, that is, the random effects model with equality constraints, is the best. Regardless of the model, these lagged endogenous variable models suggest a smaller penalty for motherhood's direct effect, particularly in later years, than that suggested by random and fixed effects models without lagged endogenous variables. However, given the lagged endogenous variable there are additional lagged effects of motherhood on wages. For instance, the number of children in say, 1983, has a direct effect on wages in 1983, but it also has an indirect effect on wages in 1984 given the impact of 1983 wages on 1984 wages. This distinction between direct, indirect, and total effects is well-known in the SEM literature (e.g., Sobel 1982; Bollen 1987) and most SEM software permits its exploration.

Other alternative models

The preceding results illustrate a few of the alternative fixed and random effects models that we can easily explore with a SEM approach, but there are more. For instance, we could estimate models in which some of the time-varying covariates were allowed to correlate with the latent time-invariant covariate while others were not. Or, we could test whether autoregressive disturbances were present. Latent covariates to measure intelligence, motivation, or other potential determinants of wages would be easy to include if we had the proper indicators. Even without these extensions our SEM models found evidence that the standard assumptions of fixed coefficients, fixed error variances, and no lagged endogenous variables were not always supported when tested in our empirical example.

CONCLUSION

Classic random and fixed effects panel models are common models applied in *ASR* and elsewhere in sociology. In this paper, we show that these models are a restrictive form of a SEM. When placed in a SEM framework, a researcher does not need to maintain these constraints but can test them and only impose those supported by the data. In addition, a wide variety of additional models and formulations are possible without much difficulty. For instance, a researcher can test whether a covariate's impact on the repeated measure stays the same across all waves of data; test whether the error variances should be allowed to vary over time; include lagged covariates or lagged dependent variables; estimate the magnitude of the covariance of the latent time-invariant variables with the observed time-varying covariates; and include observed time-invariant variables in a fixed effects model either as uncorrelated with the latent time-invariant variable or as a determinant of the latent variable. For these and other models that we discussed, we have useful tests of model fit and fit indices that are not part of the classic random and fixed effects models. Indeed it would be interesting to know how many of the fixed or random effects models that have appeared in the literature would have adequate model fit if some of these tools were applied. We also have a likelihood ratio test of the fixed vs. random effects model and a variety of fit indices as alternatives to or supplements to the Hausman Test.

Our empirical example of the impact of the number of children on women's wages illustrated some of the advantages that flow by casting fixed and random effects panel models as SEMs. For one thing, we had access to a more complete set of model fit statistics that revealed flaws in both the classic fixed and random effects models that were not evident in the usual approaches and the publications based on them. Specifically, neither model fully reproduced the covariance matrix and means of the observed variables as they should if the models were correct. Furthermore, we found evidence that the random effects models were more competitive than the Hausman test alone revealed. In fact, the Hausman test from past research unambiguously supported the fixed effects model over the random effects model. The primary distinction between the fixed and random effects models is whether the covariates correlate with the latent time-invariant variable. With the SEM approach we saw that the correlations of the covariates with the latent time invariant variable were close to zero-- information unavailable with usual methods. Furthermore, the SEM approach suggested that the impact of number of children on wages was not the same across all years and that the unexplained variances were not constant over time in all models. The classic fixed and random effects models assume constant effects regardless of the year of the panel data.

A further departure from the published models for these data was that we looked at whether lagged wages impacted current wages net of the other determinants. Given the degree to which current salary is closely tied to past salary, this is a substantively plausible effect and it was easy to explore in a SEM. We found strong evidence that the lagged endogenous variable models were superior to the models without them. A related substantive point is that these models show the importance of prior wages on current wages and this implies that any variables that impact wages in a given year have an indirect effect on later years as well. Thus, the number of children has direct as well as indirect effects on mothers' wages.

The empirical example did not exhaust the types of models for panel data that could be applied using SEM. For instance, we explained a couple of ways in which we could include observed time-invariant variables into a fixed effects model. Or, it would be straightforward to develop a model that permits the dependent variable to be latent with several indicators and to have a fixed or random effects-like model for it. We could allow for measurement error in the time-varying or time-invariant covariates and include them in the model. In addition, latent curve models or Autoregressive Latent Trajectory (ALT) models might be applied to panel data (Bollen and Curran, 2006). In brief, researchers can build a broader range of models than is commonly applied, some of which might better capture the theory that they wish to test. In addition, specialized software packages for panel data are not required since these models are estimable in standard SEM software.

Although the SEM approach offers considerable flexibility, it as well as the classic fixed and random effects models do not adequately handle all situations that researchers might encounter. For instance, if the latent time-invariant variable has a different correlation with the covariates for different individuals, these models will not work. Similarly, the models we treat permit the covariate's effect on the repeated measure to differ over time, but the models we consider assume that these coefficients are constant over individuals. Some SEM publications have shown fixed and random effects models for cross-sectional data with clusters such as families (e.g., Teachman et al. 2001) or for multilevel models (Bauer 2003). Thus, although we have expanded on the usual fixed and random effects alternatives for analyzing panel data, we have not exhausted the possibilities.

---

[1]Lagged effects of covariates and the dependent variable could be included in the usual implementations of fixed and random effect models by treating them as additional covariates. In practice, this is rarely done and only contemporaneous effects are considered.

[2]Some prior work has drawn connections between SEM and these models. An unpublished paper by Allison and Bollen (1997) and a SAS publication by Allison (2005) discuss SEM set-ups of the classic fixed and random effects model. Teachman, Duncan, Yeung, and Levy (2001) look at fixed effect models in SEM, but concentrate their discussion and example on cross sectional data with clusters of families. Finally, Ejrnaes and Holm (2006) look at different types of fixed effects estimators in panel data models, but do not cover random effects models, lagged dependent variable models, or some of the other variants that we include here. However, they do look at models where the covariance of the time-varying covariates and the latent time invariant variable differs by cases, an issue that we do not consider.

[3]Though we use the term "individual" to refer to a case, the cases do not have to be individual people. They could be groups, organizations, nations, etc.

[4]The weight matrix has the following form:

$$\mathbf{\Omega}_i = \begin{pmatrix} \sigma_\varepsilon^2 + \sigma_\eta^2 & \sigma_\eta^2 & \cdots & \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\varepsilon^2 + \sigma_\eta^2 & \cdots & \sigma_\eta^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\eta^2 & \sigma_\eta^2 & \cdots & \sigma_\varepsilon^2 + \sigma_\eta^2 \end{pmatrix}$$

with

$$\mathbf{\Omega} = \begin{pmatrix} \mathbf{\Omega}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Omega}_N \end{pmatrix}$$

. An estimated version of $\mathbf{\Omega}$ is then used in a GLS procedure.

[5]For this test, $H_o \ : \ \sigma_\eta^2 = 0$ and $H_1 \ : \ \sigma_\eta^2 \neq 0$ A Lagrangian Multiplier test is formed as

$$LM = \frac{NT}{2(T-1)} \left[\frac{\mathbf{e}'\mathbf{D}\mathbf{D}'\mathbf{e}}{\mathbf{e}'\mathbf{e}} - 1 \right]^2$$

where $\mathbf{e}$ is from the regression of y_{it} on $\mathbf{x}_{it}$ and $\mathbf{z}_i$. The

$$\mathbf{D} = \left[\mathbf{d}_1 \quad \mathbf{d}_2 \quad \mathbf{d}_3 \quad \cdots \quad \mathbf{d}_N \right]$$

where $\mathbf{d}_i$ is a dummy variable indicating the i th case. The test statistic asymptotically follows a chi-square distribution with one degree of freedom (see Greene 1997, Chapter 14).

[6]Budig and England's (2001) analysis resulted in a final sample of 5,287 women.

[7]Budig and England also examined the wage penalty with three dichotomous measures indicating one child, two children, and three or more children. They find that the effects are monotonic, although not perfectly linear, and prefer the continuous indicator of number of children for all other analyses.

[8]We thank Michelle Budig for sharing with us the experience and seniority variables from the Budig and England (2001) analysis.

[9]These slight differences reflect the difficulty when attempting to replicate all the decisions made in variable construction of an analysis reported in a research article.

[10]A possible source of confusion is that path analysis is sometimes used to refer just to recursive or nonrecursive models of only observed variables where measurement error and latent variables are not considered. This is an inaccurate usage of the term. In fact, Sewall Wright the inventor of path analysis considered latent variables as do the more contemporary users of path analysis.

[11]In this formulation the indirect effects of $\mathbf{z}_i$ would equal their direct effects on η_i since η_i's effect on y_{it} is 1.

[12]MLEs are consistent, asymptotically unbiased, asymptotically normally distributed, asymptotically efficient among asymptotically unbiased estimators, and the inverse of the expected information matrix is available to estimate the asymptotic covariance matrix of the parameter estimator that we use for significance testing.

[13]Strictly speaking, the least restrictive model should be a good fitting model for the chi square distribution of the differences to hold.

[14]Freeing of the child coefficient for Models 1 and 2 demonstrate an upward trend over time in the wage penalty for motherhood that levels off in 1988 when women are

between 23 and 31 years old, suggesting a form of cumulative disadvantage associated with motherhood on women's wages. When we control for women's human capital in Model 3, the penalty trend is no longer steeply increasing, suggesting that the increasing wage penalty for motherhood over time is at least partly a function of the differential human capital reflective of the number of children women bear.

[15]This statement recognizes the role that the large sample size plays in elevating statistical power for the LR chi-square tests and hence downplays the statistical significance of these tests. However, the chi-square tests mean that these is still room for improvement despite the favorable results with all other fit indices for these models.

REFERENCES

Alderson, Arthur S. 1999. "Explaining Deindustrialization: Globalization, Failure, or Success?" *American Sociological Review* 64:701-721.

Alderson, Arthur S. and Francois Nielsen. 1999. "Income Inequality, Development, and Dependence: A Reconsideration." *American Sociological Review* 64:606-631.

Allison, Paul D. 2005. *Fixed Effects Regression Methods for Longitudinal Data Using SAS.* Cary, NC:SAS Institute.

Allison, Paul D. 1994. "Using Panel Data to Estimate the Effects of Events." *Sociological Methods and Research* 23:179-199.

Allison, Paul D. and Kenneth A. Bollen. 1997. "Change Score, Fixed Effects and Random Component Models: A Structural Equation Approach." Paper presented at the Annual Meetings of the American Sociological Association.

Baltagi, Badi H. 2001. *Econometric Analysis of Panel Data.* NY:Wiley.

Bauer, Daniel J. 2003. "Estimating Multilevel Linear Models as Structural Equation Models." *Journal of Educational and Behavioral Statistics* 28:135-167.

Beckfield, Jason. 2006. "European Integration and Income Inequality." *American Sociological Review* 71:964-985.

Bentler, Peter M. 1990. "Comparative Fit Indexes in Structural Models." *Psychometrika* 107:238-46.

Bollen, Kenneth A. 1987. "Total, Direct, and Indirect Effects in Structural Equation Models." *Sociological Methodology* 17:115-140.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables.* NY:Wiley.

Bollen, Kenneth A. and Patrick J. Curran. 2006. *Latent Curve Models: A Structural*

Equation Perspective. NY: Wiley.

Bollen, Kenneth A. and Robert Stine. 1992. "Bootstrapping Goodness-of-fit Measures in Structural Equation Models." *Sociological Methods and Research* 21:205-229.

Bollen, Kenneth A. and Robert Stine. 1990. "Direct and Indirect Effects: Classical and Bootstrap Estimates of Variability." *Sociological Methodology 20:*115-140.

Budig, Michelle J, and Paula England. 2001. "The Wage Penalty for Motherhood." *American Sociological Review* 66:204-225.

Conley, Dalton and Neil G. Bennett. 2000. "Is Biology Destiny? Birth Weight and Life Chances." *American Sociological Review* 65:37-69.

Erjrnaes, Mette and Anders Holm. 2006. "Comparing Fixed Effects and Covariance Structure Estimators for Panel Data. *Sociological Methods and Research* 35:61-83.

Greeene, William H. 1997. *Econometric Analysis.* 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Halaby, Charles N. 2004. "Panel Models in Sociological Research: Theory into Practice." *Annual Review of Sociology* 30:507-44.

Hausman, Jerry A. 1978. "Specification Tests in Econometrics." *Econometrica* 46:1251-1272.

Hsiao, Cheng. 2003. *Analysis of Panel Data*. Second edition. Cambridge: Cambridge University Press.

Jeffreys, Harold. 1961. *Theory of Probability.* 3rd edition. Oxford:Oxford University Press.

Jöreskog, Karl G. 1973. "A General Method for Estimating a Linear Structural Equation

System." Pp. 85-122 in *Structural Equation Models in the Social Sciences*, edited by A. S. Goldberger and O. D. Duncan. NY:Academic Press.

Little, Roderick J. A. and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data.* NY: Wiley.

Matsueda, Ross L. and William T. Bielby. 1986. "Statistical Power in Covariance Structure Models." *Sociological Methodology* 16:120-158.

Matsueda, Ross L., Derek A. Kreager, and David Huizinga. 2006. "Deterring Delinquents: A Rational Choice Model of Theft and Violence." *American Sociological Review* 71:95-122.

McDonald, Roderick P. and Herbert W. Marsh. 1990. "Choosing a Multivariate Model: Noncentrality and Goodness of Fit." *Psychological Bulletin* 107:247-255.

Mouw, Ted. 2000. "Job Relocation and the Racial Gap in Unemployment in Detroit and Chicago, 1980 to 1990." *American Sociological Review* 65:730-753.

Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46:69-85.

Raftery, Adrian E. 1995. Bayesian Model Selection in Social Research." *Sociological Methodology 25:*111-63.

Raftery, Adrian E. 1993. "Bayesian Model Selection in Structural Equation Models." Pp. 163-180 in *Testing Structural Equation Models*, edited by K. A. Bollen and J. S. Long. Newbury, CA: Sage.

Satorra, Albert and Willem Saris. 1985. "Power of the Likelihood Ratio Test in Covariance Structure Analysis." *Psychometrika* 50:83-90.

Schafer, Joseph L. 2000. *Analysis of Incomplete Multivariate Data.* Boca Raton, FL:

Chapman & Hall/CRC.

Schwarz, Gideon. 1978. "Estimating the Dimensions of a Model." *Annals of Statistics* 6:461-464.

Sobel, Michael. 1986. "Some New Results on Indirect Effects and their Standard Errors in Covariance Structure Models." *Sociological Methodology* 16:159-186.

Steiger, James H. and John C. Lind. 1980. "Statistically Based Tests for the Number of Common Factors." Presented at the Psychometric Society, Iowa City, IA.

Teachman, Jay, Greg J. Duncan, W. Jean Yeung, and Dan Levy. 2001. "Covariance Structure Models for Fixed and Random Effects." *Sociological Methods and Research* 30:242-70.

Tucker, Ledyard R. and Charles Lewis. 1973. "A Reliability Coefficient for Maximum Likelihood Factor Analysis." *Psychometrika* 38:1-10.

Wheaton, Blair and Phillipa Clarke. 2003. "Space Meets Time: Integrating Temporal and Contextual Influences on Mental Health in Early Adulthood." *American Sociological Review* 68:680-706.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Yakubovich, Valery. 2005. "Weak Ties, Information, and Influence: How Workers Find Jobs in a Local Russian Labor Market." *American Sociological Review* 70:408-421.

TABLES

Table 1. Coefficients for the Effect of Total Number of Children (Continuous Variable) on Women's Log Hourly Wage (Stata): NLSY 1982 to 1993

Control Variables in Model	Random Effects Model		Fixed Effects Model	
Model 1: Gross (no controls)	-0.082 (0.003)	***	-0.068 (0.004)	***
Model 2: Marital status	-0.089 (0.003)	***	-0.072 (0.004)	***
Model 3: Marital status and human capital variables	-0.037 (0.003)	***	-0.043 (0.004)	***

Notes: Numbers in parentheses are standard errors. Measures of human capital include education, full-time seniority, part-time seniority, full-time experience, part-time experience, number of breaks in employment, and whether currently enrolled in school. Number of observations = 41757; Number of groups = 5285;
*p<.05 **p<.01 ***p<.001 (two-tailed tests)

Table 2. SEM Coefficients for the Effect of Total Number of Children (Continuous Variable, Constrained to be Equal) on Women's Log Hourly Wage (Mplus): NLSY 1982 to 1993

Control Variables in Model	Random Effects Model		Fixed Effects Model	
Model 1: Gross (no controls)	-0.082	***	-0.068	***
	(0.003)		(0.004)	
T_m (LR chi-square)	6721.31		6647.85	
df	219		207	
IFI	0.9316		0.9323	
RNI	0.9316		0.9322	
RMSEA	0.0750		0.0767	
BIC	4843.91		4873.32	
Model 2: Marital status	-0.089	***	-0.072	***
	(0.003)		(0.004)	
T_m (LR chi-square)	7055.17		6903.80	
df	505		469	
IFI	0.9669		0.9675	
RNI	0.9668		0.9673	
RMSEA	0.0495		0.0510	
BIC	2725.99		2883.24	
Model 3: Marital status and human capital variables	-0.037	***	-0.042	***
	(0.003)		(0.004)	
T_m (LR chi-square)	8118.40		7597.55	
df	1506		1386	
IFI	0.9922		0.9927	
RNI	0.9921		0.9926	
RMSEA	0.0288		0.0291	
BIC	-4791.98		-4284.11	

Notes: Numbers in parentheses are standard errors. Measures of human capital include education, full-time seniority, part-time seniority, full-time experience, part-time experience, number of breaks in employment, and whether currently enrolled in school. Number of observations = 41757; Number of groups = 5285;
*p<.05 **p<.01 ***p<.001 (two-tailed tests)

Table 3. SEM Coefficients for Fixed Effects Model 3

Control Variables	1982 Covariance with η		1982 Correlation with η	…	1993 Covariance with η		1993 Correlation with η
	-	-		-		-	
Number of children	0.002		0.011		0.001		0.003
	(0.004)				(0.007)		
Married	0.002		0.017		0.007	**	0.052
	(0.002)				(0.002)		
Divorced	0.002		0.031		-0.002		-0.018
	(0.001)				(0.002)		
Educational Attainment*	0.030	*	0.062		0.038	*	0.061
	(0.014)				(0.019)		
Currently enrolled in school*	0.001		0.008		0.002	*	0.035
	(0.002)				(0.001)		
Full-time seniority*	0.000	***	0.000		0.140	***	0.141
	(0.000)				(0.017)		
Part-time seniority*	0.000		0.000		0.011		0.029
	(0.000)				(0.007)		
Full-time experience*	0.040	***	0.113		0.218	***	0.202
	(0.006)				(0.022)		
Part-time experience	0.016	*	0.053		-0.038	*	-0.051
	(0.006)				(0.015)		
Num. of employment breaks	0.027	***	0.092		-0.093	***	-0.148
	(0.006)				(0.013)		

Notes: Numbers in parentheses are standard errors.
*p<.05 **p<.01 ***p<.001 (two-tailed tests)

Table 4. Model Fit Statistics (N=5285)

Control Variables in Model	**Model Specification**	**Log Lik.**	**T_m (LR chi-square)**	***df***	**IFI**	**RNI**	**RMSEA**	**BIC**
	Saturated for BIC	-186947.29		0				
	Baseline model for CFI, IFI	-611001.12	848108	8646				
Model 3: Marital status and human capital variables	Random Effects	-191006.49	8118.40	1506	0.9922	0.9921	0.0288	-4791.98
	Fixed Effects	-190746.07	7597.55	1386	0.9927	0.9926	0.0291	-4284.11
	Random Effects, All Coef. Freed	-190508.70	7122.82	1396	0.9932	0.9932	0.0279	-4844.57
	Fixed Effects, All Coef. Freed	-190238.97	6583.34	1276	0.9937	0.9937	0.0281	-4355.33
	Random Effects, All Coef. & Error Var. Freed	-190454.59	7014.59	1385	0.9934	0.9933	0.0277	-4858.50
	Fixed Effects, All Coef. & Error Var. Freed	-190184.31	6474.03	1265	0.9938	0.9938	0.0279	-4370.35
	Saturated for BIC, Lagged End. Var. Models	-204457.31		0				
	Baseline model for CFI, IFI, Lagged End. Var. Models	-599267.35	789620	7381				
	RE, Lagged End. Var.	-206204.46	3494.31	1273	0.9972	0.9972	0.0182	-7418.65
	FE, Lagged End. Var.	-205918.25	2921.89	1163	0.9978	0.9978	0.0169	-7048.07

RE, Lagged End. Var., All Coef. Freed	-205905.22	2895.82	1163	0.9978	0.9978	0.0168	-7074.15
FE, Lagged End. Var., All Coef. Freed	-205633.45	2352.29	1053	0.9984	0.9983	0.0153	-6674.69
RE, Lagged End. Var. Model, All Coef. & Error Freed	-205872.79	2830.97	1153	0.9979	0.9979	0.0166	-7053.27
FE, Lagged End. Var. Model, All Coef. & Error Freed	-205601.95	2289.29	1043	0.9984	0.9984	0.0150	-6651.96

FIGURES

Figure 1. Classic Random Effects Model in Path Diagram

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \mathbf{B}_{y\mathbf{z}}\mathbf{z}_i + \eta_i + \varepsilon_{it}$$

$$E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$$

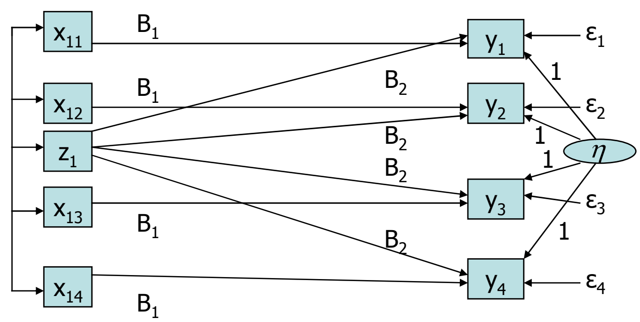


Figure 2. Classic Fixed Effects Model in Path Diagram

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \eta_i + \varepsilon_{it}$$

$$E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$$

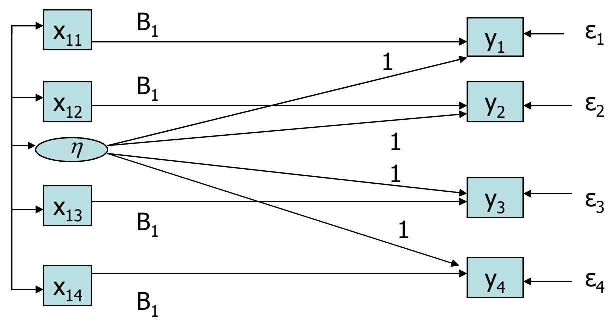


Figure 3. Fixed Effects Model with Lagged Dependent Variables

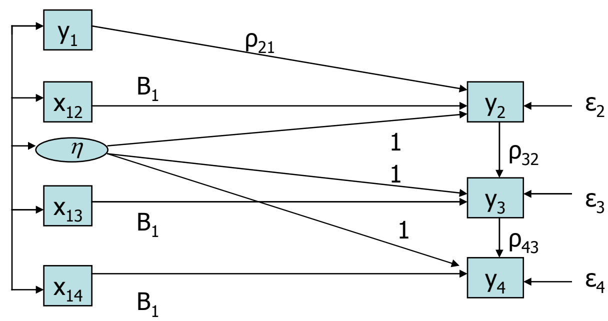


Figure 4. Fixed Effects Model with Observed Influencing Latent Time Invariant variable

$$y_{it} = \mathbf{B}_{y\mathbf{x}}\mathbf{x}_{it} + \eta_i + \varepsilon_{it} \qquad \eta_i = \mathbf{B}_{\eta\mathbf{z}}\mathbf{z}_i + \varepsilon_{\eta i}$$

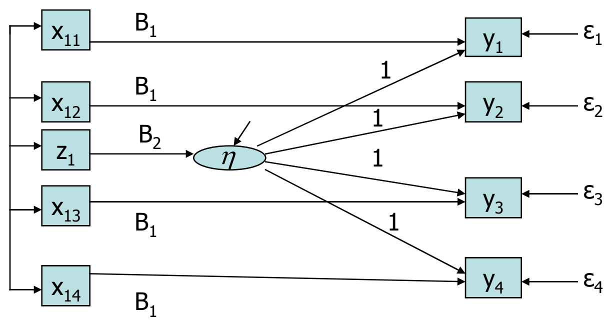
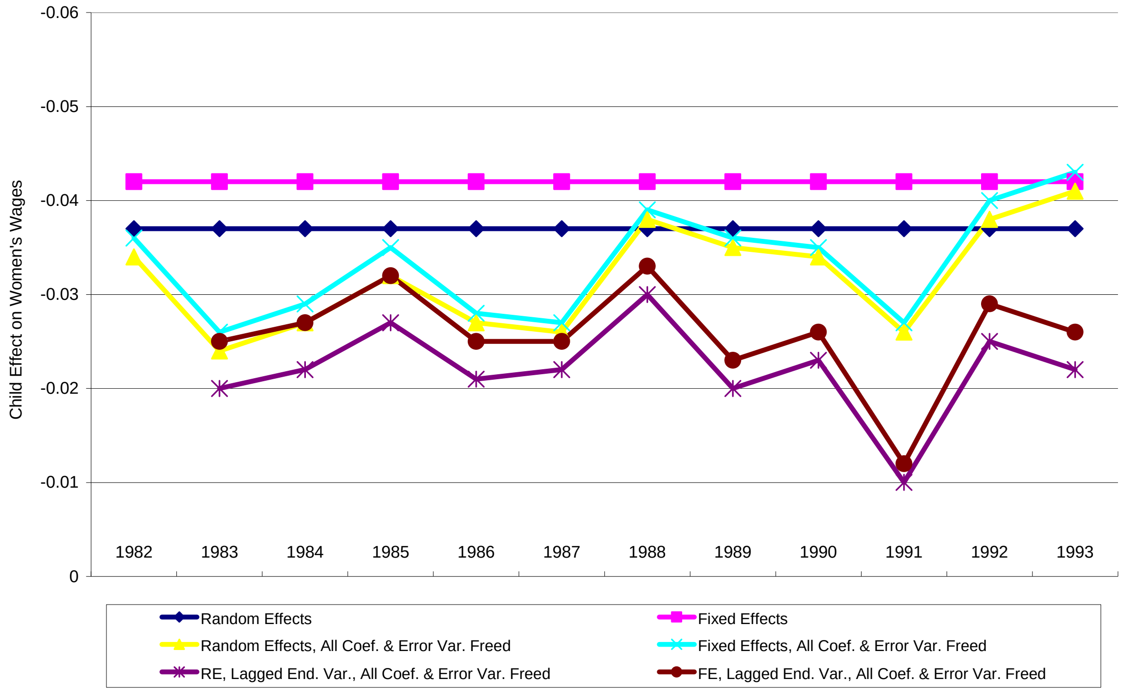


Figure 5. SEM Coefficients for the Effect of Total Number of Children on Women's Log Hourly Wage: Model 3, Marital Status and Human Capital Variables



APPENDIX

Fixed and Random Effects Models as Structural Equation Models

Classic Fixed and Random Effects Models

We represent the classic fixed and random effects models in the following matrix equation:

$$\mathbf{y}_i = \boldsymbol{\alpha} + \boldsymbol{\Gamma}\mathbf{w}_i + \boldsymbol{\varepsilon}_i \qquad \text{(A1)}$$

where

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} \quad \mathbf{w}_i = \begin{bmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \\ \vdots \\ \mathbf{x}_{iT} \\ \mathbf{z}_i \\ \eta_i \end{bmatrix} \quad \boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix} \qquad \text{(A2)}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_T \end{bmatrix} \quad \boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{B}_{y_1\mathbf{x}_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{y_1\mathbf{z}} & 1 \\ \mathbf{0} & \mathbf{B}_{y_2\mathbf{x}_2} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{y_2\mathbf{z}} & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{y_3\mathbf{x}_3} & \cdots & \mathbf{0} & \mathbf{B}_{y_3\mathbf{z}} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{y_T\mathbf{x}_T} & \mathbf{B}_{y_T\mathbf{z}} & 1 \end{bmatrix}$$

The $\mathbf{x}_{it}$ vector contains the values of the time varying covariates for the ith case at the tth time, $\mathbf{z}_i$ is the vector of observed time-invariant variables for the ith case, and η_i is the latent time-invariant variable for the ith case. We assume that the mean of the disturbance is zero [$E(\boldsymbol{\varepsilon}_i) = \mathbf{0}$ for all i], that they are not autorcorrelated over cases[$COV(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j^{'}) = \mathbf{0}$ for $i \neq j$], and that the covariance of the disturbance with the covariates in $\mathbf{w}_i$ is zero [$COV(\mathbf{w}_i, \boldsymbol{\varepsilon}_j^{'}) = \mathbf{0}$ for all i, j].

In SEMs the vector of means ($\boldsymbol{\mu}$) and the covariance matrix ($\boldsymbol{\Sigma}$) of the observed variables are functions of the parameters of the researcher's model. If we place all model parameters (coefficients, intercepts, variances, covariances) in a vector $\boldsymbol{\theta}$, then these implied functions are the model *implied covariance matrix* ($\boldsymbol{\Sigma}(\boldsymbol{\theta})$) and *implied mean vector* [$\boldsymbol{\mu}(\boldsymbol{\theta})$]. When the model is valid, then

$$H_o \ : \ \boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}) \ \& \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}) \tag{A3}$$

That is, we will exactly reproduce the means and covariance matrix of the observed variables by knowing the model parameter values and substituting them into the implied mean vector and implied covariance matrix. For equation (A1), the implied mean vector [$\boldsymbol{\mu}(\boldsymbol{\theta})$] is

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\mu}_{\mathbf{w}} \\ \boldsymbol{\mu}_{\mathbf{w}} \end{bmatrix} \tag{A4}$$

and the implied covariance matrix [$\boldsymbol{\Sigma}(\boldsymbol{\theta})$] is

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\mathbf{ww}}\boldsymbol{\Gamma}^{'} + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}} & \boldsymbol{\Sigma}_{\mathbf{ww}}\boldsymbol{\Gamma}^{'} \\ \boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\mathbf{ww}} & \boldsymbol{\Sigma}_{\mathbf{ww}} \end{bmatrix} \tag{A5}$$

where $\boldsymbol{\Sigma}_{\mathbf{ww}}$ is the covariance matrix of the covariates in $\mathbf{w}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}}$ is the covariance matrix of the disturbances ($\boldsymbol{\varepsilon}$).

In the classic fixed effects model, we would drop $\mathbf{z}_i$ from $\mathbf{w}_i$ and the corresponding coefficients from $\boldsymbol{\Gamma}$, set $\mathbf{B}_{y_1\mathbf{x}_1} = \mathbf{B}_{y_2\mathbf{x}_2} = \cdots = \mathbf{B}_{y_T\mathbf{x}_T}$, and make $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}}$ a diagonal matrix with all elements of the main diagonal equal. The $\boldsymbol{\Sigma}_{\mathbf{ww}}$ covariance matrix allows all covariates to correlate, including the latent time-invariant variable. For the classic random effects model, we can return $\mathbf{z}_i$ to $\mathbf{w}_i$, but now we must constrain $\boldsymbol{\Sigma}_{\mathbf{ww}}$ so that all covariances of η with $\mathbf{x}_t$ and $\mathbf{z}$ are zero and we maintain the equality constraints on the coefficients so that $\mathbf{B}_{y_1\mathbf{x}_1} = \mathbf{B}_{y_2\mathbf{x}_2} = \cdots = \mathbf{B}_{y_T\mathbf{x}_T}$ and $\mathbf{B}_{y_1\mathbf{z}} = \mathbf{B}_{y_2\mathbf{z}} = \cdots = \mathbf{B}_{y_T\mathbf{z}}$. As explained in the text, we can easily test these restrictions in SEMs.

The Maximum Likelihood Estimator (MLE) is the most widely used estimator in SEM software. The fitting function that incorporates the MLE is

$$F_{ML} = \ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \ln|\mathbf{S}| + tr[\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{S}] - p + (\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta}))'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta})) \qquad \text{(A6)}$$

where $\mathbf{S}$ is the sample covariance matrix, $\bar{\mathbf{z}}$ is the vector of the sample means of the observed variables, p is the number of observed variables, $\ln$ is the natural log, $|\cdot|$ is the determinant, and tr is the trace of a matrix. The MLE estimator, $\hat{\boldsymbol{\theta}}$, is chosen so as to minimize F_{ML}. Like all MLEs, $\hat{\boldsymbol{\theta}}$, has several desirable properties. It is consistent, asymptotically unbiased, asymptotically efficient, asymptotically normally distributed, and the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$ is the inverse of the expected information matrix.

The MLE estimator as implemented in F_{ML} leads to a consistent estimator of all intercepts, means, coefficients, variances, and covariances in the model under a broad range of conditions. This means that in larger samples, the estimator will converge on the true parameters for valid models. However, if we wish to develop appropriate significance tests, then we need to make assumptions about the distributions of the observed variables. The classic assumption is that the observed variables come from a multivariate normal distribution. A slightly less restrictive distributional assumption that maintains the properties of the MLE and its significance tests is that the observed variables come from a multivariate distribution with no excess multivariate kurtosis (Browne 1984). Multivariate skewness is permitted as long as the multivariate kurtosis does not differ from that of a normal distribution.

Fortunately, even when there is excess multivariate kurtosis there are a variety of alternative ways to obtain asymptotically accurate signficance tests including bootstrapping techniques (e.g., Bollen and Stine 1990, 1993), corrected standard errors and chi-squares (e.g., Satorra and Bentler, 1994), or arbitrary distribution estimators (e.g., Browne 1984). See Bollen and Curran (2006:55-57) for further discussion and references. These options provide a broader range of choices than is true in the usual implementation of the classic fixed and random effects models.

Dynamic Fixed and Random Effects Models

In the econometric literature, "dynamic" models refers to fixed and random effects models with lagged dependent variables included among the covariates. In the usual implementations, the lagged dependent variable model creates considerable difficulties and are the source of much discussion (see, e.g., Hsiao 2003, Ch.4). Fortunately, these models are relatively straightforward in the SEM approach. A modification of equation (A1) permits lagged endogenous variables,

$$\mathbf{y}_i = \boldsymbol{\alpha} + \mathbf{R}\mathbf{y}_i + \boldsymbol{\Gamma}\mathbf{w}_i + \boldsymbol{\varepsilon}_i \qquad \text{(A7)}$$

where because of using a lagged dependent variable we need to redefine vectors to take account of treating the first time wave variable, y_{i1}, as predetermined and included among the other covariates and the presence of lagged y influences, so that

$$\mathbf{y}_i = \begin{bmatrix} y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} \quad \mathbf{w}_i = \begin{bmatrix} y_{i1} \\ \mathbf{x}_{i2} \\ \vdots \\ \mathbf{x}_{iT} \\ \mathbf{z}_i \\ \eta_i \end{bmatrix} \quad \boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix} \tag{A8}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_2 \\ \vdots \\ \alpha_T \end{bmatrix} \quad \boldsymbol{\Gamma} = \begin{bmatrix} \rho_{21} & \mathbf{B}_{y_2\mathbf{x}_2} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{y_1\mathbf{z}} & 1 \\ 0 & \mathbf{0} & \mathbf{B}_{y_3\mathbf{x}_3} & \cdots & \mathbf{0} & \mathbf{B}_{y_2\mathbf{z}} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{y_T\mathbf{x}_T} & \mathbf{B}_{y_T\mathbf{z}} & 1 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \rho_{32} & 0 & \cdots & 0 & 0 \\ 0 & \rho_{43} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \rho_{T,T-1} & 0 \end{bmatrix}$$

In this model, y_{i1} is predetermined and uncorrelated with $\boldsymbol{\varepsilon}_i$ as are the other covariates. However, there is a correlation between $y_{i2} \ldots y_{iT}$ and at least some elements of $\boldsymbol{\varepsilon}_i$ (e.g., y_{i2} correlates with ε_{i2}) so we need to consider all but the first wave (y_{i1}) as endogenous. For this model, the implied mean and covariance matrices become,

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} (\mathbf{I} - \mathbf{R})^{-1}(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\mu}_{\mathbf{w}}) \\ \boldsymbol{\mu}_{\mathbf{w}} \end{bmatrix} \tag{A9}$$

and the implied covariance matrix [$\boldsymbol{\Sigma}(\boldsymbol{\theta})$] is

$$\boldsymbol{\Sigma}(\theta) = \begin{bmatrix} (\mathbf{I}-\mathbf{R})^{-1}[\boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\mathbf{ww}}\boldsymbol{\Gamma}' + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}}](\mathbf{I}-\mathbf{R})^{-1} & \boldsymbol{\Sigma}_{\mathbf{ww}}\boldsymbol{\Gamma}'(\mathbf{I}-\mathbf{R})^{-1} \\ (\mathbf{I}-\mathbf{R})^{-1}\boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\mathbf{ww}} & \boldsymbol{\Sigma}_{\mathbf{ww}} \end{bmatrix} \qquad \text{(A10)}$$

Fortunately, we can continue to use the ML fitting function in equation (Fml) and the resulting estimator maintains the properties of an MLE under the precedingly described distributional assumptions and the corrected test statistics are also available when needed (see above).