On the Estimation of Panel-Data Models With Serial Correlation When Instruments Are Not Strictly Exogenous

Michael P. Keane
Industrial Relations Center and Department of Economics, University of Minnesota, Minneapolis, MN 55455, and Research Department, Federal Reserve Bank of Minneapolis, MN 55440

David E. Runkle
Research Department, Federal Reserve Bank of Minneapolis, MN 55440, and Department of Finance, Carlson School of Management, University of Minnesota, Minneapolis, MN 55455

In recent years, researchers in many disciplines, including economics, accounting, finance, and marketing, have increasingly relied on panel data to model the behavior of individuals and firms. They have done so because panel data allow them to control for temporally persistent unobserved differences among individuals or firms that in many instances may bias estimates obtained from cross-sections.

Since the original work of Balestra and Nerlove (1966), panel-data econometricians have commonly assumed that the instruments used to identify model parameters are strictly exogenous with respect to the time-varying error component. For example, Hausman and Taylor (1981), Amemiya and MacCurdy (1986), and Breusch, Mizon, and Schmidt (1989) studied models of the form

\[ Y_{it} = X_{it} \beta + \eta_i + \nu_{it}, \tag{1} \]

where \( \eta_i \) is an individual fixed or random effect (i.e., \( \eta_i \) is fixed or random with respect to \( X_{it} \)) for person \( i \) that is invariant over time \( t \). In the fixed-effects case, they proposed an instrumental variable (IV) estimator for (1) that is obtained by taking deviations from individual means (to eliminate \( \eta_i \) from the equation) and then using the orthogonality conditions \( E[(v_{it} - \tilde{v}_i)X_{it}] = 0 \). In the random-effects case, quasi-demeaning is applied to (1) to obtain the orthogonality conditions \( E[(v_{it} - \gamma \tilde{v}_i)X_{it}] = 0 \), where \( 0 < \gamma < 1 \). Of course, in either case the \( X \)'s are not valid instruments unless \( X_{it} \) is strictly exogenous with respect to \( \nu_{it} \); that is, \( E(X_{it} \nu_{is}) = 0 \) for all \( i \) and \( s \).

There are, however, many important cases in which the regressors or other potential instruments are only predetermined with respect to the time-varying error component. In other words, we only have \( E(X_{it} \nu_{is}) = 0 \) for \( t \leq s \). Some examples are models with lagged dependent variables, rational-expectations models in which potential instruments are functions only of variables in the time \( t - 1 \) information set so that there are never any strictly exogenous instruments, and models with predetermined choice variables as regressors, such as hours equations that include children.

Previously, Anderson and Hsiao (1981), Bhargava and Sargan (1983), Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bover (1990), and Ahn and Schmidt (1990) considered models with individual effects and lagged dependent variables as elements of \( X_{it} \). These authors noted that taking deviations (or quasi-deviations) from individual means in such models leads to inconsistency even if \( X_{it} \) is strictly exogenous because \( Y_{it-1} \) is correlated with \( \tilde{v}_i \) by construction. They proposed instead the IV estimator obtained by first-differencing (1) to eliminate the individual effect (or, in the work of Arellano and Bover, the use of deviations from means that are taken over only current and future values) and then including only \( Y_{it-s} \) for \( s \leq t - 2 \) in the instrument set.

In this article, we propose an alternative procedure for IV estimation of models with predetermined but not strictly exogenous instruments that is more generally applicable and has other important advantages over this first-differencing approach. We propose a new estimator for panel data, based on the time series work on forward filtering by Hayashi and Sims (1983), that eliminates any form of serial correlation in models with predetermined instruments yet does not cause parameter estimates to be inconsistent.

In the fixed-effects case, we propose that, following first-differencing, the estimating equation be forward filtered to eliminate any serial correlation in the time-varying error component. Instruments that were only predetermined prior to forward filtering remain valid after forward filtering, so the procedure does not require instruments to be strictly exogenous. Our forward-filtering procedure provides a potential efficiency gain over procedures that only first-difference.

In the random-effects case, the advantages of forward filtering are more important. Of course, one may obtain
consistent estimates of random-effects models that do not contain lagged dependent variables by ignoring the random effects and simply applying ordinary least squares (OLS) or two-tape least squares (2SLS). Consistent but inefficient estimates of the covariance matrix of the parameters are also easy to compute. But compared to OLS or 2SLS, the forward-filtering procedure provides a potential efficiency gain because it eliminates serial correlation. Alternatively, one may first-difference to eliminate random effects, but this entails unnecessary loss of efficiency. The forward-filtering procedure provides a potential efficiency gain by eliminating serial correlation without differencing.

In the case of a lagged dependent variable, the simple procedure of ignoring the random effects and applying OLS will produce inconsistent parameter estimates. 2SLS will only produce consistent parameter estimates if instruments that are uncorrelated with the random effect are available for the lagged dependent variable. This may be difficult in practice because the \( y_{it} \) for all \( s \leq t - 1 \) are correlated with the random effect by construction, ruling them out as potential instruments. The lagged dependent variable model may of course be estimated by first-differencing and using lagged instruments, but, as in the fixed-effects case, application of forward filtering provides a potential efficiency gain if there is serial correlation in \( r_{i,t} \). More important, several of the authors cited previously have noted that dynamic random-effects models can be estimated via generalized least squares (GLS) (or quasi-demeaning) procedures that preserve identification of coefficients on time-invariant regressors so long as a sufficient number of strictly exogenous instruments are available (in particular, instruments that can be used for the lagged dependent variables). By applying our forward-filtering procedure, this condition is relaxed to the requirement that sufficient predetermined instruments be available. Then the \( y_{it} \) for \( s \leq t - 1 \) are in the set of valid instruments.

In applied panel-data work it is, of course, crucial to determine whether fixed effects, random effects, or no individual effects are present. Methods for testing hypotheses concerning the nature of the individual effects were developed by Hausman (1978). Both the theoretical and applied work on such specification testing, however, has universally maintained the assumption of strict exogeneity. In this article, we show how to perform specification tests when instruments are predetermined and not strictly exogenous. We also show how to test for strict exogeneity.

These tests are of practical importance, for two reasons. First, with failure of strict exogeneity, the inconsistent fixed-effects estimates obtained by demeaning will tend to differ from random-effects estimates even when random effects or no individual effects are present. Since the specification tests for fixed effects that have been applied by other researchers are based on the difference between these two estimators, failure of strict exogeneity can falsely lead researchers to conclude that fixed effects are present when they are not. Use of our tests avoids this problem. Second, many researchers have unnecessarily reduced the efficiency of their parameter estimates by using first-differenced estimators in panel-data models with predetermined instruments to avoid the parameter inconsistency that would occur if fixed effects were present. Our tests would result in potentially more efficient estimates if the individual effects were random, because they would indicate that first-differencing is unnecessary.

Finally, we use our estimator based on the forward-filtering procedure to estimate the consumption Euler equation for a simple version of the rational-expectations life-cycle consumption model. We show that instruments are not strictly exogenous in this model and that incorrect application of a specification test that assumes strict exogeneity leads to false acceptance of the fixed-effects hypothesis. The incorrect application of a fixed-effects estimator obtained by demeaning leads to false rejection of the model. We show, however, that with proper application of our forward-filtering estimation procedure the simple life-cycle model cannot be rejected.

1. A REVIEW OF STANDARD PANEL–DATA ESTIMATION METHODS

Suppose that we have a panel of \( N \) people (indexed by \( i \)), for whom we observe the variables \( Y_{it} \) and \( X_{it} \), in each of time \( T \) periods (indexed by \( t \)). Consider the linear model

\[
Y_{it} = X_{it}\beta + \epsilon_{it},
\]

\( i = 1, 2, \ldots, N; t = 1, 2, \ldots, T \) \( (1') \)

where \( \epsilon_{it} \) is a mean-zero residual. Estimating this model is simple if \( X_{it} \) is exogenous and \( \epsilon_{it} \) is homoscedastic and serially uncorrelated \( (E(\epsilon_{it}|X_{it}) = 0, E(\epsilon_{it}^2) = \sigma^2, \) and \( E(\epsilon_{it}\epsilon_{js}) = 0 \) for all \( i \neq j \) or \( t \neq s \)). In that case, the model can be estimated using OLS. The resulting coefficient estimates and standard errors will be consistent.

Researchers have long recognized, however, that the strong conditions necessary to use OLS estimation on this model are not likely to be satisfied in most empirical applications using panel data. In particular, with unobserved individual heterogeneity the errors in Equation \( (1') \) are likely to be correlated across time for each individual, invalidating the assumption that \( E(\epsilon_{it}\epsilon_{js}) = 0 \), for all \( t \neq s \). Balestra and Nerlove (1966) proposed instead the assumption that the error term \( \epsilon_{it} \) can be decomposed into an individual-specific component \( \eta_i \) and the remaining time-varying error component \( \epsilon_{it}, \) where \( E(\eta_i|X_{it}) = 0, E(\eta_i^2) = \sigma_\eta^2, E(\eta_i\eta_j) = 0 \) for all \( i \neq j, E(\epsilon_{it}^2) = \sigma^2, E(\epsilon_{it}\epsilon_{js}) = 0 \) for all \( i \neq j \) or \( t \neq s \), and where the strict exogeneity assumption
\(E(v_{is} | X_{is}) = 0\) for all \(t\) and \(s\) holds. The individual-specific component captures persistent deviations of \(Y_{it}\) from its predicted value, based on the explanatory variables \(X_{it}\). For example, if Equation (1') were explaining individual wages, \(\eta_i\) would represent those aspects of individual ability that are not captured by observable variables such as experience, education, and past employment status.

If the Balestra-Nerlove assumptions are correct, then OLS will provide consistent but inefficient parameter estimates, and the OLS covariance matrix will be inconsistent. (Of course, a correction exists that provides consistent but inefficient standard errors.) However, the GLS estimator

\[
Y_{it} = \gamma \bar{Y}_t = (X_{it} - \gamma \bar{X}_t)\beta + (\varepsilon_{it} - \gamma \bar{\varepsilon}_t),
\]

(2)

where \(\gamma = 1 - \sigma_\eta^2 / \sum_{t=1}^{T} \sigma_{\eta_t}^2\) provides both efficient parameter estimates and consistent estimates of the covariance matrix of the parameters. The Balestra-Nerlove estimator has often been called a random-effects estimator because of its assumption that the individual effect is random with respect to the observed explanatory variables [meaning that \(E(\eta_i | X_{it}) = 0\)].

But if \(E(\eta_i | X_{it}) \neq 0\), then the individual effect is correlated with the explanatory variables and, as Maddala (1971) first explained, neither OLS nor the random-effects estimator would be consistent. In this case, as long as \(X_{it}\) is strictly exogenous with respect to \(v_{it}\), consistent parameter estimates can be obtained by taking deviations from individual means in (1) to obtain the transformed equation

\[
Y_{it} - \bar{Y}_t = (X_{it} - \bar{X}_t)\beta + (\varepsilon_{it} - \bar{\varepsilon}_t),
\]

(3)

where \(\bar{Y}_t\) and \(\bar{X}_t\) are the individual means for \(Y_{it}\) and \(X_{it}\). Note that subtracting individual means eliminates any fixed effect from Equation (1). (Exactly the same results would be obtained by estimating individual-specific intercepts using dummy variables.) Because of the assumption that the individual effect is correlated with the explanatory variables, which leads to the interpretation of the individual effects as individual constants, Equation (2) is often called the fixed-effects estimator. Although the fixed-effects estimates are consistent, that consistency comes at a price: Coefficients cannot be estimated for variables that are constant for each individual across all time periods, and the remaining estimates are likely to have large standard errors.

Although Equations (1'), (2), and (3) may be consistently estimated by OLS if \(X_{it}\) is exogenous, in the more general case \(X_{it}\) is endogenous and (1'), (2), and (3) must be estimated by 2SLS. Suppose we have available instruments \(Z_{it}\) that satisfy \(E(v_{it} | Z_{it}) = 0\). In the case of no individual effects, Equation (1') may be estimated consistently by 2SLS using the instrument \(Z_{it}\). In the fixed- and random-effects cases, however, Equations (2) and (3) cannot generally be estimated consistently using such instruments. This is because the errors of these equations involve the term \(\bar{Y}_t\). This point is crucial. To have \(E(\bar{Y}_t | Z_{it}) = 0\), it is not sufficient to have the contemporaneous no-correlation condition \(E(v_{it} | Z_{it}) = 0\). Rather, we need the much stronger strict-exogeneity condition \(E(v_{it} | Z_{is}) = 0\) for all \(t\) and \(s\).

Unfortunately, there are many cases in which such a strict exogeneity condition will not hold. One example is the case of a lagged dependent variable. Suppose we have the equation

\[
Y_{it} = \gamma Y_{it-1} + Z_{it}\beta + \eta_i + \nu_{it}.
\]

(4)

If demeaning or quasi-demeaning is applied to this equation, then lagged \(Y_{it}\) are not valid instruments because they are correlated with \(\bar{Y}_t\) by construction.

As noted in the introduction, several authors have observed that if Equation (4) is first-differenced, giving

\[
Y_{it} - Y_{it-1} = \gamma (Y_{it-1} - Y_{it-2}) + (X_{it} - X_{it-1})\beta + (\nu_{it} - \nu_{it-1}),
\]

(5)

then the predetermined variable \(Y_{it-1}\) is a valid instrument. This observation raises an important point: There are some transformations of regression equations (e.g., demeaning) that render predetermined variables such as \(Y_{it-2}\) invalid as instruments, leaving only strictly exogenous variables as valid instruments. But there are other transformations (like differencing) that leave predetermined variables valid as instruments. In Section 2, we describe a forward-filtering transformation that also leaves predetermined variables valid as instruments but that has some important advantages over first-differencing.

2. GLS ESTIMATORS WHEN INSTRUMENTS ARE PREDETERMINED BUT NOT STRICTLY EXOGENOUS

Suppose that we want to estimate the equation

\[
Y_{it} = X_{it}\beta + \varepsilon_{it},
\]

(6)

using instruments \(Z_{it}\) that are predetermined but not strictly exogenous; that is, \(E(v_{it} | Z_{is}) = 0\) for all \(s \leq t\), but \(E(v_{it} | Z_{is}) \neq 0\) for all \(s > t\). If the errors in Equation (6) are uncorrelated with the instruments, that equation can be estimated using 2SLS. Given random effects or other sources of serial correlation, however, \(\Omega_F = (I_N \otimes \Sigma_F) = E(\varepsilon \varepsilon')\) will not be diagonal and 2SLS will be inefficient.

We propose a new GLS estimator that remains consistent if the instruments in Equation (6) are predetermined rather than strictly exogenous but that is potentially more efficient than 2SLS when the errors in Equation (6) are serially correlated. The estimator is constructed by first obtaining a consistent estimate of
\[ \Sigma_{TS}^{-1} \] and then computing its upper-triangular Cholesky decomposition, which we will call \( \hat{P}_{TS} \). Then premultiply Equation (6) by \( \hat{Q}_{TS} = (I_N \otimes \hat{P}_{TS}) \) and estimate the transformed equation using 2SLS, while still using the original instruments. We call this new estimator \( \hat{\beta}_{KR} \).

\[
\hat{\beta}_{KR} = (X'\hat{Q}_{TS}Z(Z'Z)^{-1}Z'\hat{Q}_{TS}X)^{-1} \\
\times X'\hat{Q}_{TS}Z(Z'Z)^{-1}\hat{Q}_{TS}Y.
\] (7)

This estimator does not impose an equicorrelation assumption on \( \Omega \) as the fixed-effects and random-effects estimators do.

This new estimator is derived by applying the insights of Hayashi and Sims (1983) concerning time series models to panel data. They showed that, if a time series equation has serially correlated errors and predetermined instruments, serial correlation can be eliminated by a transformation that makes the transformed dependent variable for time \( t \) a linear combination of the values of the original dependent variable for time periods \( t \) and later. So long as the dating of the instruments is left unchanged, this transformation preserves the orthogonality conditions implied by the time series model and yields consistent and potentially more efficient estimates of the parameters. Note that it was exactly these considerations that led Arellano and Bover (1990) to suggest forward demeaning (i.e., taking deviations from means taken only over current and future values) in models with predetermined but not strictly exogenous instruments. Unlike premultiplication by \( \hat{Q}_{TS} \), however, their procedure does not eliminate all forms of serial correlation. Thus forward filtering has more general applicability.

In time series models, the covariance matrix \( \Omega \) must be tightly parameterized because the number of elements in \( \Omega \) is much larger than the number of observations available in the time series model. In a panel-data model, however, the number of unique elements in \( \Omega_{TS} \)—namely, \( T(T + 1)/2 \)—is usually far smaller than \( N \), so \( \Omega_{TS} \) can be estimated directly. Denoting the residuals for individual \( i \) in the 2SLS equation as \( \hat{u}_{it} \), a consistent estimate of \( \Sigma_{TS} \) is obtained by constructing

\[
\hat{\Sigma}_{TS} = \frac{1}{N} \sum_{i=1}^{N} \hat{u}_{it}' \hat{u}_{it}.
\]

As a result, a consistent estimate of \( \Omega_{TS} \) would be

\[
\hat{\Omega}_{TS} = (I_N \otimes \hat{P}_{TS}).
\]

Note that we did not have to specify the covariance structure of the residuals for each individual to compute the estimates of \( \Sigma_{TS} \) that we will use in computing \( \hat{\beta}_{KR} \). (Of course, we could parameterize \( \Sigma_{TS} \) as a Toeplitz matrix or include individual effects as the only source of serial correlation.)

The key difference between our GLS estimator and those suggested by Hausman and Taylor (1981), Amemiya and MacCurdy (1986), and Breusch et al. (1989), is the form of \( \hat{P}_{TS} \) that we use to create \( \hat{Q}_{TS} \). Those authors use \( \Sigma^{-1/2}_{TS} \) to premultiply Equation (1), where \( \Sigma^{-1/2}_{TS} = Q + \delta P_{TS} \), \( \delta^2 = \sigma^2(\alpha^2 + \tau^2) \), \( \sigma \) is the \( NT \times N \) matrix of individual dummy variables, \( P_{TS} = v'(v'v)^{-1}v \), and \( Q_{TS} = I - P_{TS} \). This transformation leads to Equation (2). As was noted previously, the error in Equation (2) for person \( i \) in time \( t \) is not orthogonal to \( Z_{it} \), because that error is a linear combination of the errors for person \( i \) in all different time periods, and \( E(\epsilon_{it} | Z_{it}) \neq 0 \) for all \( i \). Therefore, the estimators proposed by those authors will lead to inconsistent parameter estimates if the instruments are predetermined and not strictly exogenous, because the transformed equation will violate the orthogonality conditions of the original equation.

Although our new estimator is certainly useful when there are random effects, it may also be used to obtain more efficient estimates of the fixed-effects first-difference model. Suppose that the instruments in Equation (6) are predetermined but that the individual effects in that equation are also correlated with the instruments. In this case, we would want to estimate a first-differenced version of Equation (5)—namely,

\[
Y_{it} - Y_{i(t-1)} = (X_{it} - X_{i(t-1)}) \beta + (\nu_{it} - \nu_{i(t-1)}),
\]

\( i = 1, 2, \ldots, N \); \( t = 1, 2, \ldots, T - 1 \), (8)

using instruments from time \( t - 1 \) and before. As we noted previously, (8) may be estimated consistently by 2SLS even when instruments are not strictly exogenous. As is obvious from Equation (8), however, the residuals in the transformed equation are serially correlated. In fact, the serial correlation in this equation is (moving average) MA(1) if \( \nu_{it} \) is itself serially uncorrelated and will take the form of a higher order autoregressive moving average process if \( \nu_{it} \) is serially correlated. Thus \( \hat{\beta}_{KR} \) may be more efficient than the 2SLS estimator. To obtain \( \hat{\beta}_{KR} \), simply estimate \( \Sigma_{TS} \) and premultiply Equation (8) by \( \hat{Q}_{TS} = (I_N \otimes \hat{P}_{TS}) \). Then estimate the transformed equation by 2SLS using the original instruments.

### 3. Specification Tests

There are two problems facing econometricians who want to estimate a panel-data model: First, they must determine whether the instruments are merely predetermined or whether they are strictly exogenous. Second, they must determine whether the individual effects are correlated with the instruments.

We can test whether the instruments are strictly exogenous by comparing the results of two different estimators: a first-difference estimator and a fixed-effects estimator. Consider the first-difference equation (8). If we use instruments from time \( t - 1 \) or before in estimating this equation, we can consistently estimate \( \beta \) whether or not the instruments are strictly exogenous. Since any potential individual effect has been eliminated by differencing, there will also be no problem
caused by correlation of an individual effect with the instruments. In contrast, the fixed-effects estimator (3) will give a consistent estimate of $\beta$ only if the instruments are strictly exogenous. As previously noted, however, if the instruments are merely predetermined, the fixed-effects estimator will be inconsistent.

Since the probability limits of the estimates of $\beta_{FE}$ and $\beta_{FD}$ differ only if the instruments are not strictly exogenous, one simple specification test for strict exogeneity is to compute the statistic $(\hat{\beta}_{FE} - \hat{\beta}_{FD})'(V(\hat{\beta}_{FE} - \hat{\beta}_{FD}))^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{FD})$, which should be distributed asymptotically as a $\chi^2_k$ random variable if $\beta_{FD}$ and $\beta_{FE}$ each contain $k$ parameters and the instruments are strictly exogenous.

Since no clear efficiency comparison is possible between the fixed-effects estimator and the first-difference estimator, computing $V(\hat{\beta}_{FE} - \hat{\beta}_{FD})$ is not as simple as it is for Hausman specification tests. Denote the residuals for individual $i$ in the fixed-effects equation as $\tilde{U}_{FE}$ and denote the residuals for individual $i$ in the first-difference equation as $\tilde{U}_{FD}$. Then consistent estimates of the covariances and cross-covariances of the residuals for each individual in those equations are given by

$$
\Sigma_{FE} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{FE}^{\prime} \tilde{U}_{FE},
$$

$$
\Sigma_{FD} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{FD}^{\prime} \tilde{U}_{FD},
$$

and

$$
\Sigma_{FDFD} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{FE}^{\prime} \tilde{U}_{FD}.
$$

As a result, consistent estimates of the covariance and cross-covariance matrices of the residuals of these two equations would be $\tilde{\Omega}_{FE} = (I_N \otimes \Sigma_{FE})$, $\tilde{\Omega}_{FD} = (I_N \otimes \Sigma_{FD})$, and $\tilde{\Omega}_{FDFD} = (I_N \otimes \Sigma_{FDFD})$. Given these estimates, we can compute $V(\hat{\beta}_{FE} - \hat{\beta}_{FD})$ as

$$
V(\hat{\beta}_{FE} - \hat{\beta}_{FD}) = (X_{FE}'Z(Z'Z)^{-1}Z_{FE})^{-1}(X_{FE}'Z(Z'Z)^{-1}Z_{FE})' \times (X_{FE}'Z(Z'Z)^{-1}Z_{FD})^{-1}
$$

$$
- (X_{FD}'Z(Z'Z)^{-1}Z_{FD})^{-1}(X_{FD}'Z(Z'Z)^{-1}Z_{FD})' \times (X_{FD}'Z(Z'Z)^{-1}Z_{FE})^{-1}
$$

$$
- (X_{FD}'Z(Z'Z)^{-1}Z_{FD})^{-1}(X_{FD}'Z(Z'Z)^{-1}Z_{FD})' \times (X_{FE}'Z(Z'Z)^{-1}Z_{FE})^{-1}
$$

$$
+ (X_{FD}'Z(Z'Z)^{-1}Z_{FD})^{-1}(X_{FD}'Z(Z'Z)^{-1}Z_{FD})' \times (X_{FE}'Z(Z'Z)^{-1}Z_{FE})^{-1}.
$$

With this estimate of the covariance matrix, we can compute the specification test to determine whether the instruments are strictly exogenous or merely predetermined. (Runkle [1991] showed how to construct an estimate of $Z'\Omega_{FD}Z$ and similar terms that are robust to conditional heteroscedasticity.) Simple extensions to deal with missing data are also available.

Since the fixed-effects estimator would unnecessarily reduce efficiency if the individual effects were unrelated with the regressors, it is important to test directly the assumption that $E(\eta_i|X_{i,t}) = 0$. Hausman (1978) first developed such a test. Assuming strict exogeneity, he noted that, under the null hypothesis that $E(\eta_i|X_{i,t}) = 0$, both the random-effects and the fixed-effects estimators are consistent, but the random-effects estimator is efficient. Under the alternative hypothesis that $E(\eta_i|X_{i,t}) \neq 0$, the fixed-effects estimator is consistent, but the random-effects estimator is inconsistent. He proposed a test of the null hypothesis based on the difference between the random-effects and the fixed-effects parameter estimates.

It is necessary to know the outcome of the strict exogeneity test, however, to properly test for correlation between individual effects and the instruments. If the instruments are strictly exogenous, then Hausman and Taylor’s (1981) specification tests can be used to determine whether there is an individual effect that is correlated with the instruments. But if the instruments are merely predetermined, the Hausman and Taylor tests may lead one to believe that fixed-effects are present when in fact they are not.

With predetermined but not strictly exogenous instruments, a valid specification test for fixed effects can be based on the difference between the common parameters from a two-stage first-difference estimator and a 2SLS estimator of Equation (6). Under the null hypothesis that there is no individual effect that is correlated with the instruments, both $\hat{\beta}_{FD}$ and $\hat{\beta}_{TS}$ will be consistent even when instruments are merely predetermined. If the null is incorrect, then $\hat{\beta}_{FD}$ will still be consistent, but $\hat{\beta}_{TS}$ will not. Consequently, we can test the null by forming the statistic $(\hat{\beta}_{TS} - \hat{\beta}_{FD})'(V(\hat{\beta}_{TS} - \hat{\beta}_{FD}))^{-1}(\hat{\beta}_{TS} - \hat{\beta}_{FD})$, which should be distributed asymptotically as a $\chi^2$ random variable under the null.

To construct the test, a consistent estimate of $V(\hat{\beta}_{TS} - \hat{\beta}_{FD})$ is needed. Let the residuals for individual $i$ in the 2SLS equation be denoted as $\tilde{U}_{TS}$, and let the residuals for individual $i$ in the first-difference equation be denoted as $\tilde{U}_{FD}$. Then consistent estimates of the covariances and cross-covariances of the residuals for each individual in those equations are

$$
\Sigma_{TS} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{TS}^{\prime} \tilde{U}_{TS},
$$

$$
\Sigma_{FD} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{FD}^{\prime} \tilde{U}_{FD},
$$

and

$$
\Sigma_{TSFD} = \frac{1}{N} \sum_{i=1}^{N} \tilde{U}_{TS}^{\prime} \tilde{U}_{FD}.
$$
Thus consistent estimates of the covariance and cross-
covariance matrices of the residuals of these two equations 
would be $\hat{\Omega}_{TS} = (I_N \otimes \Sigma_{TS}), \hat{\Omega}_{TD} = (I_N \otimes 
\Sigma_{TD}),$ and $\hat{\Omega}_{TSDT} = (I_N \otimes \Sigma_{TSDT})$. Given these 
estimates, we can compute an estimate of $(V(\hat{\beta}_{TS} - \hat{\beta}_o))$ 
that is guaranteed to be positive definite as follows:

$$
V(\hat{\beta}_{TS} - \hat{\beta}_o)
= (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}(X_{i\tau} Z(Z, Z)^{-1} Z \hat{\Omega}_{TS} Z(Z, Z)^{-1} Z X_{i\tau})
\times (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}
- (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}(X_{i\tau} Z(Z, Z)^{-1} Z \hat{\Omega}_{TD} Z(Z, Z)^{-1} Z X_{i\tau})
\times (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}
- (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}(X_{i\tau} Z(Z, Z)^{-1} Z \hat{\Omega}_{TSDT} Z(Z, Z)^{-1} Z X_{i\tau})
\times (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}
+ (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}(X_{i\tau} Z(Z, Z)^{-1} Z \hat{\Omega}_{TSDT} Z(Z, Z)^{-1} Z X_{i\tau})
\times (X_{i\tau} Z(Z, Z)^{-1} Z X_{i\tau})^{-1}.
$$

This estimate of the covariance matrix will allow us to 
compute the specification test to determine whether 
there is an individual effect that is correlated with 
instrumentals. (Here, as previously, a heteroscedastic-
consistent version is available by using the estimator 
for $Z^t \Omega_{TS} Z$, and similar terms, suggested by Runkle 
[1991].)

4. EMPIRICAL EXAMPLE: TESTING THE 
RATIONAL-EXPECTATIONS-LIFE-CYCLE 
CONSUMPTION MODEL

The following example will illustrate why failing to 
determine whether instruments are strictly exogenous 
or merely predetermined can cause problems with sta-
tistical inference in panel-data models. Several recent 
articles have used panel data to test the permanent-
inecome hypothesis. Among them are those of Shapiro 
permanent-income hypothesis states that consumption growth from period $t$ to period $t + 1$ should depend only 
on the real-interest rate from period $t$ to period $t + 1$, these authors have all estimated some version of the equation

$$
\Delta C_{i,t+1} = \ln(C_{i,t+1}) - \ln(C_{i,t})
= \beta_0 + \beta_1 r_{i,t} + \varepsilon_{i,t+1},
$$

$i = 1, 2, \ldots, N; t = 1, 2, \ldots, T - 1, (9)$

where $C_{i,t}$ is the level of real consumption for person $i$ in period $t$ and $r_{i,t}$ is the one-period after-tax real interest 
rate for person $i$ from period $t$ to period $t + 1$. (Of 
course, this equation may also include demographic 
factors such as age.) The variable $r_{i,t}$ is uncertain because 
the price level at time $t + 1$ is unknown at time $t$. 
Therefore, Equation (9) must be estimated using some 
IV method. A list of valid instruments for estimating 
Equation (9) can be found by invoking the assumption of 
Rational expectations—namely, $E(\varepsilon_{i,t+1} | I_t) = 0$, where 
$I_t$ is the information available to person $i$ at 
time $t$.

If there is no serial correlation in $\varepsilon_{i,t+1}$, Equation (9) 
can be estimated using 2SLS. But serial correlations 
could exist if there were an error in the measurement of $\ln(C_{i,t})$—an unnecessary complication here (see Runkle 
1991)—or if there were persistent individual differences in 
the discount rate. Such differences in the discount rate 
would cause individual-specific variations in $\beta_0$, 
which would cause a persistent individual effect in $\varepsilon_{i,t+1}$.

If there is a persistent individual effect $\eta_i$ in $\varepsilon_{i,t+1}$ in 
Equation (9), the standard 2SLS estimator will be in-
consistent or, at best, inefficient, depending on whether 
$\eta_i$ is a fixed or a random effect. But standard GLS 
transformations cannot be used to obtain consistent and 
efficient estimates. Because of the rational-expectations 
assumption needed to estimate Equation (9), both the 
GLS random-effects estimator and the fixed-effects esti-
mator will be inconsistent because the available instru-
ments are predetermined, rather than strictly ex-
gogenous. Specifically $\varepsilon_i$ is correlated with elements of 
all the information sets $I_s$ for $s = 1, \ldots, T - 1$, so 
no valid instruments are available to estimate a standard 
GLS transformed version of (9). It has long been noted in the time-series literature (for example by Hansen and Hodrick 1980) that 
conventional GLS estimators of linear rational-
expectations models will produce inconsistent parameter 
estimates because those estimators violate some of the 
orthogonality conditions imposed by rational expectations. For exactly the same reasons, the IV GLS 
estimators proposed by Hausman and Taylor (1981), 
Amemiya and MaCurdy (1986), and Breusch et al. (1989) will be inconsistent for estimating Equation (9). All of 
those estimators would start by quasi-demmeaning the 
equation, as in the work of Maddala (1971). Then they 
would use an IV estimator that requires all instruments to 
be strictly exogenous. But rational-expectations models 
have no such instruments; rational expectations imposes 
only predeterminedness. And the GLS transformations 
that all of the estimators employ would violate the 
orthogonality conditions imposed by Equation (9) 
because those transformations assume that a linear com-
bination of current and past errors is orthogonal to 
current instruments. That condition is not satisfied in 
Equation (9). For the same reason, a standard fixed-effects esti-
mator will yield inconsistent estimates of Equation (9). 
(A similar point was made by Chamberlain [1984] and 
Runkle [1991]). Writing down a fixed-effects version of 
Equation (9) shows why this is true. That equation is 

$$
\Delta C_{i,t+1} - \bar{\Delta} C_i = \beta_i (\bar{r}_{i,t} - \bar{r}) + \varepsilon_{i,t+1} - \bar{\varepsilon}_i,
$$

$i = 1, 2, \ldots, N; t = 1, 2, \ldots, T - 1, (10)$
which would be estimated using instruments from time
\( t \) and before. Unfortunately, the error term for person
\( i \) in period \( t \) in this transformed equation, \( \hat{e}_{i,t+1} = \hat{e}_i \),
is a function of the errors for person \( i \) in each period
in Equation (9). As a result, the errors in the trans-
formed equation are not orthogonal to the proposed
instruments, and the parameter estimates will be in-
consistent.

Thus, in rational-expectations panel-data models,
neither conventional random-effects estimates nor con-
ventional fixed-effects estimators will be consistent,
because instruments are predetermined rather than strictly
exogenous. Furthermore, the specification tests
described by Hausman (1978) and Hausman and Taylor
(1981) will give misleading results. Since each of the
estimators used in those tests will be inconsistent when
estimating rational-expectations panel-data models, those
tests will also be inconsistent.

Because of these problems, we have reexamined the
rational-expectations–life-cycle consumption model us-
ing the new estimation methods and specification tests
described in Sections 2 and 3. As with all of the studies
we have cited, we use data from the Michigan Panel
Study on Income Dynamics (PSID). Our sample in-
cludes 3,762 observations on 627 households surveyed
between 1975 and 1982. For a description of our data-
screening criteria and an explanation of our constructed
variables, see the Appendix. These computations closely

We estimate a modified version of Equation (9),
\[
\ln(C_{i,t}) - \ln(C_{i,t}) = \beta_0 + \beta_1 r_{i,t} + \beta_2 \text{age}_{i,t} + \epsilon_{i,t+1},
\]
\( i = 1, 2, \ldots, N; t = 1, 2, \ldots, T - 1, \) \( (9') \)
where \( C_{i,t} \) is the level of real consumption for household
\( i \) in period \( t \), and \( r_{i,t} \) is the one-period after-tax real
interest rate for household \( i \) from period \( t \) to period \( t + 1 \). In the PSID, the only measure of consumption is
food consumption. We use that as our measure of con-
sumption.

We examine three sets of results in this example.
First, we test to see whether the instruments for esti-
mating Equation (9') are predetermined or strictly ex-
genous. Second, we test whether there is an individual
effect that is correlated with the instruments. Finally,
we examine the implications of these tests for previous
work testing for liquidity constraints using panel data.
All reported standard errors and test statistics are cor-
corrected for conditional heteroscedasticity and serial
 correlation. (For a discussion of the source of the serial
correlation in this model, see Runkle [1991].)

Table 1 shows the results of estimating Equation (9').

\begin{table}
<table>
<thead>
<tr>
<th>Estimator</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SLS</td>
<td>.103</td>
<td></td>
<td>096</td>
<td></td>
</tr>
<tr>
<td>FE</td>
<td></td>
<td>1.551</td>
<td>.719</td>
<td>439</td>
</tr>
<tr>
<td>FE</td>
<td>(.145)</td>
<td>(.419)</td>
<td>(.233)</td>
<td>(.139)</td>
</tr>
<tr>
<td>FE</td>
<td>-.002</td>
<td>-.022</td>
<td>.063</td>
<td>-.002</td>
</tr>
<tr>
<td>FE</td>
<td>(.0003)</td>
<td>(.007)</td>
<td>(.003)</td>
<td>(.0002)</td>
</tr>
</tbody>
</table>

NOTES: Standard errors are in parentheses. The number of observations is 3,762.

Our specification test for strict exogeneity rejects ex-
genous. If we compute that test based on the differ-
ence between the parameter estimates of the fixed-
effects and the first-difference models, we find that
the value of the test statistic is 12.05. Under the null
hypothesis of exogeneity, that statistic should be distrib-
uted asymptotically as a \( \chi^2 \) random variable. Thus we
must reject the null hypothesis of exogeneity. This means
that it is \textit{not} legitimate to estimate any version of Equa-
tion (9') using a standard fixed-effects estimator.

Comparing the 2SLS estimates with the first-differ-
ence estimates allows us to test whether there is an
individual effect that is correlated with the instruments.
The estimates appear quite close, and a specification
test reveals no significant difference between the two
sets of parameters. If we compute our specification test
based on the difference between the common param-
eters in the 2SLS and first-difference estimators, the
value of the statistic is .55. Under the null hypothesis
of no correlation, that test statistic is distributed asym-
ptotically as a \( \chi^2 \) random variable. Therefore, we cannot
reject the null hypothesis of no correlation between the
individual effect and the instruments. This means that
it is unnecessary to use a first-difference estimator for
estimating (9') and that severe efficiency losses will
result if one is used. The difference in the standard errors
between columns 1 and 3 shows how severe those losses
are.

Note that the failure of strict exogeneity causes the
2SLS and two-stage fixed-effects estimates to be quite
different even though no fixed effects are present. Given
this type of bias in two-stage fixed-effects estimates,
traditional Hausman–Taylor type specification tests
would lead one to falsely conclude that fixed effects are
present.

The results of these two specification tests also sug-
gest that our new estimator is appropriate to use in this
case and that it may increase efficiency. A comparison of
columns 1 and 4 shows the efficiency gains from using
\( \hat{\beta}_{KR} \). The standard errors are smaller for every coef-
cient using the new estimator than using 2SLS. The
reductions in the estimated standard deviations of the
coefficients range from 7% to 13%.
We have assumed here, as does all of the literature, that there are no time-period-specific error components. If such error components were present, they would cause problems for our new estimator, but not for the specification tests. Runkle (1991), however, failed to reject the hypothesis that there were no such time-period-specific errors in this model, so our assumption seems appropriate.

Although Table 1 and the specification tests showed the statistical importance of the assumption of exogeneity, we have not yet demonstrated its economic importance. We do that now by considering tests for liquidity constraints using a second modification to Equation (9). This final equation, which we will call Equation (9'), is as follows:

\[
\ln(C_{i,t}) - \ln(C_{i,t-1}) = \beta_0 + \beta_1 t + \beta_2 \ln(\frac{Y_{i,t}}{Y_{i,t-1}}) + \epsilon_{i,t-1}
\]

\(i = 1, 2, \ldots, N; t = 1, 2, \ldots, T - 1.\) \( (9') \)

Both Runkle (1991) and Zeldes (1989) estimated Equation (9') using a sample that is similar to the one used in this article. Table 2 reports estimates of Equation (9') using \(\hat{\beta}_{KR}\), which eliminates serial correlation by transforming Equation (9'), as well as estimates obtained using a standard two-stage fixed-effects estimator. The difference in the results is striking. In fact, the value of the \(\chi^2\) test statistic comparing the common coefficients is 11.07. Thus we must reject the hypothesis that the common parameters are the same.

The inconsistent fixed-effects estimates suggest that lagged income is significant, both statistically and economically, in explaining consumption growth; that is, there are important liquidity constraints. The results imply that consumption grows faster from this period to next period if this period's income is lower, because an inability to borrow constrains today's consumption. These estimates are qualitatively similar to those in Zeldes's (1989) article, but the estimated effect of income on consumption growth is even larger than the one he found. The consistent estimator \(\hat{\beta}_{KR}\), however, shows that lagged income is neither statistically nor economically significant in explaining consumption growth. These results are similar to those of Runkle (1991). When coupled with the results in Table 1 this suggests that one possible reason for the difference between Zeldes's and Runkle's estimates stems from Zeldes's use of a fixed-effects estimator in a case that yields inconsistent parameter estimates.

5. CONCLUSION

This article developed two new specification tests for panel-data models. The first is a test to determine whether a panel-data model with individual-specific effects can be estimated using conventional estimators that assume that all instruments are strictly exogenous. The second is a test for the presence of individual fixed effects that is valid even when instruments are predetermined but not strictly exogenous. It also developed a new estimator that may yield more efficient estimates for panel-data models when instruments are predetermined but not strictly exogenous. Our empirical example demonstrated the importance of this work for determining the validity of the permanent-income hypothesis.

APPENDIX: DATA CONSTRUCTION

It is important to understand both the criteria used to exclude specific observations and the methods used to construct variables used in the empirical analysis. As mentioned in the text of the article, the data for this study come from the PSID. The sample used in the following analysis consists of a balanced panel of 3,762 observations on 627 households between 1975 and 1982. Because the model is estimated in first differences and we use lagged variables, eight years of data yield six annual observations per family. Since this article is concerned with consumption, our criteria for sample selection minimize the noise in the consumption data. We included an observation for a household only if all data on consumption and income were available for that period. Since all data records were kept for heads of households, measurement error in the consumption series would occur if the head of a household divorced, stayed single for a year, and remarried. Therefore, if a couple married or divorced in a given year, we discarded the data for that year and treated the resulting household as a new household. To reduce errors in the measurement of income, we excluded farmers and self-employed heads of households. Since the selection is based on variables exogenous to the estimated equations, no sample selection bias will occur.

The most important household variables used in this study were food consumption, disposable income, the annual number of hours worked by the household, and the household's after-tax real interest rate. Annual hours worked are provided directly in the survey, but all of the other variables must be computed from data in the survey.

Real food consumption was computed as the sum of real food consumption at home and real food consumption away from home. The nominal data for each element of food consumption was deflated by the ap-

<table>
<thead>
<tr>
<th>Table 2. The Results of Estimating Equation (9')</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimator</strong></td>
</tr>
<tr>
<td>(\beta_0)</td>
</tr>
<tr>
<td>(\beta_1)</td>
</tr>
<tr>
<td>(\beta_2)</td>
</tr>
<tr>
<td>(\beta_3)</td>
</tr>
</tbody>
</table>

**NOTE:** Standard errors are in parentheses. The number of observations is 3,762.
propriate Consumer Price Index (CPI) component. Since the food-consumption data refer to consumption during a week in the first quarter of the survey year, the average CPI component for the first three months of the year was used to deflate the nominal quantities. The net cash value of food stamps was included in the nominal value of food consumption at home.

Disposable income was computed as reported family-unit income plus the net cash value of food stamps minus federal income and Social Security taxes paid by the householders. Taxes are computed for both husband and wife if both are present in the family. Social Security taxes are computed from published Social Security tax schedules and reported labor income.

The after-tax interest rate for each household was computed by multiplying the interest rate by \((1 - \theta_{i,t})\), where \(\theta_{i,t}\) is the marginal tax rate for household \(i\) in period \(t\). The average annual passbook savings rate for the year before the panel interview was used as the interest rate. The real after-tax interest rate for each household was computed by subtracting the ex post inflation rate from the after-tax interest rate.

For the equations in Table 1, the instrument list includes a constant, the householder's hours worked in period \(t - 2\), the natural log of the family's disposable income in period \(t - 2\), and the value of the after-tax real interest rate for passbook and T-bill interest rates in period \(t - 2\). This common instrument list is chosen because all information must be from period \(t - 2\) or before the first-difference estimator to be consistent.

For the equations in Table 2, the instrument list includes a constant, the householder's age, the householder's hours worked in period \(t - 1\), the natural log of the family's disposable income in period \(t - 1\), the value of the after-tax real interest rate for passbook and T-bill interest rates in period \(t - 1\), and the log-difference of the family's disposable income in periods \(t - 1\) and \(t - 2\). The instrument list is different from that used for Table 1 because the information constraints are not as strict as those for estimating the first-difference model in Table 1.

[Received January 1991. Revised August 1991.]

REFERENCES


Reply

Michael P. Keane and David E. Runkle

We wish to thank all of our discussants for their thoughtful commentary on our article. Their comments have helped us to organize our thoughts about some important issues in panel-data inference. In our response, we want to answer two questions raised by the discussants: First, what are the most powerful specification tests to use for panel data? Second, what are the most efficient estimators to use for panel data? We believe that the answer to the first question is more important, because incorrect specification may lead to inconsistent parameter estimates, whereas choice among estimators, for a given specification, will only affect efficiency.

1. POWER OF SPECIFICATION TESTS IN PRACTICE

In our article, we have proposed a specification test for strict exogeneity based on comparison of the fixed-effects and first-difference estimators and a test for individual effects based on comparison of the two-stage least squares (2SLS) and first-difference estimators. In each case, we construct the covariance matrix of the difference of the estimators by using the sample residuals.

In their discussion, both Hayashi and Schmidt, Ahn, and Wybowski (SAW) suggest alternative specification tests based on estimators that use additional orthogonality conditions beyond those we use. Specifically, they argue that an efficient estimator could be constructed by applying the generalized method of moments (GMM) to our Equation (1), with all of the available instruments (i.e., all instruments that are predetermined at time t) being used. They suggest that specification tests based on this estimator would be simpler than the tests we propose because the covariance matrix for the tests could be constructed using Hausman's method of subtracting the covariance matrix of the efficient estimator from the covariance matrix of the inefficient estimator.

Although under ideal circumstances it is true that the tests they suggest would be both more powerful and simpler than the ones we suggest, there are four prac-
tical reasons to prefer our tests. First, since the estimator they propose uses all predetermined instruments available at time $t$, including those dated all the way back to the first observation in the panel, each of the tests that they propose requires that there are no missing observations for any individual in the sample. By contrast, our tests require, at most, two consecutive observations for each individual. (Of course, if the model contains lagged dependent variables or regressions then additional periods are necessary.) Since missing observations are extremely frequent in most panel-data sets, using their tests would greatly restrict the number of observations that could be used in conducting specification tests. In fact, in the six panel-data sets with which we are familiar (PSID, NLS Young Men, NLS Youth, ASANBER forecasts, Compustat, IBES Forecast Database), eliminating all individuals with missing data would leave only 10%–20% as many observations available for the tests that they propose as for the tests that we advocate. As a result, in practice their tests are not likely to be as powerful as ours.

Second, even if there were no missing observations, the additional orthogonality conditions they suggest may generate substantial bias in parameter estimates obtained from finite samples. Although in general more orthogonality conditions increase efficiency, Tauchen’s (1986) work on the small-sample properties of GMM estimators suggests that additional orthogonality conditions may increase bias in finite samples. Note, however, that Tauchen’s results are for much smaller samples than those typically used in panel-data work. Thus further work is needed to assess the practical importance of this problem.

Third, in our specification test for fixed effects with predetermined instruments, it is not our failure to use all lagged instruments that generates the increased complexity of our covariance matrix estimator. Our tests are based on comparisons of 2SLS and first-difference estimators. Since it is known that the 2SLS estimator is efficient relative to the first-difference estimator, Hausman’s simplified form may be used. However, we believe that great caution should be used in applying this simplified form. There is no guarantee that the Hausman form of the covariance matrix will be positive definite in finite samples. In fact, applied researchers often find that it is not positive definite. The covariance matrix estimators we propose are guaranteed to be positive definite and, given recent advances in econometric software, are relatively simple to compute.

Finally, our tests are based on estimators that are easy to compute using existing econometric software. Two-stage least squares estimators, fixed-effects estimators, and first-difference estimators are easy to compute for panel-data models. Therefore, it is likely that applied researchers could compute our tests without much extra effort. It would be much more onerous to compute the specification tests proposed by Hayashi and SAW.

2. EFFICIENCY

By far the majority of the discussants’ comments address the new estimator based on forward filtering that we propose. All of the discussants propose alternative estimators to the one we suggest and argue that the estimators they propose are more efficient. Among the discussants there is also concern about the consistency of the estimator and about whether we are merely repeating previous researchers’ results.

Although we still believe that our estimator provides an easy alternative to 2SLS that will often increase efficiency, we think that the entire debate about efficiency is likely to have limited practical importance, because of the missing-data problems discussed earlier. Our new estimator, as well as all of those mentioned by the discussants, generally requires that there are no missing observations for any individual in the sample. Since missing-data problems are likely to be serious in most panels, it is likely that the 2SLS estimator will often be more efficient than either our estimator or those mentioned by the discussants, simply because the 2SLS will use 5–10 times as many observations. There is an important special case in which our forward-filtering procedure may prove especially valuable, however. That is when the errors have an MA($K$) structure. In that case, our estimator requires only $K$ additional periods of available data beyond that needed for 2SLS, rather than complete data. Thus, for small $K$, it may provide important efficiency gains over any of the alternatives.

Assuming complete data, and given their statistical assumptions, the estimator that Hayashi and SAW propose would be efficient relative to the estimator based on forward filtering that we propose in the article (although theirs would involve an increased computational burden and nonstandard software). With more realistic statistical assumptions, however, there are estimators that are efficient relative to those proposed by Hayashi and SAW. As we shall describe, in some instances it is necessary to use forward filtering to construct these estimators.

If potential small-sample bias problems arising from use of large instrument sets can be ignored, more efficient estimates are possible by adding additional orthogonality conditions beyond those suggested by Hayashi and SAW. Note that both Hayashi and SAW assume that there is a fixed set of instruments for each time period. Using the SAW notation, at time $t$, the instruments for individual $i$ are $Z_{it} = (Z_{i1},Z_{i2},\ldots,Z_{ik})$. This assumes that the only information that we have is that $E(e_{i1}\mid Z_{i}) = 0$ for all $s \leq t$. In many of the models we consider, however, such as rational-expectations models, there are stronger conditions that we can impose, namely: $E(e_{it}\mid Z_{i}) = 0$ for all $s \leq t$. As a result, there are many more instruments available than the ones suggested by Hayashi and SAW. Any measurable function of $Z_{it}$ would also be a valid instrument. So, for example, cross-products of elements of $Z_{it}$ from
s. is one in which $y$ is now assumed to be correlated only with $a$ and not with $e$. For this model, they again proposed IV estimation on (2) using $Q_Y X$, $Q_Y X$, and $X^* = \ell T$ as instruments. To show that this estimator is asymptotically efficient, they referred to the results in their appendix. To use those results, they wrote the model in the transformed form

$$
\begin{bmatrix}
\dot{y}_1 \\
\vdots \\
\dot{y}_{T-1} \\
\dot{y}_T
\end{bmatrix} =
\begin{bmatrix}
\dot{x}_1 w' \\
\vdots \\
\dot{x}_{T-1} w' \\
x
\end{bmatrix} \beta +
\begin{bmatrix}
0 \\
\vdots \\
0 \\
Z
\end{bmatrix} \gamma \\
\dot{e}
$$

where the superscript $*$ denotes demeaning. AM called the demeaned variables $\dot{y}_i$, through $\dot{y}_{T-1}$, "predetermined."

By "predetermined," they meant that $\dot{y}_1, \ldots, \dot{y}_{T-1}$ are uncorrelated with $\dot{e}_{T+1}, \ldots, \dot{e}_T$ but correlated with $\dot{e}_{T-1}$, $\ldots, \dot{e}_T$. Note that this is the reverse (in time) of the standard definition of predetermined. Clearly this definition applies to the model (3) because $\dot{y}_1, \ldots, \dot{y}_{T-1}$ are uncorrelated with $\dot{e}_{T-1}$ but correlated with $\alpha + \dot{e}$, the error in the equation indexed by $T$. Note, however, that for $\dot{y}_1, \ldots, \dot{y}_{T-1}$ to be uncorrelated with $\dot{e}_{T-1}$, $\ldots, \dot{e}_T$ it is necessary that, $y$ be strictly exogenous with respect to $e$ in our terminology. Thus this estimation method would not be applicable to the case we consider.

In the fourth and final model considered by AM, in their section 6, they extended the previous model to allow for serial correlation in the $e$. They showed that this is not a problem, because if the covariance matrix of $(\dot{e}_1, \ldots, \dot{e}_{T-1}, e)$ in (3) is $\Omega$, then premultiplication of (3) by a lower triangular matrix $Q$, such that $Q \Omega Q' = I_{NT}$, preserves the "predetermined properties" of the $\dot{y}_i$. But again, the predetermined properties require our strict exogeneity condition with respect to $e$. Thus we feel that strict exogeneity assumptions are being made to construct all four estimators considered by AM. Note also that the additional restrictions proposed by MacCurdy (1982) on the higher moments of panel-data models will yield increased efficiency if errors are homoscedastic. If there is heteroscedasticity, however, which we believe is likely with most panel-data sets, imposing those additional restrictions will cause all parameter estimates to be inconsistent (see Newey 1990).

### 4. Aggregate Shocks

Finally, Hayashi raises the issue of whether any of the estimators we use are consistent for estimating rational-expectations panel-data models in the presence of aggregate shocks. Hayashi cites Chamberlain's (1984) concern that aggregate shocks may result in inconsistent parameter estimates in such models. We are well aware of this issue, having devoted large parts of two recent articles to addressing it (Keane and Runkle 1990, Runkle 1991). The reason we did not address it in this article is that we felt it would be an unnecessary complication in this case. In one of our previous articles (Runkle 1991), we tested for aggregate shocks using essentially the same data set that we used in this article. Since the null hypothesis of no aggregate shocks was not rejected, we felt that discussing aggregate shocks in this article would be an unnecessary distraction. In addition, that previous article found no evidence for any kind of individual effect. Therefore, we believe that the probability of an interaction effect between an aggregate shock and a random individual effect is quite unlikely. We refer the reader to our previous articles for discussions of how to obtain consistent estimates when aggregate shocks are present.

### ADDITIONAL REFERENCES


