ESTIMATING AUTOCORRELATIONS IN FIXED-EFFECTS MODELS

Gary Solon

Technical Working Paper No. **32**

NBER TECHNICAL PAPER SERIES

ESTIMATING AUTOCORRELATIONS IN FIXED-EFFECTS MODELS

Gary Solon

Technical Working Paper No. 32

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

February 1984

The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Technical Working Paper #32
February 1984

Estimating Autocorrelations in Fixed-Effects Models

ABSTRACT

This paper discusses the estimation of serial correlation in fixed-effects models for longitudinal data. Like time series data, longitudinal data often contain serially correlated error terms, but the autocorrelation estimators commonly used for time series, which are consistent as the length of the time series goes to infinity, are not consistent for a short time series as the size of the cross-section goes to infinity. This form of inconsistency is of particular concern because a short time series of a large cross-section is the typical case in longitudinal data.

This paper extends Nickell's method of correcting for the inconsistency of autocorrelation estimators by generalizing to higher than first-order autocorrelations and to error processes other than first-order autoregressions. The paper also presents statistical tables that facilitate the identification and estimation of autocorrelation processes in both the generalized Nickell method and an alternative method due to MaCurdy. Finally, the paper uses Monte Carlo methods to explore the finite-sample properties of both methods.

Gary Solon
Department of Economics
The University of Michigan
Ann Arbor, MI 48109
(313) 763-1306

ESTIMATING AUTOCORRELATIONS IN FIXED-EFFECTS MODELS

1. Introduction

The recent development of several major longitudinal data bases has led to an explosion in the analysis of pooled cross-section and time series data.[1] A common technique in such analysis is to assume that y_{it}, the dependent variable observed for the i^{th} individual at time t, is generated by the equation

$$y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it} \qquad (1)$$

where α_i is an individual-specific constant, X_{it} is a vector of explanatory variables, β is the associated parameter vector, and ε_{it} is an error term. The term α_i represents the combined effect of unobserved variables that remain fixed for the individual over time. This "fixed-effects" approach is sometimes referred to as the dummy variable technique because it can be estimated by entering individual-specific dummies as explanatory variables. It is also known as the covariance approach because an equivalent estimation technique is to subtract the time mean of (1) from (1) itself to "difference out" the fixed effects and produce the equation

$$y_{it} - \bar{y}_i = \beta'(X_{it} - \bar{X}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \qquad (2)$$

where, for any variable z_{it}, $\bar{z}_i = \sum_{t=1}^{T} z_{it}/T$.

An alternative approach, the "random-effects" technique, treats α_i as a random component of the error term and applies generalized least squares. This technique, however, yields an inconsistent estimator of β when some of the variables underlying α_i are correlated with some of the variables in X_{it}. Indeed, in many longitudinal analyses, such as those examining union wage effects,[2] the very purpose of using longitudinal data is to difference out unobserved individual-specific variables thought to be correlated with other explanatory variables. Therefore, the fixed-effects technique is clearly to be preferred in many applications.

Of course, if the error term ε_{it} is serially correlated because of the omission of variables that change gradually over time, ordinary least squares (OLS) estimation of the fixed-effects model (with either individual-specific dummies or subtraction of time means) is inefficient and also yields incorrect standard error estimators and hypothesis tests. This is probably a common problem in practice. Of the few longitudinal analyses that have checked for autocorrelation, most have found it to be present.[3]

In such circumstances, if the true autocorrelation structure and parameters were known, the analysis of the pooled cross-section and time series data could then proceed with the same methods commonly used to correct for autocorrelation in a single time series. If they are not known, however, new problems arise. As shown below, the conventional auto-correlation estimators, such as those generated by autoregressions of the OLS residuals, are consistent as the length of the time series T goes to infinity, but are not consistent for small T as the size of the cross-section N goes to infinity. This is particularly troubling because a short

time series of a large cross-section is the typical case in longitudinal data.

Nickell (1981) has already derived an exact expression for the inconsistency of the first-order autocorrelation estimator in the particular case where the error follows a first-order autoregression. As Nickell (1980) has suggested, this expression can be used to correct the inconsistency of the estimator, and the corrected estimate can be used to reestimate equation (1) in "quasi-first-differenced" form. Sections 2 and 3 of this paper extend Nickell's analysis to higher-order autocorrelation estimators and to more general autocorrelation structures. This extension, along with the accompanying statistical tables, enables an assessment of whether the observed autocorrelation pattern is well described as a first-order autoregressive process or whether some other structure is more appropriate. The tables also facilitate correction for the inconsistency of the estimators. Section 4 discusses an alternative approach due to MaCurdy (1982) and presents additional statistical tables that simplify the identification of the error structure under his approach.[4] Section 5 presents Monte Carlo evidence on the finite-sample properties of the alternative methods, and the final section briefly summarizes the paper.

2. Inconsistency of Autocorrelation Estimators

This section discusses the estimation of the correlation between ε_{it} and $\varepsilon_{i,t-K}$. It is assumed that, for any i and t, ε_{it} has zero mean and constant variance σ^2, and the correlation of ε_{it} and $\varepsilon_{i,t-K}$ is ρ_K. If $i \neq j$, ε_{it} and ε_{js} are assumed to be independently distributed. Now define r_K, the estimator of the K^{th}-order autocorrelation ρ_K, by

$$r_K = \frac{\sum_{i=1}^{N} \sum_{t=K+1}^{T} (e_{it} - \bar{e}_{i,K})(e_{i,t-K} - \bar{e}_{i,-K})}{\sum_{i=1}^{N} \sum_{t=K+1}^{T} (e_{i,t-K} - \bar{e}_{i,-K})^2}$$

where the e's denote the residuals from OLS estimation of equation (2),

$$\bar{e}_{i,K} = \frac{1}{T-K} \sum_{t=K+1}^{T} e_{it} ,$$

and $$\bar{e}_{i,-K} = \frac{1}{T-K} \sum_{t=K+1}^{T} e_{i,t-K} .$$

This estimator can easily be computed by an autoregression (with constant term) of the OLS residuals.

Because longitudinal data typically consist of a short time series of a large cross-section, we will now investigate the probability limit of r_K as the size of the cross-section N grows large while the length of the time series T is fixed. This probability limit is

$$\operatorname*{plim}_{N\to\infty} r_K = \frac{\operatorname*{plim}_{N\to\infty} \frac{1}{N(T-K)} \sum_{i=1}^{N} \sum_{t=K+1}^{T} (e_{it} - \bar{e}_{i,K})(e_{i,t-K} - \bar{e}_{i,-K})}{\operatorname*{plim}_{N\to\infty} \frac{1}{N(T-K)} \sum_{i=1}^{N} \sum_{t=K+1}^{T} (e_{it-K} - \bar{e}_{i,-K})^2} .$$

Since the OLS estimator of β in equation (2) is consistent, the probability limit of the difference between e_{it} and $\varepsilon_{it} - \bar{\varepsilon}_i$ is zero. It follows that

$$\operatorname*{plim}_{N\to\infty} r_K = \frac{\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T-K} \sum_{t=K+1}^{T} (\varepsilon_{it} - \bar{\varepsilon}_{i,K})(\varepsilon_{i,t-K} - \bar{\varepsilon}_{i,-K}) \right]}{\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T-K} \sum_{t=K+1}^{T} (\varepsilon_{i,t-K} - \bar{\varepsilon}_{i,-K})^2 \right]} .$$

Now define the terms

$$A_i = \frac{1}{T-K} \sum_{t=K+1}^{T} (\varepsilon_{it} - \bar{\varepsilon}_{i,K})(\varepsilon_{i,t-K} - \bar{\varepsilon}_{i,-K}) ,$$

$$B_i = \frac{1}{T-K} \sum_{t=K+1}^{T} (\varepsilon_{i,t-K} - \bar{\varepsilon}_{i,-K})^2 ,$$

$$\bar{A}_N = \frac{1}{N} \sum_{i=1}^{N} A_i ,$$

and $$\bar{B}_N = \frac{1}{N} \sum_{i=1}^{N} B_i .$$

Then

$$\operatorname*{plim}_{N\to\infty} r_K = \frac{\operatorname*{plim}_{N\to\infty} \bar{A}_N}{\operatorname*{plim}_{N\to\infty} \bar{B}_N} .$$

Furthermore, given the conditions described in the beginning of this section, the law of large numbers implies that

$$\operatorname*{plim}_{N\to\infty} \bar{A}_N = E(A_i)$$

and $$\operatorname*{plim}_{N\to\infty} \bar{B}_N = E(B_i),$$

so that

$$\operatorname*{plim}_{N\to\infty} r_K = \frac{E(A_i)}{E(B_i)} .$$

Following Kendall's (1954) analysis of a single time series, let $V = T-K$ and, without loss of generality, let $\sigma^2 = 1$. Then, dropping i subscripts,

$$E(B_i) = \frac{1}{V} E\left[\sum_{t=K+1}^{T} (\varepsilon_{t-K} - \bar{\varepsilon}_{-K})^2\right]$$

$$= \frac{1}{V} E\left[\sum_{t=K+1}^{T} \varepsilon^2_{t-K} - V\bar{\varepsilon}^2_{-K}\right]$$

$$= 1 - E\left[\frac{\sum_{t=K+1}^{T} \varepsilon_{t-K}}{V}\right]^2$$

$$= 1 - \frac{1}{V^2}\left[V + 2\sum_{j=1}^{V-1} (V-j)\,\rho_j\right]$$

$$= 1 - \frac{1}{V} - \frac{2}{V^2}\sum_{j=1}^{V-1} (V-j)\,\rho_j \quad , \qquad (3)$$

$$\text{and} \quad E(A_i) = \frac{1}{V} E\left[\sum_{t=K+1}^{T} (\varepsilon_t - \bar{\varepsilon}_K)(\varepsilon_{t-K} - \bar{\varepsilon}_{-K})\right]$$

$$= \frac{1}{V} E\left[\sum_{t=K+1}^{T} \varepsilon_t \varepsilon_{t-K} - V \bar{\varepsilon}_K \bar{\varepsilon}_{-K}\right]$$

$$= \rho_K - \frac{1}{V^2} E\left[\sum_{t=K+1}^{T} \varepsilon_t \sum_{t=K+1}^{T} \varepsilon_{t-K}\right]$$

$$= \rho_K - \frac{1}{V^2}\left[\sum_{j=0}^{V-1} (V-j)\rho_{K+j} + \sum_{j=1}^{K} (V-j)\rho_{K-j} + \sum_{j=1}^{V-K-1} (V-K-j)\rho_j\right]. \qquad (4)$$

This last expression assumes $V > K + 2$. If $V < K + 2$, the last summation vanishes from equation (4). Similarly, if $V < K$, the upper bound of the index in the second summation should be V instead of K.

With expressions (3) and (4) for $E(A_i)$ and $E(B_i)$, we now have a general solution for

$$\underset{N\to\infty}{\text{plim}} \quad r_K = \frac{E(A_i)}{E(B_i)} .$$

This expression encompasses as a special case Nickell's (1981) equation (17) for the probability limit of r_1 when ε_{it} follows a first-order autoregression. It is more general, however, in that it applies to autocorrelation structures other than first-order autoregressions as well as to higher than first-order autocorrelations.

Inspection of equations (3) and (4) reveals that as T and hence V go to infinity, $\underset{N\to\infty}{\text{plim}}\ r_K$ converges in probability to ρ_K. For small T and V, however, $\underset{N\to\infty}{\text{plim}}\ r_K$ does not equal ρ_K. This inconsistency of r_K in short time series is illustrated in the next section for the cases where ε_{it} is serially independent, follows a first-order autoregression, or follows a first-order moving-average process. In each case, an exact expression for the inconsistency of the autocorrelation estimators enables correction of the estimates. Where serial correlation is indeed present, the corrected estimates can be used to achieve more efficient estimation of β and more appropriate hypothesis tests.

3. Some Special Cases

Serial independence: $\rho_K = 0$ for $K \neq 0$.

For $K \neq 0$, applying equations (3) and (4) to this case yields

$$E(B_i) = 1 - \frac{1}{V}$$

and

$$E(A_i) = -\frac{1}{V^2}(V-K),$$

so that

$$\operatorname*{plim}_{N\to\infty} r_K = \frac{E(A_i)}{E(B_i)}$$

$$= -\frac{V-K}{V(V-1)}.$$

For example, as N (but not T) goes to infinity, the probability limit of the first-order autocorrelation estimator is

$$\operatorname*{plim}_{N\to\infty} r_1 = -\frac{1}{V},$$

and

$$\operatorname*{plim}_{N\to\infty} r_2 = -\frac{V-2}{V(V-1)}.$$

If T = 6, for instance, the probability limit of r_1 is -.20 and that of r_2 is -.17 even though the true ρ_1 and ρ_2 are both zero. To put it another way, if autoregressions of the OLS residuals give estimated first- and second-order autocorrelations of approximately -.2, this result is entirely in accordance with the hypothesis that ε_{it} is serially uncorrelated.

AR(1): $\rho_K = \rho^K$.

For any $K \neq 0$,

$$E(B_i) = 1 - \frac{1}{V} - \frac{2}{V^2}\left[(V-1)\rho + (V-2)\rho^2 + \ldots + 2\rho^{V-2} + \rho^{V-1}\right]$$

$$= 1 - \frac{1}{V} - \frac{2}{V^2}\left[(\rho+\rho^2+\ldots+\rho^{V-1})+(\rho+\rho^2+\ldots+\rho^{V-2})+\ldots+(\rho+\rho^2)+\rho\right]$$

$$= 1 - \frac{1}{V} - \frac{2}{V^2}\left[\frac{\rho-\rho^V}{1-\rho} + \frac{\rho-\rho^{V-1}}{1-\rho} + \ldots + \frac{\rho-\rho^2}{1-\rho}\right]$$

$$= 1 - \frac{1}{V} - \frac{2}{V^2}\left[\frac{(V-1)\rho - \rho^V - \rho^{V-1} - \ldots - \rho^2}{1-\rho}\right]$$

$$= 1 - \frac{1}{V} - \frac{2}{V^2}\left[\frac{(V-1)\rho}{1-\rho} - \frac{\rho^2 - \rho^{V+1}}{(1-\rho)^2}\right]$$

$$= 1 - \frac{1}{V} - \frac{2\rho}{V^2(1-\rho)}\left[V - \frac{1-\rho^V}{1-\rho}\right]$$

$$= 1 - \frac{1}{V} - \frac{2\rho}{V(1-\rho)} + \frac{2\rho(1-\rho^V)}{V^2(1-\rho)^2}$$

$$= 1 - \frac{1+\rho}{V(1-\rho)} + \frac{2\rho(1-\rho^V)}{V^2(1-\rho)^2}$$

$$\text{and} \quad E(A_i) = \rho^K - \frac{1}{V^2}\left[\sum_{j=0}^{V-1}(V-j)\,\rho^{K+j} + \sum_{j=1}^{K}(V-j)\,\rho^{K-j} + \sum_{j=1}^{V-K-1}(V-K-j)\,\rho^{j}\right]$$

$$= \rho^K - \frac{\rho^K}{V} - \frac{1}{V^2}\left[\sum_{j=1}^{V-1}(V-j)\rho^{K+j} + \sum_{j=1}^{K}(V-j)\rho^{K-j} + \sum_{j=1}^{V-K-1}(V-K-j)\rho^{j}\right]$$

$$= \rho^K\left[1 - \frac{1}{V} - \frac{2}{V^2}\sum_{j=1}^{V-1}(V-j)\,\rho^{j}\right]$$

$$+ \frac{1}{V^2}\left[\sum_{j=1}^{V-1}(V-j)\rho^{K+j} - \sum_{j=1}^{K}(V-j)\rho^{K-j} - \sum_{j=1}^{V-K-1}(V-K-j)\rho^{j}\right]$$

$$= \rho^K\,E(B_i) + \frac{1}{V^2}\left[\sum_{j=1}^{V-1}(V-j)\rho^{K+j} - \sum_{j=1}^{K}(V-j)\rho^{K-j} - \sum_{j=1}^{V-K-1}(V-K-j)\rho^{j}\right].$$

Hence,

$$\operatorname*{plim}_{N\to\infty} r_K = \rho^K + \frac{C}{D} \tag{5}$$

where $$C = \sum_{j=1}^{V-1}(V-j)\rho^{K+j} - \sum_{j=1}^{K}(V-j)\rho^{K-j} - \sum_{j=1}^{V-K-1}(V-K-j)\rho^{j}$$

and $$D = V^2 - \frac{V(1+\rho)}{1-\rho} + \frac{2\rho\,(1-\rho^V)}{(1-\rho)^2}\,.$$

If K = 1, then

$$C = \sum_{j=1}^{V-1} (V-j)\rho^{j+1} - (V-1) - \sum_{j=1}^{V-K-1} (V-j-1)\rho^{j}$$

$$= 1 - V + \left[(V-1)\rho^2 + (V-2)\rho^3 + \ldots + 2\rho^{V-1} + \rho^V \right]$$

$$- \left[(V-2)\rho + (V-3)\rho^2 + \ldots + 2\rho^{V-3} + \rho^{V-2} \right]$$

$$= 1 - V - V\rho + 2\left(\rho + \rho^2 + \ldots + \rho^V\right) - \rho^V$$

$$= 1 - V - V\rho - \rho^V + 2\,\frac{\rho - \rho^{V+1}}{1-\rho}$$

and $$\operatorname*{plim}_{N\to\infty} r_1 = \rho + \frac{C}{D}$$

$$= \rho + \frac{1 - V - V\rho - \rho^V + 2\,\frac{\rho - \rho^{V+1}}{1-\rho}}{V^2 - \frac{V(1+\rho)}{1-\rho} + \frac{2\rho(1-\rho^V)}{(1-\rho)^2}} \quad .$$

This last expression can be shown to be exactly equivalent to Nickell's (1981) equation (17).

Proceeding to higher-order autocorrelations not considered by Nickell, if K = 2, the appropriate algebraic manipulation shows that

$$C = 2 - V - 2V\rho - V\rho^2 + 3\rho^{V-1} + 2\rho^V + \rho^{V+1} + 4\left[\frac{\rho - \rho^{V-1}}{1-\rho}\right],$$

which can be substituted into equation (5) to obtain $\underset{N\to\infty}{\text{plim}}\ r_2$. Similarly, if K = 3,

$$C = 3 - V - 2V\rho - 2V\rho^2 - V\rho^3 + 5\rho^{V-2} + 4\rho^{V-1} + 3\rho^V + 2\rho^{V+1} + \rho^{V+2} + 6\left[\frac{\rho - \rho^{V-2}}{1-\rho}\right],$$

which can be substituted into (5) to obtain $\underset{N\to\infty}{\text{plim}}\ r_3$.

To facilitate the use of these results, the probability limits of r_1, r_2, and r_3 as $N\to\infty$ are displayed in Table 1 for selected values of ρ and T. For example, if T = 10 and autoregressions of the OLS residuals yield autocorrelation estimates of $r_1 = .3$, $r_2 = 0$, and $r_3 = -.2$, this pattern would be quite consistent with an assumption that the error follows a first-order autoregressive process with parameter $\rho = .5$. One could then reestimate equation (1) in "quasi-first-differenced" form, as is frequently done with single time series, to achieve more efficient estimation of β and more accurate estimation of standard errors.

Table 1 is useful not only for correcting for the inconsistency of the conventional autocorrelation estimators, but also for examining whether the autocorrelation pattern in the residuals accords with the assumption of an AR(1) error structure. For example, with T = 10 again, an observed pattern of $r_1 = .4$, $r_2 = -.2$, and $r_3 = -.2$ would not be consistent with any AR(1) structure. It is, however, approximately what one would expect to see if the error follows a first-order moving-average process with first-order autocorrelation $\rho = .5$, as will be demonstrated below.

MA(1): $\rho_1 = \rho$, $\rho_K = 0$ for $K > 1$.

For any $K \neq 0$,

$$E(B_i) = 1 - \frac{1}{V} - \frac{2}{V^2}(V-1)\rho .$$

For $K > 2$,

$$E(A_i) = -\frac{1}{V^2}[(V-K+1)\rho + V - K + (V-K-1)\rho]$$

$$= -\frac{1}{V^2}[2(V-K)\rho + V - K]$$

$$= -\frac{V-K}{V^2}(2\rho + 1) ,$$

so that, for $K > 2$,

$$\operatorname*{plim}_{N\to\infty} r_K = -\frac{(V-K)(2\rho+1)}{V^2-V-2(V-1)\rho} .$$

For $K = 1$,

$$E(A_i) = \rho - \frac{1}{V^2}[V\rho + V - 1 + (V-2)\rho]$$

$$= \rho\, E(B_i) + \frac{1}{V^2}[(2-V)\rho + 2(V-1)\rho^2 - V + 1],$$

so that

$$\text{plim } r_1 = \rho + \frac{(2-V)\rho + 2(V-1)\rho^2 - V + 1}{V^2 - V - 2(V-1)\rho} .$$

Table 2 presents, for the MA(1) case, the probability limits of r_1, r_2, and r_3 as $N \to \infty$ for selected values of ρ and T. For example, if T = 10 and autoregressions of the OLS residuals yield autocorrelation estimates of $r_1 = .4$, $r_2 = -.2$, and $r_3 = -.2$, this pattern would be consistent with an assumption that the error follows a first-order moving-average process with first-order autocorrelation $\rho = .5$. On the other hand, with T = 10, an observed pattern of $r_1 = .6$, $r_2 = .2$, and $r_3 = 0$ would not be consistent with any of the displayed MA(1) structures. Reference to Table 1, however, reveals that this pattern is approximately what one would expect if the error is AR(1) with $\rho = .8$.

4. An Alternative Approach

The approach described above involves analysis of the residuals from OLS estimation of equation (2). An alternative approach, suggested by MaCurdy (1982), begins by taking first differences of equation (1) to obtain

$$y_{it} - y_{i,t-1} = \beta'(X_{it} - X_{i,t-1}) + \varepsilon_{it} - \varepsilon_{i,t-1}. \qquad (6)$$

MaCurdy suggests estimating equation (6) by OLS and then using the resulting residuals e^*_{it} to estimate autocorrelations of the differenced error term $\varepsilon_{it} - \varepsilon_{i,t-1}$. Since the OLS estimator of β in equation (6) is consistent as N goes to infinity, the probability limit of the difference between e^*_{it} and $\varepsilon_{it} - \varepsilon_{i,t-1}$ is zero. As a result, conventional autocorrelation estimators applied to e^*_{it} are consistent for the autocorrelations of the differenced error term.

For example, consider the estimator

$$r^*_K = \frac{\sum_{i=1}^{N} \sum_{t=K+2}^{T} e^*_{it}\, e^*_{i,t-K}}{\sum_{t=K+2}^{T} e^{*\,2}_{i,t-K}},$$

which can easily be computed by an autoregression (without constant term) of the OLS residuals from equation (6).

Then

$$\operatorname*{plim}_{N\to\infty} r_K^* = \frac{\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T-K-1} \sum_{t=K+2}^{T} (\varepsilon_{it} - \varepsilon_{i,t-1})(\varepsilon_{i,t-K} - \varepsilon_{i,t-K-1})}{\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T-K-1} \sum_{t=K+2}^{T} (\varepsilon_{i,t-K} - \varepsilon_{i,t-K-1})^2} .$$

Manipulation of this expression similar to that in Section 2 produces the result

$$\operatorname*{plim}_{N\to\infty} r_K^* = \frac{2\rho_K - \rho_{K+1} - \rho_{K-1}}{2(1-\rho_1)} = \rho_K^* \qquad (7)$$

where ρ_J is the true J^{th}-order autocorrelation of the original error term. The expression in equation (7) for the probability limit of r_K^* is the true K^{th}-order autocorrelation ρ_K^* of the differenced error term.

In principle, once one has used the estimates r_K^* to identify and estimate the autocorrelation structure of the differenced error term, one may "integrate back" to obtain the autocorrelation structure of the original error term. In practice, however, the differencing procedure may obscure the original error structure. For example, if the original error follows an AR(1) process with autoregressive parameter ρ, the differenced error follows an ARMA(1,1) process with autoregressive parameter ρ and moving-average parameter -1. This more complicated autocorrelation structure makes it difficult to recognize that, in levels, the error follows a simple AR(1) process. Indeed, MaCurdy's analysis of longitudinal earnings data from the Panel Study of Income Dynamics concludes that the differenced

error follows an MA(2) process and the original error therefore follows an ARMA(1,2) process with a unit autoregressive parameter. As will be shown below, however, a much simpler AR(1) process for the original error fits the observed residuals pattern almost as well.

The remainder of this section develops statistical tables to facilitate recognition of simple error structures in MaCurdy's method.[5] The tables give ρ_1^*, ρ_2^*, and ρ_3^* -- the probability limits of the autocorrelation estimators for the differenced error term -- when the original error is serially independent, AR(1), AR(2), MA(1), or MA(2).

Serial independence : $\rho_K = 0$ for $K \neq 0$.

Applying equation (7) yields

$$\rho_1^* = -\frac{1}{2}$$

and

$$\rho_K^* = 0$$

for $K > 1$. Hence, if the residuals from OLS estimation of equation (6) show a first-order autocorrelation of about $-\frac{1}{2}$ and higher-order autocorrelations close to zero, this result accords with the hypothesis that ε_{it} is serially uncorrelated.

AR(1) : $\rho_K = \rho^K$.

Applying equation (7),

$$\rho_K^* = \frac{2\rho^K - \rho^{K+1} - \rho^{K-1}}{2(1-\rho)}$$

$$= -\frac{\rho^{K-1}(\rho^2 - 2\rho + 1)}{2(1-\rho)}$$

$$= -\frac{\rho^{K-1}(1-\rho)}{2} .$$

Hence,

$$\rho_1^* = -\frac{1-\rho}{2} , \qquad (8)$$

$$\rho_2^* = -\frac{\rho(1-\rho)}{2} ,$$

$$\rho_3^* = -\frac{\rho^2(1-\rho)}{2} ,$$

and so forth.

To facilitate the use of these results, the values of ρ_1^*, ρ_2^* and ρ_3^* are displayed in Table 3. Suppose, for example, that $r_1^* = -.32$, $r_2^* = -.09$, and $r_3^* = -.01$. Examination of Table 3 shows this pattern is roughly consistent with an assumption that ε_{it} follows an AR(1) process with $\rho = .3$, in which case the probability limits of the first three autocorrelation estimators are respectively -.35, -.105, and -.0315. The supposed values

for r_1^*, r_2^*, and r_3^* are in fact the ones MaCurdy observed and modeled as an ARMA(1,2) process. An AR(1) process with $\rho = .3$, however, fits the residuals pattern nearly as well and is much more tractable for purposes of estimating equation (2) with a correction for autocorrelation.

AR(2) : $\varepsilon_{it} = \lambda_1 \varepsilon_{i,t-1} + \lambda_2 \varepsilon_{i,t-2} + v_{it}$ where v_{it} is independently, identically distributed with zero mean and constant variance.

In this case, the true autocorrelations are[6]

$$\rho_1 = \frac{\lambda_1}{1 - \lambda_2} ,$$

$$\rho_2 = \frac{\lambda_1^2 + \lambda_2 - \lambda_2^2}{1 - \lambda_2} ,$$

and, for $K > 2$,

$$\rho_K = \lambda_1 \rho_{K-1} + \lambda_2 \rho_{K-2} .$$

Then, use of equation (7) and some algebraic manipulation gives

$$\rho_1^* = \frac{2\lambda_1 - \lambda_1^2 + \lambda_2^2 - 1}{2(1 - \lambda_1 - \lambda_2)} ,$$

$$\rho_2^* = \frac{2\lambda_1^2 + 2\lambda_2 - 2\lambda_2^2 - \lambda_1^3 - 2\lambda_1\lambda_2 + \lambda_1\lambda_2^2 - \lambda_1}{2(1 - \lambda_1 - \lambda_2)} ,$$

and $$\rho_3^* = \frac{2\lambda_1^3 + 4\lambda_1\lambda_2 - 2\lambda_1\lambda_2^2 - \lambda_1^4 - 3\lambda_1^2\lambda_2 + \lambda_1^2\lambda_2^2 + \lambda_2^3 - \lambda_1^2 - \lambda_2}{2\,(1 - \lambda_1 - \lambda_2)}.$$

Table 4 presents ρ_1^*, ρ_2^*, and ρ_3^* for selected values of λ_1 and λ_2 consistent with a stationary AR(2) process. For example, if autoregressions of the OLS residuals from equation (6) yield $r_1^* = -.35$, $r_2^* = -.1$, and $r_3^* = -.3$, this pattern would be roughly consistent with an assumption that ε_{it} follows an AR(2) process with $\lambda_1 = 1.2$ and $\lambda_2 = -.5$.

MA(1) : $\rho_1 = \rho$, $\rho_K = 0$ for $K > 1$.

In this case,

$$\rho_1^* = \frac{2\rho - 1}{2\,(1 - \rho)},$$

$$\rho_2^* = \frac{\rho}{2\,(1 - \rho)},$$

and, for $K > 2$,

$$\rho_K^* = 0.$$

Table 5 displays ρ_1^*, ρ_2^*, and ρ_3^* for selected admissible values of ρ.

MA(2) : $\rho_1, \rho_2 \neq 0, \quad \rho_K = 0$ for $K > 2$.

In this case,

$$\rho_1^* = \frac{2\rho_1 - \rho_2 - 1}{2(1 - \rho_1)},$$

$$\rho_2^* = \frac{2\rho_2 - \rho_1}{2(1 - \rho_1)},$$

$$\rho_3^* = \frac{\rho_2}{2(1 - \rho_1)},$$

and, for $K > 3$,

$$\rho_K^* = 0$$

Table 6 displays ρ_1^*, ρ_2^*, and ρ_3^* for selected admissible values of ρ_1 and ρ_2.

5. Monte Carlo Results

The preceding sections have derived the probability limits (as $N \to \infty$) of various autocorrelation estimators. Two remaining questions are how closely, in finite samples, do these estimators cluster around their probability limits, and to what extent does using these estimators to correct for serial correlation improve the estimation of the original regression equation (1)? This section explores these questions with Monte Carlo methods.

In each of the Monte Carlo experiments, the regression model takes the form

$$y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it} \cdot \tag{9}$$

In all experiments, $\beta = 1$ and the α_i are independently and approximately normally distributed[7] with mean 0 and variance 1.8. For each individual i, X_{it} follows the first-order autoregressive process

$$X_{it} = .5\, X_{i,t-1} + w_{it}$$

where the w_{it} are independently and approximately normally distributed with mean 0 and variance 1. The same set of α_i and X_{it} is reused in every trial of every experiment.

Each trial generates a new sample of ε_{it} according to the first-order autoregressive equation

$$\varepsilon_{it} = \rho \varepsilon_{i,t-1} + v_{it}$$

where the v_{it} are independently and approximately normally distributed with mean 0. The autocorrelation parameter ρ is varied between experiments. The variance of v_{it} also is altered between experiments so that, in all experiments, the variance of ε_{it} is 1/.84.

The first series of experiments uses time series length T = 10 and cross-section size N = 50. Each experiment uses a different value for ρ and is replicated 50 times. In each replication, once the data are generated, equation (9) is estimated by OLS in both its level and first-differenced forms. Autocorrelations are then estimated from the residuals of both regressions and compared with the corresponding probability limits in Tables 1 and 3. After the appropriate adjustments have been made, these estimates are used to reestimate equation (9) with correction for serial correlation. The remainder of this section summarizes the findings of these and related experiments.

The first finding is that, as expected, estimation of equation (9) without appropriate correction for serial correlation tends to produce misleading standard error estimates for $\hat{\beta}$. Table 7 reports the relevant summary statistics for experiments with T = 10, N = 50, and ρ set at 0, .4, .8, or .99. The first column gives the average over 50 replications of the estimated standard error of $\hat{\beta}$ from OLS estimation of equation (9). The second column shows the standard deviation actually observed for $\hat{\beta}$ in the 50 replications.[8] Columns 3 and 4 report the analagous statistics for OLS estimation of the first-differenced version of equation (9).

Of course, in the experiment with ρ = 0, OLS is the appropriate pro-

cedure, and the average standard error estimate is nearly identical to the observed standard error of $\hat{\beta}$. Applying OLS to the first-differenced equation, however, erroneously assumes that ε_{it} follows a random walk, and consequently the average standard error estimate considerably understates the variability actually observed for $\hat{\beta}$. The table also shows that, as ρ grows larger and the error process becomes more like a random walk, the standard error estimator in the first-differenced version performs better while the standard error estimator from applying OLS in levels tends to understate the actual variability in $\hat{\beta}$.

The second set of findings concerns how closely the autocorrelation estimators cluster around their probability limits. Table 8 gives the relevant results from the experiments with T = 10, N = 50, and ρ equal to 0, .4, .8, or .99. The first three columns pertain to the first-, second-, and third-order autocorrelation estimators based on the OLS residuals. For each value of ρ, Table 8 shows the probability limits of r_1, r_2, and r_3, as previously displayed in Table 1. Table 8 also shows the average values of r_1, r_2, and r_3 actually observed over 50 replications, as well their observed standard deviations.[9] Similarly, columns 4-6 give the probability limits (from Table 3) and the sample means and standard deviations of the autocorrelation estimators r_1^*, r_2^*, and r_3^* for the first-differenced equation.

As can be seen in the table, for all the values of ρ, the probability limits of the autocorrelation estimators for the first-differenced equation appear to be very good measures of the central tendency of these

estimators in finite samples. For example, with $\rho = 0$, the probability limits of r_1^*, r_2^*, and r_3^* are -.5, 0, and 0, while the observed sample averages are -.49, -.00, and -.01. The observed standard errors in the sample are large enough (usually about .05) so that the true AR(1) error structure will occasionally be difficult to recognize. This result, however, pertains to experiments with a cross-section size of only 50. In most actual applications, the cross-section will be much larger and the standard errors correspondingly smaller.

The results are somewhat different for the autocorrelation estimators from the level equation. When ρ is 0 or .4, the average autocorrelation estimates in the experiment are quite close to the corresponding probability limits. When ρ is larger, however, the higher-order autocorrelation estimators tend to diverge from their probability limits. When $\rho = .8$, the probability limit of r_3 is -.04, but the sample mean of r_3 is -.14. When $\rho = .99$, the probability limit of r_2 is .41, but its sample mean is .31, and the probability limit of r_3 is .16, but its sample mean is -.04. For all values of ρ, the probability limits and sample means of r_1 agree reasonably well, but the discrepancies observed for the higher-order autocorrelation estimators suggest that, when ρ is large, it may be difficult to identify the true autocorrelation structure with the residuals from the level equation.

The experiments reported in Table 9 further explore the finite-sample properties of the autocorrelation estimators by reducing the time series length to $T = 6$. In each of these experiments, $\rho = .4$. In the

experiment with N = 50, the sample averages of the autocorrelation estimators from the first-differenced equation continue to agree well with their probability limits. The shortening of the time series, however, appears to exacerbate the problems of the estimators from the level equation. Even with a relatively small ρ of .4, the sample mean of r_2 is -.28, compared to its probability limit of -.22, and the sample mean of r_3 is -.37, compared to its probability limit of -.10.

The rest of Table 9 reports the results of experiments with N increased to 100 and 200. The successive doublings of the cross-section size reduce the observed standard errors of the autocorrelation estimators as expected, but do not move the sample means of r_2 and r_3 appreciably closer to their probability limits. At least for data bases of these sizes, a comparison of the autocorrelation estimates from the level residuals with the tabulated probability limits may not give an accurate means of identifying the true autocorrelation structure.

The experiments reported in Table 10 address the question of whether attempting to correct for serial correlation improves the estimation of the original regression equation. These experiments <u>assume</u> that the error structure has been correctly identified as AR(1). Then r_1, the first-order autocorrelation estimate based on the residuals from the level regression, or r_1^*, the estimate from the first-differenced regression, is used to generate an estimate $\hat{\rho}$ of the autocorrelation parameter, which is then used in the reestimation of equation (9). The estimate of ρ based on r_1 is obtained by interpolating the observed r_1 into the relevant section of Table 1. The estimate of ρ based on r_1^* is obtained by inverting

equation (8) in Section 4 to get

$$\hat{\rho} = 1 + 2r_1^* .$$

Both estimates of ρ are then used to reestimate equation (9) in quasi-first-differenced form, i.e., with new variables z_{it}^* obtained from the transformation $z_{it}^* = z_{it} - \hat{\rho} z_{i,t-1}$. The experiments of Table 10 use N = 50, T = 10, and ρ equal to .4 or .8.

The first two rows of the table reproduce from Table 7 the average estimated standard errors of $\hat{\beta}$ from OLS and first-differencing, along with the sample standard deviations actually observed for these estimators of β. The next two rows give the analagous statistics for the quasi-first-difference estimators based respectively on r_1 and r_1^*.

When $\rho = .4$, estimating in first differences tends to produce moderate underestimates of the standard error of $\hat{\beta}$. OLS estimation in levels appears to be a little more efficient for β and to give smaller underestimates of the standard error. Both methods of correcting for serial correlation appear to eliminate the underestimation of the standard error of $\hat{\beta}$ and to be more efficient than first-differencing (but not OLS) for estimating β. There is no clear difference between the estimators based on r_1 and those based on r_1^*.

When $\rho = .8$, OLS gives substantial underestimates of the standard error of $\hat{\beta}$. Estimating in first differences is much more efficient for β

and seems to underestimate the standard error of $\hat{\beta}$ only slightly. The quasi-first-difference methods further reduce the underestimation of standard errors and achieve large efficiency gains relative to OLS similar to those obtained by first-differencing. Again, there is no clear difference between the estimators based on r_1 and those based on r_1^*.

As always in Monte Carlo research, it is unclear to what extent the results can be generalized beyond the specific conditions of the experiments. Within these conditions, though, the following patterns have emerged:

(1) The results have confirmed that failure to correct for serial correlation may lead to inefficient estimation and misleading hypothesis tests.

(2) In experiments that assumed correct identification of autocorrelation structure, the methods developed in this paper to estimate and correct for serial correlation sometimes achieved major efficiency gains and more reliable standard error estimates. For example, in experiments with $\rho = .8$, the observed standard deviation of the coefficient estimator $\hat{\beta}$ was more than one-third less with correction for serial correlation than with OLS. No major differences in efficiency appeared between the alternative correction methods.

(3) Important differences did appear in the reliability of the alternative methods for identifying autocorrelation structure. In certain circumstances (i.e., with larger ρ or shorter T), the probability limits

derived for the higher-order autocorrelation estimators based on the OLS residuals appeared to be poor measures of the estimators' central tendency in finite samples. In such circumstances, it may be difficult to infer from the estimated autocorrelations what the true autocorrelation structure is. In contrast, the probability limits of the autocorrelation estimators based on residuals from estimating in first differences proved to be good measures of these estimators' central tendency through the full range of experimental conditions. If these results are generalizable, applying OLS to the first-differenced regression equation and then examining the autocorrelation patterns of the resulting residuals may be the most reliable way to identify the autocorrelation structure of the error process.

6. Summary

This paper has discussed methods of estimating autocorrelations in fixed-effects models for longitudinal data. It has extended Nickell's method by deriving a more general expression for the inconsistency of conventional autocorrelation estimators as the size of the cross-section N (but not the length of the time series T) goes to infinity, and it has presented statistical tables to facilitate identification of autocorrelation structures and correction for inconsistency of autocorrelation estimators. The paper also has presented tables to simplify identification of autocorrelation processes in MaCurdy's first-differencing method. Finally, the paper has reported Monte Carlo results suggesting that, in finite samples, the first-differencing approach may be a more reliable method for identifying autocorrelation structure.

TABLE 1

Probability Limits of r_1, r_2, and r_3 as $N \to \infty$ in AR(1) Model with Selected Values of ρ and T

	T = 6			T = 8			T = 10		
ρ	r_1	r_2	r_3	r_1	r_2	r_3	r_1	r_2	r_3
0	-.20	-.17	.00	-.14	-.13	-.10	-.11	-.11	-.10
.1	-.12	-.20	-.02	-.06	-.16	-.13	-.02	-.12	-.12
.2	-.05	-.22	-.05	.02	-.16	-.15	.06	-.12	-.14
.3	.03	-.23	-.07	.11	-.16	-.18	.15	-.10	-.16
.4	.10	-.22	-.10	.19	-.13	-.20	.23	-.07	-.17
.5	.17	-.20	-.12	.26	-.10	-.21	.32	-.02	-.17
.6	.24	-.16	-.13	.34	-.05	-.20	.40	.04	-.15
.7	.31	-.12	-.12	.42	.02	-.18	.48	.12	-.11
.8	.37	-.06	-.10	.49	.09	-.13	.56	.20	-.04
.9	.44	.02	-.06	.56	.18	-.05	.63	.31	.05
.99	.49	.09	-.01	.62	.28	.04	.69	.41	.16

	T = 15			T = 30			T = 100		
ρ	r_1	r_2	r_3	r_1	r_2	r_3	r_1	r_2	r_3
0	-.07	-.07	-.07	-.03	-.03	-.03	-.01	-.01	-.01
.1	.02	-.08	-.08	.06	-.03	-.04	.09	-.00	-.01
.2	.11	-.06	-.10	.16	-.01	-.04	.19	.03	-.01
.3	.20	-.03	-.10	.25	.03	-.04	.29	.07	.01
.4	.30	.01	-.10	.35	.09	-.01	.39	.14	.04
.5	.39	.08	-.08	.45	.17	.03	.48	.23	.10
.6	.48	.16	-.03	.54	.27	.10	.58	.33	.18
.7	.56	.25	.04	.64	.38	.20	.68	.46	.30
.8	.65	.36	.13	.73	.51	.34	.78	.61	.47
.9	.73	.48	.26	.82	.66	.51	.88	.77	.67
.99	.79	.59	.40	.89	.79	.68	.96	.92	.89

TABLE 2

Probability Limits of r_1, r_2, and r_3 as $N \to \infty$
in MA(1) Model with Selected Values of ρ and T

	T = 6			T = 8			T = 10		
ρ	r_1	r_2	r_3	r_1	r_2	r_3	r_1	r_2	r_3
0	-.20	-.17	.00	-.14	-.13	-.10	-.11	-.11	.10
.1	-.12	-.21	-.02	-.06	-.17	-.13	-.02	-.13	-.12
.2	-.03	-.26	-.04	.04	-.20	-.15	.07	-.16	-.14
.3	.06	-.31	-.06	.13	-.24	-.18	.17	-.19	-.17
.4	.17	-.38	-.09	.24	-.28	-.21	.27	-.21	-.19
.5	.28	-.44	-.13	.35	-.32	-.25	.38	-.24	-.22

	T = 15			T = 30			T = 100		
ρ	r_1	r_2	r_3	r_1	r_2	r_3	r_1	r_2	r_3
0	-.07	-.07	-.07	-.03	-.03	-.03	-.01	-.01	-.01
.1	.02	-.09	-.08	.06	-.04	-.04	.09	-.01	-.01
.2	.12	-.10	-.10	.16	-.05	-.05	.19	-.01	-.01
.3	.22	-.12	-.11	.26	-.06	-.06	.29	-.02	-.02
.4	.32	-.14	-.13	.36	-.06	-.06	.39	-.02	-.02
.5	.43	-.15	-.15	.46	-.07	-.07	.49	-.02	-.02

TABLE 3

Autocorrelations of First-Differenced Error Term
When Original Error Is AR(1) with Selected Values of ρ

ρ	ρ_1^*	ρ_2^*	ρ_3^*
0	-.50	0	0
.1	-.45	-.05	-.00
.2	-.40	-.08	-.02
.3	-.35	-.11	-.03
.4	-.30	-.12	-.05
.5	-.25	-.13	-.06
.6	-.20	-.12	-.07
.7	-.15	-.11	-.07
.8	-.10	-.08	-.06
.9	-.05	-.05	-.04
Random walk	0	0	0

TABLE 4

Autocorrelations of First-Differenced Error Term
When Original Error Is AR(2) with Selected Values of λ_1 and λ_2

$\lambda_2 = -.9$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.05	-.90	.05
.2	.05	-.89	-.22
.4	.15	-.84	-.47
.6	.25	-.75	-.68
.8	.35	-.62	-.81
1.0	.45	-.45	-.86
1.2	.55	-.24	-.78
1.4	.65	.01	-.57
1.6	.75	.30	-.20
1.8	.85	.63	.37

$\lambda_2 = -.7$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.15	-.70	.11
.2	-.05	-.71	-.11
.4	.05	-.68	-.31
.6	.15	-.61	-.47
.8	.25	-.50	-.58
1.0	.35	-.35	-.60
1.2	.45	-.16	-.51
1.4	.55	.07	-.29
1.6	.65	.34	.09

$\lambda_2 = -.5$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.25	-.50	.13
.2	-.15	-.53	-.03
.4	-.05	-.52	-.18
.6	.05	-.47	-.31
.8	.15	-.38	-.38
1.0	.25	-.25	-.38
1.2	.35	-.08	-.27
1.4	.45	.13	-.04

$\lambda_2 = -.3$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.35	-.30	.11
.2	-.25	-.35	.01
.4	-.15	-.36	-.10
.6	-.05	-.33	-.18
.8	.05	-.26	-.22
1.0	.15	-.15	-.20
1.2	.25	0	-.08

$\lambda_2 = -.1$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.45	-.10	.05
.2	-.35	-.17	.00
.4	-.25	-.20	-.06
.6	-.15	-.19	-.10
.8	-.05	-.14	-.11
1.0	.05	-.05	-.06

$\lambda_2 = .1$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.55	.10	-.06
.2	-.45	.01	-.04
.4	-.35	-.04	-.05
.6	-.25	-.05	-.06
.8	-.15	-.02	-.03

$\lambda_2 = .3$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.65	.30	-.20
.2	-.55	.19	-.13
.4	-.45	.12	-.09
.6	-.35	.09	-.05

$\lambda = .5$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.75	.50	-.38
.2	-.65	.37	-.25
.4	-.55	.28	-.16

$\lambda = .7$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.85	.70	-.60
.2	-.75	.55	-.42

$\lambda_2 = .9$

λ_1	ρ_1^*	ρ_2^*	ρ_3^*
0	-.95	.90	-.86

TABLE 5

Autocorrelations of First-Differenced Error Term
When Original Error Is MA(1) with Selected Values of ρ

ρ	ρ_1^*	ρ_2^*	ρ_3^*
0	-.50	0	0
.1	-.44	-.06	0
.2	-.38	-.13	0
.3	-.29	-.21	0
.4	-.17	-.33	0
.5	0	-.50	0

TABLE 6

Autocorrelations of First-Differenced Error Term
When Original Error Is MA(2) with Selected Values of ρ_1 and ρ_2

	$\rho_2 = -.4$			$\rho_2 = -.3$			$\rho_2 = -.2$		
ρ_1	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*
0	-.30	-.40	.20	-.35	-.30	.15	-.40	-.20	.10
.1	-.22	-.50	.22	-.28	-.39	.17	-.33	-.28	.11
.2				-.19	-.50	.19	-.25	-.38	.13
.3							-.14	-.50	.14

	$\rho_2 = -.1$			$\rho_2 = 0$			$\rho_2 = .1$		
ρ_1	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*
0	-.45	-.10	.05	-.50	0	0	-.55	.10	-.05
.1	-.39	-.17	.06	-.44	-.06	0	-.50	.06	-.06
.2	-.31	-.25	.06	-.38	-.13	0	-.44	0	-.06
.3	-.21	-.36	.07	-.29	-.21	0	-.36	-.07	-.07
.4	-.08	-.50	.08	-.17	-.33	0	-.25	-.17	-.08
.5				0	-.50	0	-.10	-.30	-.10
.6							.13	-.50	-.13

	$\rho_2 = .2$			$\rho = .3$			$\rho = .4$		
ρ_1	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*	ρ_1^*	ρ_2^*	ρ_3^*
0	-.60	.20	-.10	-.65	.30	-.15	-.70	.40	-.20
.1	-.56	.17	-.11	-.61	.28	-.17	-.67	.39	-.22
.2	-.50	.13	-.13	-.56	.25	-.19	-.63	.38	-.25
.3	-.43	.07	-.14	-.50	.21	-.21	-.57	.36	-.29
.4	-.33	0	-.17	-.42	.17	-.25	-.50	.33	-.33
.5	-.20	-.10	-.20	-.30	.10	-.30	-.40	.30	-.40
.6	0	-.25	-.25	-.13	0	-.38			
.7	.33	-.50	-.33						

TABLE 7

Monte Carlo Results on Standard Error Estimation in Experiments with N = 50 and T = 10

	OLS		First differences	
	Sample mean of estimated standard error of $\hat{\beta}$	Sample standard deviation of $\hat{\beta}$	Sample mean of estimated standard error of $\hat{\beta}$	Sample standard deviation of $\hat{\beta}$
$\rho = 0$	.049	.047	.060	.080
$\rho = .4$	.046	.048	.047	.054
$\rho = .8$	.035	.047	.027	.030
$\rho = .99$	.0094	.0108	.0060	.0051

TABLE 8

Monte Carlo Results on Central Tendency and Variability of Autocorrelation Estimators in Experiments with N = 50 and T = 10

	r_1	r_2	r_3	r_1^*	r_2^*	r_3^*
$\rho = 0$						
Probability limit of estimator	-.11	-.11	-.10	-.50	0	0
Sample mean (and standard deviation) of estimator	-.11 (.046)	-.11 (.041)	-.11 (.044)	-.49 (.037)	-.00 (.059)	-.01 (.059)
$\rho = .4$						
Probability limit of estimator	.23	-.07	-.17	-.30	-.12	-.05
Sample mean (and standard deviation) of estimator	.22 (.046)	-.08 (.048)	-.20 (.044)	-.30 (.039)	-.11 (.055)	-.05 (.061)
$\rho = .8$						
Probability limit of estimator	.56	.20	-.04	-.10	-.08	-.06
Sample mean (and standard deviation) of estimator	.54 (.036)	.16 (.052)	-.14 (.043)	-.10 (.046)	-.08 (.053)	-.07 (.053)
$\rho = .99$						
Probability limit of estimator	.69	.41	.16	-.005	-.005	-.005
Sample mean (and standard deviation) of estimator	.66 (.036)	.31 (.053)	-.04 (.043)	-.00 (.050)	.01 (.054)	-.00 (.045)

TABLE 9

Monte Carlo Results on Central Tendency and Variability of Autocorrelation Estimators in Experiments with $\rho = .4$ and $T = 6$

	r_1	r_2	r_3	r_1^*	r_2^*	r_3^*
Probability limit of estimator	.10	-.22	-.10	-.30	-.12	-.05
N =50						
Sample mean (and standard deviation) of estimator	.07 (.058)	-.28 (.050)	-.37 (.074)	-.30 (.065)	-.12 (.074)	-.03 (.098)
N = 100						
Sample mean (and standard deviation) of estimator	.08 (.038)	-.28 (.044)	-.39 (.057)	-.30 (.042)	-.11 (.062)	-.06 (.065)
N = 200						
Sample mean (and standard deviation) of estimator	.08 (.029)	-.28 (.025)	-.38 (.039)			

TABLE 10

Monte Carlo Results on Correcting for Serial Correlation in in Experiments with N = 50 and T = 10

	$\rho = .4$		$\rho = .8$	
	Sample mean of estimated standard error of $\hat{\beta}$	Sample standard deviation of $\hat{\beta}$	Sample mean of estimated standard error of $\hat{\beta}$	Sample standard deviation of $\hat{\beta}$
OLS	.046	.048	.035	.047
First differences	.047	.054	.027	.030
Quasi-first differences: $\hat{\rho}$ based on r_1	.049	.049	.029	.030
$\hat{\rho}$ based on r_1^*	.049	.049	.029	.030

Footnotes

1. Ashenfelter and Solon (1982) review some of the sources and uses of longitudinal data.

2. See Mellow (1981) and Mincer (1981) for examples.

3. See Bhargava, Franzini, and Narendranathan (1982), Nickell (1982), Solon (forthcoming), Lillard and Willis (1978), and MaCurdy (1982) for examples.

4. Still another approach, suggested by Kiefer (1980) and Chamberlain (1982), is to stack the equations from different time periods and apply a seemingly-unrelated-regressions estimator with an unrestricted intertemporal covariance matrix. This approach will often be more cumbersome than the other methods discussed in this paper, particularly if the time series are of at least moderate length. For example, a time series length of T=10 would require the estimation of 45 intertemporal covariance parameters. Furthermore, if the other methods discussed here indicate that the error structure is well described as a low-order ARMA process, imposing such a structure in the estimation should improve efficiency in finite samples.

5. The author thanks Takeshi Amemiya for the idea of generating these tables.

6. See Box and Jenkins (1970), pp. 59-61.

7. The α_i and other approximately normal variables used in the experiments were generated by the SAS package's NORMAL function. This function uses a multiplicative congruential method to generate uniformly distributed pseudo-random variables, shuffles the variables, and averages series of 12 uniformly distributed variables to obtain approximately normal pseudo-random variables. See Hammersley and Handscomb (1964) and Quandt (forthcoming) for more detailed discussions of pseudo-random number generation.

8. This standard deviation was computed as the square root of

$$\sum_{j=1}^{50} (\hat{\beta}_j - 1)^2/50 \quad \text{where } j \text{ indexes the replication.}$$

9. These standard deviations were computed as the square root of

$$\sum_{j=1}^{50} (r_j - \bar{r})^2/49 \quad \text{where } j \text{ indexes the replication and } \bar{r} \text{ is the sample}$$

mean over all 50 replications.

References

Ashenfelter, Orley, and Gary Solon. "Longitudinal Labor Market Data: Sources, Uses, and Limitations." What's Happening to American Labor Force and Productivity Measurements?, pp. 109-26. Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research, 1982.

Bhargava, A., L. Franzini, and W. Narendranathan. "Serial Correlation and the Fixed Effects Model." Review of Economic Studies 49 (October 1982): 533-49.

Box, George E. P., and Gwilym M. Jenkins. Time Series Analysis, Forecasting and Control. San Francisco: Holden-Day, 1970.

Chamberlain, Gary. "Panel Data." Working Paper No. 8209, Social Systems Research Institute, University of Wisconsin, April 1982.

Hammersley, J. M., and D. C. Handscomb. Monte Carlo Methods. London: Methuen, 1964.

Kendall, M. G. "Note on Bias in the Estimation of Autocorrelation." Biometrika 41 (December 1954): 403-4.

Kiefer, Nicholas M. "Estimation of Fixed Effect Models for Time Series of Cross-Sections with Arbitrary Intertemporal Covariance." Journal of Econometrics 14 (October 1980): 195-202.

Lillard, Lee A., and Robert J. Willis. "Dynamic Aspects of Earning Mobility." Econometrica 46 (September 1978): 985-1012.

MaCurdy, Thomas E. "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." Journal of Econometrics 18 (January 1982): 83-114.

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

Mincer, Jacob. "Union Effects: Wages, Turnover, and Job Training." Working Paper No. 808, National Bureau of Economic Research, Inc., November 1981.

Nickell, Stephen. "Correcting the Biases in Dynamic Models with Fixed Effects." Working Paper No. 133, Industrial Relations Section, Princeton University, March 1980.

_____. "Biases in Dynamic Models with Fixed Effects." Econometrica 49 (November 1981): 1417-26.

_____. "The Determinants of Occupational Success in Britain." Review of Economic Studies 49 (January 1982): 43-53.

Quandt, Richard. "Computational Problems and Methods." Handbook of Econometrics, ed. Zvi Griliches and Michael D. Intriligator. Amsterdam: North-Holland, forthcoming.

Solon, Gary. "The Effects of Unemployment Insurance Eligibility Rules on Job Quitting Behavior." Journal of Human Resources (forthcoming).