Pooling in Dynamic Panel-Data Models: An Application to Forecasting GDP Growth Rates

André J. Hoogstrate
CentER, Tilburg University, NL-5000 LE Tilburg, The Netherlands (*a.j.hoogstrate@kub.nl*)

Franz C. Palm
Faculty of Economics and Business Administration, Maastricht University, NL-6200 MD Maastricht, The Netherlands (*f.palm@ke.unimaas.nl*)

Gerard A. Pfann
Faculty of Economics and Business Administration, Maastricht University, NL-6200 MD Maastricht, The Netherlands, and Centre for Economic Policy Research, London NW1 45A, United Kingdom (*g.pfann@ke.unimaas.nl*)

In this article, we analyze issues of pooling models for a given set of N individual units observed over T periods of time. When the parameters of the models are different but exhibit some similarity, pooling may lead to a reduction of the mean squared error of the estimates and forecasts. We investigate theoretically and through simulations the conditions that lead to improved performance of forecasts based on pooled estimates. We show that the superiority of pooled forecasts in small samples can deteriorate as the sample size grows. Empirical results for postwar international real gross domestic product growth rates of 18 Organization for Economic Cooperation and Development countries using a model put forward by Garcia-Ferrer, Highfield, Palm, and Zellner and Hong, among others illustrate these findings. When allowing for contemporaneous residual correlation across countries, pooling restrictions and criteria have to be rejected when formally tested, but generalized least squares (GLS)-based pooled forecasts are found to outperform GLS-based individual and ordinary least squares-based pooled and individual forecasts.

KEY WORDS: Contemporaneous disturbance correlation; Linear restrictions; Systems of regression equations.

Panel data are used more and more frequently in business and economic studies. Sometimes a given number of entities is observed over a longer period of time, whereas traditionally panel data are available for a large and variable number of entities observed for a fixed number of time periods (e.g., see Baltagi 1995 for a recent overview; Maddala 1991; Maddala, Trost, and Li 1994). In this article, we analyze issues of pooling models for a given set of N individual units observed over T periods of time. Large T asymptotics, with N fixed, provide the benchmark against which to evaluate the methods considered. Pooling estimates in panel-data models is appropriate if parameters are the same for the individual units observed. When the parameters are different but exhibit some similarity, pooling may also lead to a reduction of the mean squared error (MSE) of the estimates. A reduction of the MSE will be achieved when the square of the bias resulting from imposing false restrictions is outweighed by the reduction of the variance of the estimator due to restricted estimation. The existence of this trade-off has generated the literature on MSE criteria and associated tests for the superiority of restricted over unrestricted least squares estimators (e.g., see Wallace and Toro-Vizčarrondo 1969; Wallace 1972; Goodnight and Wallace 1972; McElroy 1977).

Pooling techniques have been successfully applied, for instance, to test the market efficiency hypothesis (e.g., see Bilson 1981) and to forecast multicountry output growth rates (e.g., see Mittnik 1990). In a study of real gross national product (GNP) growth rates of nine Organization for Economic Cooperation and Development (OECD) countries for the period 1951–1981, Garcia-Ferrer, Highfield, Palm, and Zellner (1987) showed that pooled estimates of an (autoregressive) AR(3) model with leading indicator (LI) variables, denoted by AR(3)LI, provided superior forecasting results. Forecasting results for an extended time period, 1974–1984, and an extended number of countries, 18 OECD countries, provided by Zellner and Hong (1989) were in favor of the earlier findings. Leading economic indicators have come to play a dominant role in forecasting business-cycle turning points on a single-country level (Stock and Watson 1989) as well as on a multicountry level (Zellner, Hong, and Min 1991). Cross-country, cross-equation restrictions have also been imposed successfully to analyze convergence of annual log real per capita output for 15 OECD countries from 1900 to 1987 (Bernard and Durlauf 1995).

The objective of this article is threefold. First, in Section 1, we investigate theoretically whether the improvement of forecasting performance using pooling techniques instead of single-country forecasts remains valid as T grows large(r) while N remains constant. The model that we investigate consists of a set of dynamic regression equations with contemporaneously correlated disturbances. It nests

Journal of Business & Economic Statistics
July 2000, Vol. 18, No. 3

the specifications put forward by Garcia-Ferrer et al. (1987) and used in many studies since (e.g., see Zellner 1994). Second, in Section 2 we present simulation results that give some insights into the importance of the gains from pooling data when the sample size is small. The simulations also provide evidence on the statistical properties of some test procedures for pooling restrictions. Third, the theoretical results are investigated empirically. The Zellner, Hong, and Min (1991) data, slightly modified to have a consistent set of real gross domestic product (GDP) growth rates of 18 OECD countries for an extended period of 1948 to 1990, are used to forecast international growth rates using individual and pooled estimates of the AR(3)LI model. Section 3 concludes.

1. THE MODEL AND THE FORECASTING PROCEDURES

Consider a linear regression model for N units/countries and T successive observations:

$$y_{it} = x'_{it}\beta_i + \varepsilon_{it}, \qquad i = 1,2,\ldots N, t = 1,2,\ldots T, \quad (1)$$

or

$$\underset{T\times 1}{y_i} = \underset{(T\times k)}{X_i}\ \underset{(k\times 1)}{\beta_i} + \underset{T\times 1}{\varepsilon_i}, \qquad i = 1,2,\ldots,N, \quad (2)$$

where y_{it} denotes the value of the endogenous variable y for country i in period t, x_{it} is a vector of explanatory variables, β_i is a vector of k regression coefficients for country i, and ε_{it} denotes a disturbance term. Alternatively, the model (2) for N countries can be written as

$$\underset{n\times 1}{y} = \underset{(n\times Nk)}{X^*}\ \underset{(Nk\times 1)}{\beta_a} + \underset{n\times 1}{\varepsilon}, \quad (3)$$

with $y = (y'_1, y'_2, \ldots, y'_N)', X^* = \text{diag}(X_1, X_2, \ldots, X_N)$, $\beta_a = (\beta'_1, \beta'_2, \ldots, \beta'_N)', \varepsilon = (\varepsilon'_1, \varepsilon'_2, \ldots, \varepsilon'_N)'$, and $n = TN$.

We allow the vector x_{it} to contain lagged values of y_{it}. The disturbances ε_{it} are assumed to be normally distributed with mean 0 and zero serial correlations, possibly contemporaneously correlated; that is $\varepsilon \sim \text{N}(0, \Omega)$ with $\Omega = \Sigma \otimes I_T$, where Σ denotes the contemporaneous covariance matrix of dimension N. The regressors x_{it} are predetermined: $E(x_{it}\varepsilon'_{js}|x_{lt-1}, \ldots, l = 1,2,\ldots,N) = 0$ for all i, j and $t \leq s$.

The system of seemingly unrelated regressions (SUR) in (3) has been extensively studied in the literature, both from a classical and a Bayesian point of view. For instance, Haytovsky (1990) generalized Zellner's (1971) Bayesian analysis of the SUR, using a linear hierarchical structure similar to that in the Lindley and Smith (1972) pooling model. Chib and Greenberg (1995) carried out a hierarchical analysis of SUR models with serially correlated errors and time-varying parameters. Nandram and Petruccelli (1997) considered pooling autoregressive time series panel data using a hierarchical framework for a model with a highly structured covariance matrix Σ, with off-diagonal elements that depend on one unknown parameter.

For $\Sigma = \sigma^2 I_N$, assuming a hierarchical structure, Garcia et al. (1987), Zellner and Hong (1989), and Min and Zellner (1993) considered shrinkage forecasts based on an estimate of β_a in (3), which is a matrix-weighted average of the equation-by-equation least squares estimates $\hat{\beta}_i$ and a pooled estimate [i.e., the least squares estimate of β restricting the β_i's in (3) to be the same].

A scalar-weighted average of these estimators results, with the scalar weight depending on the ratio of σ^2 and the prior variance of the β_i's, if the "g-prior" approach of Zellner (1986) is adopted (see Zellner and Hong 1989).

In this article, we consider the problem of choosing among forecasts based on various estimators of β_a in (3) from a sampling theory point of view using an F statistic that has been proposed to test for pooling. Note that there is a direct link between an F statistic for linear restrictions on the coefficients of a linear regression model and the posterior odds used to choose among the models associated with the hypothesis to be tested, although their interpretations will differ (see Zellner 1984).

The reasons for considering the problem of choosing among forecasts are fourfold.

First, under a diagonal loss structure, it is optimal to select a forecast rather than to combine forecasts (e.g., see Min and Zellner 1993).

Second, as described previously, the problem of combining forecasts has been extensively studied in the literature, using a hierarchical structure, whereas the problem of choosing among forecasts based on alternative estimators for an SUR has received less attention. In particular, SUR models for N as large as 18 have not been extensively used.

Third, adopting a full-fledged hierarchical Bayesian procedure in a model with unrestricted covariance matrix Σ, when N is large, requires integration in high dimensions [at least $N(N+1)/2$]. For applications with N as large as 18, this requirement may make it prohibitive to use such procedures if it is not appropriate to impose some structure on Σ.

Fourth, unpooled and pooled forecasts are the ingredients required for combining forecasts in a hierarchical structure. Our results can be interpreted as a benchmark against which combined forecasts can be judged.

We consider the following one-step-ahead forecasts of $y_{iT+1}, \hat{y}_{iT+1} = x'_{iT+1}\hat{\beta}_i$, with $\hat{\beta}_i$ being an estimate of β_i:

1. The individual forecast is based on the least squares estimator of β_i,

$$\hat{\beta}_i = (X'_i X_i)^{-1} X'_i y_i. \quad (4)$$

2. The pooled (p) forecast is based on the OLS estimator for the pooled data,

$$\hat{\beta}^p = (X'X)^{-1} X'y, \quad (5)$$

where $X = (X'_1, X'_2, \ldots, X'_N)'$.

3. The forecast (g) is based on a feasible SUR estimator of $\beta_i, \hat{\beta}^g_i$, being the ith subvector of the generalized least squares (GLS) estimator of β_a in (3),

$$\hat{\beta}^g_a = (X^{*\prime}\hat{\Omega}^{-1}X^*)^{-1}X^{*\prime}\hat{\Omega}^{-1}y, \quad (6)$$

with $\hat{\Omega}$ being a consistent estimate of Ω.

4. The forecast (pg) is based on a pooled feasible GLS estimator of β_i,

$$\hat{\beta}^{pg} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y. \quad (7)$$

The predictors based on the estimators (4)–(5) were used by Garcia-Ferrer et al. (1987) and Zellner and Hong (1989) to obtain point forecasts of GNP growth rates and by Zellner et al. (1991) to forecast turning points in GNP growth. We investigate the behavior of the four previously mentioned predictors when the time dimension of the sample becomes large. We compare the performance of individual forecasts with the pooled forecasts when T grows while N remains fixed and investigate the conditions under which pooling leads to improved forecast performance.

To compare the forecast performance, we use the mean squared forecast error (MSFE) criterion

$$\text{MSFE}(\hat{y}_{iT+1}) \doteq x'_{iT+1}\text{MSE}(\tilde{\beta}_i)x_{iT+1} + \sigma_i^2, \quad (8)$$

where $\tilde{\beta}_i$ denotes an estimator of β_i [one of the estimators (4)–(7)] and σ_i^2 is the ith diagonal element of Σ. The cross-term $x'_{iT+1}(\beta_i - \tilde{\beta}_i)\varepsilon_{iT+1}$ has been deleted from the rightside of (8) as it vanishes asymptotically. Comparing the MSFE of various forecasts basically reduces to comparing the MSE's of the estimators used to compute the forecasts.

The results obtained by McElroy (1977) can be applied to compare the GLS estimator $\hat{\beta}_a^g$ (6) for β_a with the pooled GLS estimator (7) $\hat{\beta}_a^{pg} = (\iota \otimes \hat{\beta}^{pg})$, with ι being an $N \times 1$ unit vector. When Ω is unknown and a consistent estimate is used, the results of McElroy (1977) hold for a large sample. When $\Omega = \sigma^2 I_n$, McElroy's (1977) results specialize accordingly and can be used to compare the OLS estimators $\hat{\beta}_i$ in (4) with the pooled estimator $\hat{\beta}^p$ in (5). When Ω is known, the restricted GLS estimator $\hat{\beta}_a^{pg}$ is preferred to $\hat{\beta}_a^g$ by the strong MSE criterion defined by McElroy if

$$l'[\text{MSE}(\sqrt{n}\hat{\beta}_a^g) - \text{MSE}(\sqrt{n}\hat{\beta}_a^{pg})]l \geq 0 \quad (9)$$

holds for all $Nk \times 1$ vectors $l \neq 0$. Notice that the condition (9) implies superiority of the pooled estimator for each country. Imposing restrictions erroneously produces a biased estimator. When the condition (9) holds, the size of the bias is outweighed by the reduction of the variance of the estimator.

Using the results of McElroy, Expression (9) can be shown to hold iff $\lambda_n \leq 1/2$ with $\lambda'_n = \delta'_n(RV_nR')^{-1}\delta_n/2$ and $\delta_n = \sqrt{n}(R\beta_a)$. Notice that $R\beta_a = 0$ denotes the restrictions on β_a when we pool across countries, and the $q \times Nk$ matrix R with $q = (N-1)k$ is

$$R = [\iota_{N-1} \otimes I_k, -I_{N-1} \otimes I_k], \quad (10)$$

with ι_{N-1} being an $(N-1)$ unit vector. $V_n = n(X^{*\prime}\Omega^{-1}X^*)^{-1}$ is the covariance matrix of $\sqrt{n}\hat{\beta}_a^g$.

The null hypothesis $R\beta_a = 0$ is true iff $\lambda_n = 0$. The restricted estimator $\hat{\beta}_a^{pg}$ is preferred to $\hat{\beta}_a^g$ by the first weak MSE criterion, which requires that the trace of the difference of the MSE matrices in (9) be nonnegative. This holds whenever

$$\lambda_n \leq \theta_n, \quad (11)$$

with $\theta_n = 1/2\mu_n\text{tr}[V_nR'(RV_nR')^{-1}RV_n]$ and μ_n being the smallest characteristic root of V_n^{-1}. Finally, the pooled GLS estimator is better than the GLS estimator by the second weak MSE criterion defined by McElroy (1977) as

$$E[n(\hat{\beta}_a^g - \beta_a)'V_n^{-1}(\hat{\beta}_a^g - \beta_a) - n(\hat{\beta}_a^{pg} - \beta_a)'V_n^{-1}(\hat{\beta}_a^{pg} - \beta_a)] \geq 0, \quad (12)$$

which holds iff $\lambda_n \leq q/2$—that is, iff $2\lambda_n$ is smaller than the number of restrictions on regression coefficients in the system.

As shown by McElroy (1977), when the X_i's are strictly exogenous (standard regression model) the test statistic

$$F^*(\Omega) = \frac{n(R\hat{\beta}_a^g)'(RV_nR')^{-1}(R\hat{\beta}_a^g)/q}{\text{SSE}(\hat{\beta}_a^g)/(n-q)} \sim F(q, n-q, \lambda_T), \quad (13)$$

where $\text{SSE}(\hat{\beta}_a^g) = (y - X^*\hat{\beta}_a^g)'\Omega^{-1}(y - X^*\hat{\beta}_a^g)$ has a noncentral F distribution. It can be used to test hypotheses about λ. Rejecting the null hypothesis when the statistic is large provides a uniformly most powerful test for $\lambda_n \leq \lambda^*$ against $\lambda_n > \lambda^*$, where $\lambda^* = 1/2, \theta_n$, or $q/2$ depending on the chosen MSE criterion.

As $n \to \infty$ (e.g., for $T \to \infty$ and fixed N), qF^* in (13) converges to a (non)central $\chi^2(q, \lambda)$ distribution with $\lambda = \lim_{n\to\infty} \lambda_n$, when the sequence of alternative hypotheses is chosen in such a way that λ_n converges to a finite limit λ. When Ω is replaced by a consistent estimator and/or in the presence of predetermined variables among the regressors, the same limiting distribution for qF^* results. Moreover, as $n \to \infty$ the covariance matrices for both $\sqrt{n}\hat{\beta}_a^g$ and $\sqrt{n}\hat{\beta}_a^{pg}$ converge to constant matrices. If the restrictions $R\beta_a = 0$ do not hold, δ_n, and hence the bias of $\sqrt{n}\hat{\beta}_a^{pg}$, increases without bound while the unrestricted estimator $\sqrt{n}\hat{\beta}_a^g$ remains unbiased. Therefore, for each of the three MSE criteria, there exists a sufficiently large n to make $\hat{\beta}_a^{pg}$ worse than $\hat{\beta}_a^g$ in terms of MSE. In the case in which the restrictions are false, the noncentrality parameter λ_n increases without bound. As a result, the power of a test of $H_0 : \lambda_n \leq \lambda^*$ tends to 1 and the test is consistent.

In Section 2, we shall report the findings of an empirical analysis of the MSFE of forecasts based on unrestricted and pooled estimators. In particular, we shall investigate under which conditions and for which sample size it pays to use a pooled estimator rather than an unrestricted estimator.

2. EMPIRICAL ANALYSES

In this section we investigate an AR(3) model and an AR(3)LI model used by Garcia-Ferrer et al. (1987) and Zellner and Hong (1989). The dataset consists of the post-World War II real GDP growth rates of 18 OECD countries. The parameters of the individual countries are estimated using samples for which the starting date varies between 1949 and 1957 and ends in 1990 (see Appendix A). In Subsection 2.1, we present the models, estimate, and test them. Subsection

2.2 is devoted to a simulation study of the finite-sample properties of McElroy's criteria for pooling. The empirical models for a subset of the 18 OECD countries are used in the simulations. Finally, in Subsection 2.3 we check the pooling restrictions for the 18 countries using McElroy's criteria.

2.1 The Models

The following AR(3)LI model was employed to generate one-year-ahead forecasts of the growth rate of real GDP, for the period 1981–1990, for 18 OECD countries:

$$y_{it} = \beta_{0i} + \beta_{1i}y_{it-1} + \beta_{2i}y_{it-2} + \beta_{3i}y_{it-3} + \beta_{4i}SR_{it-1} + \beta_{5i}SR_{it-2} + \beta_{6i}GM_{it-1} + \beta_{7i}WR_{it-1} + \varepsilon_{it}, \quad (14)$$

where y_{it} denotes the first difference of the logarithm of real output, SR_{it} denotes the first difference of the log of a stock price index divided by a general price index, GM_{it} denotes the first difference of the log of the nominal money supply divided by a general price index, and WR_{it} denotes world return, which equals the median of countries' real stock return in period t.

The AR(3) model arises as a special case of Model (14) when $\beta_{4i} = \beta_{5i} = \beta_{6i} = \beta_{7i} = 0$.

To illustrate the gains from pooling when T is small, we performed the actual one-step-ahead forecasts for the period 1983–1990 using an AR(3) model. The results are shown in Table 1, in which columns 1 and 2 are based on models estimated by all data, columns 3 and 4 are from models estimated by excluding the first 5 observations for each country, columns 5 and 6 are from models estimated by excluding the first 10 observations, and finally columns 7 and 8 are from models estimated by excluding the first 15 observations. As expected, as T gets smaller, the forecasting performance of the pooled forecast dominates the individual predictor. This is apparent in two ways in Table 1. First, the number of countries for which the individual forecast is better declines as T becomes smaller. Second, the difference in median root mean squared errors (RMSE's) increases as T decreases.

Next we compare the forecasting performance for different periods of time. Table 2 reports RMSE's for forecasts (RMSFE's) based on country-specific OLS parameter estimates and pooled parameter estimates for the forecasting periods 1974–1987, 1974–1990, and 1983–1990. For the three forecast periods and for both models, the pooled forecasts dominate the individual forecasts in most instances. Moreover, the median of the RMSFE's is lowest for the pooled forecasts in the six cases. This finding is in line with those of Garcia-Ferrer et al. (1987) for a forecast period 1974–1981 and Zellner and Hong (1989) for the periods 1974–1981 and 1974–1984. Notice also that for the forecast period 1983–1990, surprisingly the AR(3) model performs better than the AR(3)LI model in terms of RMSFE. For the forecast period 1974–1987, the AR(3)LI clearly performs better than the AR(3) model in most instances. There are small differences with results reported by Min and Zellner (1993) for the same forecast period. These are because they used GNP and GDP data and an estimation period 1954–1973 with data for 1951–1953 serving as initial values whereas in the present study strictly GDP data are used for all countries and the estimation period is 1961–1973 with data for 1958–1960 serving as initial values.

The finding that, as T grows, the difference between the MSFE's of forecasts based on OLS and pooled forecasts becomes larger suggests that the restrictions of identical parameters across countries are not literally true. Before testing these restrictions, we shall examine the presence of contemporaneous correlation between the disturbances for the 18 countries in the AR(3) and AR(3)LI models, respectively.

Table 1. The Root Mean Squared Forecast Error for the Individual Country Forecast and for the Pooled Forecast Using the AR(3) Model When up to 15 Observations Are Excluded From the Estimation Period

	Number of observations excluded							
	0		5		10		15	
	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled
Australia	3.66	3.36	3.66	3.41	3.60	3.35	3.26	3.30
Austria	1.53	.92	1.30	.89	.92	.74	1.03	.75
Belgium	1.38	1.79	1.54	1.71	1.59	1.69	1.97	1.74
Canada	3.90	3.47	4.26	3.44	4.39	3.43	4.59	3.35
Denmark	1.86	1.81	2.04	1.77	2.45	1.81	2.66	1.72
Finland	1.49	1.37	1.29	1.38	1.23	1.52	1.26	1.69
France	1.44	1.59	1.67	1.52	1.69	1.47	1.70	1.46
Germany	4.27	4.50	4.40	4.30	4.47	4.39	4.37	4.34
Ireland	2.23	1.68	2.48	1.67	3.78	1.64	3.95	1.74
Italy	2.40	1.55	2.39	1.47	2.42	1.52	2.43	1.52
Japan	1.09	.99	1.05	.98	1.14	1.06	1.32	1.23
Netherlands	1.67	1.46	1.60	1.42	1.42	1.26	1.61	1.27
Norway	4.95	3.98	5.00	3.91	5.52	3.86	5.66	3.71
Spain	1.03	1.19	2.03	1.22	2.00	1.34	1.79	1.54
Sweden	1.41	1.19	1.47	1.16	1.72	1.20	1.69	1.22
Switzerland	1.54	1.48	1.56	1.48	1.65	1.40	2.00	1.43
U.K.	2.38	1.90	2.51	1.88	2.77	1.93	2.90	1.96
U.S.	3.15	2.74	3.59	2.64	4.20	2.76	4.13	2.75
Median	1.76	1.63	2.04	1.59	2.21	1.58	2.22	1.71

Table 2. Eighteen Countries RMSFE's of One-Year-Ahead Forecasts, 1974–1987, 1974–1990, and 1983–1990

Period	AR(3)						AR(3)LI					
	1974–1987		1974–1990		1983–1990		1974–1987		1974–1990		1983–1990	
γ	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled
Australia	4.73	3.21	4.34	2.92	3.66	3.36	4.94	2.63	4.77	2.42	3.55	1.91
Austria	3.30	2.60	3.02	2.36	1.53	.92	3.42	2.51	3.15	2.29	3.02	1.88
Belgium	3.19	3.24	2.95	2.98	1.38	1.79	1.87	2.41	2.31	2.38	2.93	2.21
Canada	3.78	3.37	3.48	3.12	3.90	3.47	3.99	2.91	3.65	2.67	4.08	3.48
Denmark	3.50	3.40	3.36	3.30	1.86	1.81	4.00	3.92	4.02	3.94	3.73	3.06
Finland	3.63	3.34	3.34	3.12	1.49	1.37	3.74	2.92	3.61	2.73	2.66	1.53
France	3.15	2.42	2.91	2.30	1.44	1.59	3.13	2.72	2.90	2.68	1.51	2.30
Germany	5.05	4.67	4.62	4.36	4.27	4.50	4.19	3.98	3.90	3.75	3.90	3.69
Ireland	4.81	4.36	4.45	4.02	2.23	1.68	4.48	4.12	4.25	3.88	2.60	1.78
Italy	3.85	3.23	3.56	3.00	2.40	1.55	3.66	2.42	3.54	2.46	3.47	1.85
Japan	4.36	2.98	3.98	2.73	1.09	.98	4.17	2.77	3.95	2.58	2.45	1.23
Netherlands	3.39	2.58	3.13	2.41	1.67	1.45	3.23	2.47	4.00	2.46	2.90	2.46
Norway	2.85	3.02	3.70	3.15	4.94	3.97	3.37	3.12	4.20	3.88	5.36	5.43
Spain	3.00	2.57	2.74	2.40	1.03	1.19	3.03	2.28	2.95	2.09	2.32	1.24
Sweden	2.91	2.76	2.66	2.52	1.41	1.18	2.90	2.48	2.65	2.34	2.58	1.75
Switzerland	3.73	3.77	3.46	3.43	1.54	1.48	3.89	3.46	3.71	3.19	2.51	2.13
U.K.	3.49	3.27	3.36	3.10	2.38	1.90	2.83	2.64	4.48	3.35	5.64	3.76
U.S.	4.01	3.79	3.66	3.50	3.15	2.74	3.50	3.42	3.49	3.30	3.75	2.62
Median	3.57	3.24	3.41	3.05	1.76	1.63	3.58	2.74	3.63	2.68	2.79	2.17

For the estimation period 1961–1980, as expected, the residuals of the AR(3) model show more contemporaneous correlation than the residuals of the AR(3)LI model. Including leading indicators, which are approximately white noise, accounts for a major part of the contemporaneous residual correlation present in the AR(3) model.

From the estimated residual correlations, most of which are positive as expected, it also appears that the countries can be clustered in regional groups exhibiting much within-group contemporaneous residual correlation and little between-group contemporaneous residual correlation indicating that shocks to real GDP growth are partly synchronized within blocks and uncorrelated between blocks. We distinguish the following seven regional blocks—(1) Canada, United States; (2) Australia, Japan; (3) Denmark, Finland, Norway, Sweden; (4) Belgium, France, Germany, Netherlands; (5) Ireland, United Kingdom; (6) Austria, Switzerland; (7) Italy, Spain.

To formally test for contemporaneous residual correlation, we use a Lagrange multiplier (LM) statistic proposed by Breusch and Pagan (1980) for testing the null hypothesis of a diagonal Σ. Under $H_0, \lambda_{\text{LM}} = T\sum_{i=2}^{N}\sum_{j=1}^{i-1} r_{ij}^2$, with r_{ij} being the sample correlation coefficients between the residuals of the OLS estimates for countries i and j, has an asymptotic $X^2[\text{N}(N-1)/2]$ distribution. The results for the LM test are given in Table 3.

The X^2 statistics reported in Table 3 clearly indicate that the diagonality of Σ is not rejected for the AR(3)LI model for the observation period 1961–1980 when tested against an unrestricted Σ matrix. When tested against a block-diagonal matrix, diagonality is rejected. For the longer period 1961–1990, we reject the null hypothesis of a diagonal Σ matrix. This is possibly due to small sample size or structural changes that occurred in the 1980s. For this latter period, the block-diagonal structure is not rejected for the AR(3)LI model. For the AR(3) model the null hypothesis has to be rejected in all instances. This is not surprising because the common leading indicator WR_{t-1}, which is approximately white noise, accounts for interdependencies among the white-noise disturbances of the countries. Note that Chib and Greenberg (1995) reported that, when a time-varying parameter version of the AR(3)LI model is employed for the output growth rate of five countries (Australia, Canada, Germany, Japan, and the United States) in the period 1960–1987, the matrix Σ is found to be diagonal.

2.2 Properties of Pooling Restriction Tests

Before we check the appropriateness of pooling in a system of 18 equations, we investigate the small-sample properties of the F statistic given in (13) for testing the pooling restriction $H_0 : R\beta_a = 0$ against $H_1 : R\beta_a \neq 0$ and of McElroy's strong and weak pooling criteria allowing for a block-diagonal matrix Σ. The simulation results have been

Table 3. Testing for Contemporaneous Error Covariances

	Hypotheses about Σ		
H_0: H_1:	Diagonal Unrestricted	Diagonal Block-diagonal*	Block-diagonal Unrestricted
AR(3)LI 1961–1980	167.68 df = 153 $p = .197$	99.40 df = 17 $p < .001$	77.28 df = 136 $p > .5$
AR(3)LI 1961–1990	290.88 df = 153 $p < .01$	149.67 df = 17 $p < .001$	141.21 df = 136 $p = .362$
AR(3) 1961–1980	353.46 df = 153 $p < .001$	105.94 df = 17 $p < .001$	247.5 df = 136 $p < .001$
AR(3) 1961–1990	474.67 df = 153 $p < .001$	137.79 df = 17 $p < .001$	336.88 df = 136 $p < .01$

* We distinguish the following seven blocks: (1) Canada, U.S.; (2) Australia, Japan; (3) Denmark, Finland, Norway, Sweden; (4) Belgium, France, Germany, the Netherlands; (5) U.K., Ireland; (6) Austria, Switzerland; (7) Italy, Spain.

obtained using models for the following sets of countries—Belgium, Germany, France, and the Netherlands, and Canada and the United States. The general model consists of a set of six third-order autoregressions with two leading indicators:

$$\underset{(6\times6)}{\Phi(L)}\ \underset{(6\times1)}{y_t} = \underset{(6\times6)}{B}\ \underset{(6\times1)}{x_{1t}} + \underset{6\times1}{\gamma}\ x_{2t} + \underset{6\times1}{\varepsilon_t}, \quad (15)$$

where $\Phi(L)$ is a diagonal lag polynomial matrix with a third-degree polynomial on the main diagonal and where B and γ denote, respectively, a matrix and a vector of coefficients. The leading indicators x_{1t} and x_{2t} and the disturbance vector ε_t satisfy the following properties:

$$x_{1t} \sim \text{IIN}(0, \Sigma_1), \quad x_{2t} \sim \text{IIN}(0, \sigma_2^2), \quad \varepsilon_t \sim \text{IIN}(0, \Sigma_3), \quad (16)$$

with Σ_1 being diagonal and Σ_3 being a block-diagonal covariance matrix. The variables x_{1t}, x_{2t}, and ε_t are mutually independent. The vector x_{1t} can be interpreted as a country-specific leading indicator. The variable x_{2t} can be interpreted as a common leading indicator (e.g., WR_{t-1}). The block-diagonal structure of Σ_3 reflects the finding that the disturbances within European and North American subgroups are correlated and that the between-subgroup correlations are 0. The model (15)–(16) implies a third-order vector autoregressive (VAR) model for y_t with a diagonal VAR matrix and a full-disturbance covariance matrix $\Sigma = B\Sigma_1 B' + \sigma_2^2\gamma\gamma' + \Sigma_3$. The error-component structure of the disturbance term of the AR(3) model will be ignored in the sequel.

Both the AR(3)LI model (15) and the implied VAR(3) model have been simulated. The models have been simulated under parameter heterogeneity across countries and under regression parameter homogeneity. The parameter values used in the simulations for the AR(3)LI model are set equal to the OLS estimates of Model (15), taking x_{1t} to be the vector of observed GM_{it-1} and $x_{2t} = WR_{t-1}$. Under parameter homogeneity, OLS estimates of the equation in (15) for the United States are used for all six countries. The parameter values of the AR(3) model are derived from those of the AR(3)LI model under, respectively, parameter heterogeneity and homogeneity. On the basis of the data for the six countries, parameter homogeneity is not rejected for model (14) when testing regression parameter equality across countries using an F test. Obviously, McElroy's criteria do not reject pooling either. The details for these tests are given in Table 4.

The empirical distribution of the F statistic in (13) and rejection frequencies were obtained by simulation. The number of runs is 1,000. Results on the empirical distribution of the F statistic in (13) are not reported here. In most instances, the empirical distributions resemble an F distribution. Rejection frequencies when the test statistic is compared with the critical value of, respectively, a central F distribution with q and $n - q$ df and that of a noncentral $F(q, n-q, \lambda_T)$ with $\lambda_T = q/2$ are reported in Table 5. A nominal significance level of 5% is used.

Table 4. Testing for Pooling for a Subset of Six Countries (Canada, the United States, Belgium, France, Germany, the Netherlands)

Model	Σ	F value	p value $\lambda = 0$	p value $\lambda = q/2$
AR(3)	$\sigma^2 I$	.5763	.9244	.9944
AR(3)	Full	1.0532	.4032	.8839
AR(3)LI	$\sigma^2 I$	1.2580	.2156	.6772
AR(3)LI	Block	1.1099	.3320	.8453

Table 5. Rejection Frequencies for the F Test at a Nominal Significance Level of 5%

			Homogeneous		Heterogeneous	
T	Estimator	Σ	$\lambda = 0$	$\lambda = q/2$	$\lambda = 0$	$\lambda = q/2$
25	AR(3)LI	$\sigma^2 I$	.024	.000	.722	.251
50	AR(3)LI	$\sigma^2 I$	.010	.000	.996	.871
100	AR(3)LI	$\sigma^2 I$	.023	.000	1.000	1.000
25	AR(3)	$\sigma^2 I$	.017	.000	.270	.038
50	AR(3)	$\sigma^2 I$	.019	.001	.672	.207
100	AR(3)	$\sigma^2 I$	.032	.000	.984	.810
25	AR(3)LI	Block	.019	.000	.369	.095
50	AR(3)LI	Block	.007	.000	.821	.354
100	AR(3)LI	Block	.005	.000	1.000	.950
25	AR(3)	Full	.254	.026	.623	.266
50	AR(3)	Full	.160	.011	.824	.388
100	AR(3)	Full	.112	.002	.983	.829

Under parameter homogeneity, the rejection frequencies are very small when McElroy's second weak criterion is tested. With the exception of the AR(3) model with unrestricted disturbance covariance matrix, the rejection frequencies when a central F distribution is used are also substantially smaller than 5% for the model under parameter homogeneity. An F test appears to be too conservative whether the correct disturbance covariance is assumed or not.

Under parameter heterogeneity, as expected, the power is found to increase as the sample size increases. For values of T equal to or larger than 50, the rejection frequency is found to be fairly large (larger than 80%). For $T = 25$, McElroy's second weak criterion rejects rather infrequently the incorrect parameter restrictions. As T increases, the gain resulting from trading off some bias against a decrease in the variance of the estimates decreases as expected on the basis of asymptotic theory.

Next, information on distributions of the mean and median RMSFE using, respectively, unrestricted and pooled parameter estimates is given in Table 6. As expected, as T increases, the distributions become more concentrated. Left-skewness of a distribution means that the median (across countries) RMSFE of the pooled forecasts is larger than that of the forecasts based on unpooled estimates. In column 4 of Table 6, we report the percentage of the number of times forecasts using unpooled estimates outperformed those based on pooled estimates. For $T = 25$, the simulations indicate that pooling is appropriate even under parameter heterogeneity. For $T = 50$, under parameter heterogeneity, it seems to be advisable to use pooled forecasts, based on a VAR model, which leaves a major part of the contemporaneous correlation in the disturbances and there-

Table 6. Mean and Median (Median) RMSFE for the Simulations

			Homogeneous					Heterogeneous				
			Mean		Median			Mean		Median		
T	Estimator	Σ	Unpooled	Pooled	Unpooled(1)	Pooled(2)	$\Delta < 0$	Unpooled	Pooled	Unpooled(1)	Pooled(2)	$\Delta < 0$
25	AR(3)LI	$\sigma^2 I$	.016	.015	.016	.014	.149	.016	.016	.016	.016	.448
50	AR(3)LI	$\sigma^2 I$	.011	.010	.010	.010	.272	.011	.011	.011	.011	.667
100	AR(3)LI	$\sigma^2 I$	.007	.007	.007	.007	.309	.007	.008	.007	.008	.782
25	AR(3)	$\sigma^2 I$	.022	.021	.022	.020	.212	.017	.017	.017	.017	.336
50	AR(3)	$\sigma^2 I$	.015	.015	.015	.014	.286	.012	.012	.012	.018	.467
100	AR(3)	$\sigma^2 I$	.010	.010	.010	.010	.348	.008	.008	.008	.008	.547
25	AR(3)LI	Block	.015	.015	.015	.015	.309	.016	.016	.016	.016	.611
50	AR(3)LI	Block	.010	.010	.010	.010	.411	.011	.011	.010	.011	.781
100	AR(3)LI	Block	.007	.007	.007	.007	.420	.007	.008	.007	.008	.832
25	AR(3)	Full	.022	.021	.022	.021	.226	.018	.017	.017	.017	.337
50	AR(3)	Full	.015	.015	.015	.014	.299	.012	.012	.012	.012	.485
100	AR(3)	Full	.010	.010	.010	.010	.354	.008	.008	.008	.008	.560

NOTE: $\Delta < 0$ represents the proportion of the number of times that the forecasts based on unpooled (1) estimates outperforms those based on pooled (2) estimates.

fore yields forecasts that are genuinely less accurate than those for the AR(3)LI model. These findings are in line with the theoretical results presented in Section 2 and the conclusions drawn previously for the F tests.

2.3 Analyses of International Data

In this section the results of an empirical analysis of the pooling restrictions for Model (14) and the associated VAR(3) model using international data for 18 countries are presented.

Tests of the pooling restrictions $H_0 : R\beta_a = 0$ against the alternative $H_1 : R\beta_a \neq 0$ are reported in Table 7 for the AR(3)LI model and the AR(3) model, respectively, for estimation periods varying from 1961–1980 to 1961–1990.

Table 7. Testing for Pooling in the AR(3)LI and AR(3) Model

	AR(3)LI				AR(3)			
Year	F value	θ	p value $\lambda = 0$	p value $\lambda = 68$	F value	θ	p value $\lambda = 0$	p value $\lambda = 34$
			$\Sigma : \sigma^2 I$				$\Sigma : \sigma^2 I$	
1981	1.11	4.14	.24	.98	.92	5.97	.66	.99
1982	1.06	3.81	.35	.99	.78	5.52	.89	1.00
1983	1.10	7.89	.26	.98	.75	7.59	.92	1.00
1984	.97	6.97	.57	1.00	.78	7.74	.90	1.00
1985	.97	7.04	.57	1.00	.78	7.51	.89	1.00
1986	.97	6.79	.58	1.00	.79	7.38	.88	1.00
1987	.94	6.53	.66	1.00	.82	7.23	.84	1.00
1988	1.01	7.03	.46	1.00	.81	7.55	.86	1.00
1989	1.03	8.35	.42	1.00	.84	7.48	.81	1.00
1990	1.04	8.11	.38	1.00	.86	7.42	.78	1.00
			Σ :diagonal				Σ :diagonal	
1981	1.39	10.35	.01	.70	1.21	10.27	.14	.87
1982	1.20	9.87	.10	.93	.98	10.68	.52	.99
1983	1.28	9.64	.05	.86	.88	10.04	.73	1.00
1984	1.20	12.46	.10	.94	.94	9.67	.62	.99
1985	1.15	11.85	.16	.97	.94	9.33	.61	.99
1986	1.13	10.14	.18	.97	.87	9.60	.76	1.00
1987	1.06	9.46	.34	.99	.87	9.10	.75	1.00
1988	1.10	8.59	.24	.99	.82	9.73	.84	1.00
1989	1.11	10.04	.23	.98	.86	9.49	.78	1.00
1990	1.15	10.12	.15	.97	.86	9.49	.77	1.00
			Σ :block				Σ :full	
1981	1.63	4.89	.00	.29	4.55	4.38	.00	.00
1982	1.34	4.75	.02	.78	3.80	5.10	.00	.00
1983	1.63	5.32	.00	.28	2.40	4.98	.00	.00
1984	1.52	2.24	.00	.45	1.75	4.80	.00	.18
1985	1.35	5.22	.02	.77	1.61	4.67	.00	.32
1986	1.41	5.81	.00	.66	2.49	5.37	.00	.00
1987	1.16	5.58	.14	.96	2.16	5.59	.00	.01
1988	1.04	6.05	.38	1.00	2.01	5.72	.00	.04
1989	.95	7.44	.64	1.00	2.12	5.72	.00	.02
1990	.95	7.34	.63	1.00	2.17	5.53	.00	.01

Table 8. Root Mean Squared Forecast Errors for the AR(3)LI and AR(3) Models

	AR(3)LI						AR(3)					
	Block		Diagonal		$\sigma^2 I$		Full		Diagonal		$\sigma^2 I$	
Country	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled
Canada	3.58	2.58	3.57	2.26	3.57	2.26	2.64	1.94	2.81	2.16	2.81	2.17
U.S.	2.45	2.75	2.64	2.34	2.64	2.39	2.04	1.80	2.13	2.23	2.13	2.27
Australia	4.23	3.09	3.97	2.73	3.97	2.67	3.32	3.69	3.69	3.53	3.69	3.50
Japan	2.91	1.52	3.21	.90	3.21	.87	1.97	.79	.78	.91	.78	.88
Denmark	3.65	3.24	3.46	2.91	3.46	2.95	1.89	1.69	2.13	1.65	2.13	1.72
Finland	5.17	3.07	4.27	2.50	4.27	2.50	2.28	2.31	2.19	2.26	2.19	2.26
Norway	4.36	3.46	4.01	3.51	4.01	3.54	7.14	2.19	4.36	2.31	4.36	2.34
Sweden	2.17	2.65	2.08	2.28	2.08	2.27	2.83	2.02	1.87	1.88	1.87	1.91
Belgium	2.93	2.67	3.19	2.03	3.19	2.05	2.43	2.01	1.39	1.53	1.39	1.51
France	3.29	2.77	1.75	2.22	1.75	2.28	3.13	1.65	1.21	1.27	1.21	1.27
Germany	3.43	2.68	3.21	2.61	3.21	2.69	3.62	2.96	2.64	2.71	2.64	2.70
Netherlands	2.23	2.27	2.03	1.81	2.03	1.88	4.24	1.32	1.23	.88	1.23	.87
Ireland	4.02	2.94	4.10	2.36	4.10	2.38	5.01	1.91	2.94	1.69	2.94	1.66
U.K.	6.40	4.10	5.67	3.93	5.67	3.89	2.83	2.56	2.45	2.55	2.45	2.57
Austria	2.17	2.17	2.52	1.76	2.52	1.74	1.89	1.10	.95	.80	.95	.79
Switzerland	2.14	2.50	2.23	1.95	2.23	1.95	1.69	1.41	1.47	1.09	1.47	1.10
Italy	3.98	2.24	3.49	1.95	3.49	1.99	1.29	1.36	1.46	1.11	1.46	1.11
Spain	3.63	1.70	3.43	1.26	3.43	1.23	3.00	1.18	.92	1.04	.92	1.02
Median	3.51	2.67	3.32	2.27	3.32	2.27	2.73	1.85	2.00	1.67	2.00	1.69

We report the values of the test statistic F^* given in (13) and the p values for the asymptotically justified tests of exact linear restrictions $R\beta_a = 0$ and of the MSE criterion for these linear restrictions. The matrix Σ is assumed to be, respectively, $\Sigma = \sigma^2 I_N$, diagonal, and block-diagonal (as explained previously).

The values of the F statistic are given in column 2 of Table 7. Column 3 contains the value of the noncentrality parameter θ_n given in (13). In the columns 4–6, the p values are reported for the tests of $H_0 : \lambda = 0$ versus $H_1 : \lambda \neq 0$, and $H_0 : \lambda \leq \lambda^*$ against $H_1 : \lambda > \lambda^*$ for λ^* being equal to $q/2$. The two criteria correspond to the tests of McElroy's strong criterion and second weak criterion, respectively. For the AR(3)LI and the AR(3) models, under the assumption that $\Sigma = \sigma^2 I_N$ or that Σ is diagonal, an F test usually does not lead to rejection of the pooling restrictions $R\beta_a = 0$. Consequently, the less stringent restrictions of the second weak MSE test for pooling given by McElroy (1977) are not rejected either in these two cases. A similar conclusion is reached if the first weak criterion is used to test for pooling. The p values are never lower than .64 when the estimation period is varied from 1961–1980 to 1961–1990. McElroy's second weak pooling criterion is not rejected in general, except for the AR(3) model using an unrestricted covariance matrix Σ. Notice that a similar conclusion holds for the first weak criterion. Moreover, the pooled AR(3) model has to be rejected when compared with the pooled AR(3)LI model.

The true size of a test based on an "F statistic" using a spherical Σ will be different from the assumed size when the true Σ is nonspherical. The evidence for the F test in Table 7 for $\Sigma = \sigma^2 I_N$ or Σ being diagonal supports the null hypothesis because neglecting residual correlation in estimation generally leads to rejection frequencies for the null hypothesis that are much lower than the nominal size of the test [e.g., see Palm and Sneek (1984) for results on the F test in a regression model when neglecting serial correlation in the disturbances].

When Σ is estimated as a block-diagonal matrix in the AR(3)LI model, $H_0 : R\beta_a = 0$ usually has to be rejected at conventional significance levels. The less restrictive second weak MSE tests do not lead to rejecting H_0 in this case. For the AR(3) model, with full disturbance covariance matrix, the p values for the pooling restrictions are very small. In this case, the second weak MSE criterion also leads to rejecting pooling (see Table 7). Note that we are using asymptotically justified procedures in relatively small samples. The procedures are asymptotically justified because Σ has to be estimated and because of the presence of lagged dependent variables among regressors. There is an earlier literature on the incorrect sizes of asymptotic tests indicating that these tests reject the null hypothesis too often in finite samples. The simulations in Subsection 3.2, however, indicate that asymptotic theory provides rather good guidance in small samples. F statistics neglecting the presence of residual correlation are expected to reject the null hypothesis less often than they should according to the nominal size. Therefore, we conclude that, on the whole, the evidence from Table 7 supports the pooling restrictions.

This conclusion is supported by the results given in Table 8. For the forecast period 1983–1990, pooling leads to a substantial reduction in the RMSFE for the AR(3) and the AR(3)LI model when a pooled GLS estimator is used with, respectively, estimated full and block-diagonal covariance matrices. For the forecast period 1983–1990, GLS-based unpooled forecasts perform slightly less than OLS-based unpooled forecasts.

3. CONCLUSIONS

In this article, we studied the problem of whether forecasts of a set of panel data generated by models with similar but not necessarily identical parameter structures can be improved by using pooled parameter estimates. Results obtained by McElroy (1977) for a regression model with nonspherical disturbances can be generalized in a straightforward way to apply to systems of regression models used to study panel data with large T and fixed N. The gain in forecast and estimator performance measured by the reduction in MSFE or MSE results from a trade-off between the bias implied by the use of (slightly) false pooling restrictions and the reduction in the covariance matrix of the estimators due to imposing these restrictions. Moreover, as the sample size increases, the covariance matrices of restricted and unrestricted estimates converge to constant matrices but the bias of the restricted estimator (multiplied by $\sqrt{n}$) increases without bound. Therefore, beyond some given sample size, the forecasts based on unrestricted estimates will outperform the pooled forecasts.

Our simulation results show that for small and moderate values of T, reductions in MSFE can be achieved through pooling, even under parameter heterogeneity. The asymptotic properties of the pooling criteria put forward by McElroy (1977) provide a fairly accurate insight into their properties for finite T.

We applied these results to growth rates for 18 OECD countries for the periods starting in the 1950s until 1991 using models put forward by Garcia-Ferrer et al. (1987) and Zellner and Hong (1989). Our empirical findings can be summarized as follows.

First, there is contemporaneous residual correlation in the form of a block-diagonal structure corresponding to regional groups present in the models for the 18 countries.

Second, when formally tested using an estimated residual covariance matrix, the pooling restrictions and the MSE criteria for pooling put forward by McElroy (1977) are rejected only for the AR(3) model. They are not rejected when a diagonal or an identity residual covariance matrix is used. We should bear in mind that asymptotically justified F test criteria tend to reject too often in finite samples. Moreover, the pooling restrictions and MSE criteria for 18 countries were jointly tested even not allowing for individual fixed effects in the form of country-specific intercepts. Comparing pooled and unpooled models using posterior odds is probably a sensible alternative in relatively small samples to asymptotically justified test criteria.

Third, in actual forecasting, the median MSFE of OLS-based pooled forecasts is found to be smaller than that of OLS-based individual forecasts. A fairly large sample size is needed for the OLS-based pooled forecasts to be outperformed by a forecast based on unrestricted estimates. Using unpooled GLS with an estimated residual covariance matrix leads to slightly improved forecast performance. Pooled GLS-based forecasts have a much lower median MSFE than pooled OLS-based forecasts. Although we did not present results for shrinkage procedures, we like to note that shrinkage forecasts are convex combinations of individual and pooled forecasts. Therefore, results for shrinkage forecasts lie between the two polar cases of forecasts based on unrestricted estimates and those based on pooled estimates. Our findings parallel results obtained by Blattberg and George (1991). When modeling sales using a chain-brand model, they found that GLS added little to their data, whereas pooling and shrinkage estimation procedures provided superior estimates to OLS. Finally, the question of whether restricting the contemporaneous residual correlations to be the same within groups of countries and possibly across groups leads to further improvement of GLS-based pooled forecasts remains to be investigated.

ACKNOWLEDGMENTS

The authors thank Arnold Zellner for kindly providing them with the data and for most helpful discussions and comments on this article. We also acknowledge the useful comments of Ruey Tsay, an associate editor, and a referee. This research was sponsored by the Economics Research Foundation, which is part of the Netherlands Organization for Scientific Research.

APPENDIX: DATA

The dataset used is an updated set as used by Min and Zellner (1993) and consists of annual postwar data for 18 OECD countries for the period 1948 to 1990. The data are obtained from the main International Monetary Funds International Financial Statistics Data Base and contain the following four variables—(1) real stock prices, (2) an index of nominal stock prices as price index, (3) nominal money M1, and (4) GDP. Because of missing values and to make a fair comparison between the AR(3) model and the AR(3)LI model, we only included those years that could be used to estimate both models. The countries, with starting year given between parentheses, are Australia (1956), Austria (1949), Belgium (1953), Canada (1955), Denmark (1950), Finland (1950), France (1950), Germany (1950), Ireland (1948), Italy (1951), Japan (1953), The Netherlands (1950), Norway (1949), Spain (1954), Sweden (1950), Switzerland (1953), United Kingdom (1957), and the United States (1955).

[Received July 1996. Revised June 1999.]

REFERENCES

Baltagi, B. H. (1995), *Econometric Analysis of Panel Data*, New York, Wiley.

Bernard, A. B., and Durlauf, S. N. (1995), "Convergence in International Output," *Journal of Applied Econometrics*, 10, 97–108.

Bilson, J. F. O. (1981), "The 'Speculative Efficiency' Hypothesis," *Journal of Business*, 54, 435–451.

Blattberg, R. C., and George, E. I. (1991), "Shrinkage Estimation of Promotional Elasticities: Seemingly Unrelated Equations," *Journal of the American Statistical Association*, 86, 304–315.

Breusch, T. S., and Pagan, A. R. (1980), "The Lagrange Multiplier Test and its Applications to Model Specifications in Econometrics," *Review of Economic Studies*, 47, 239–253.

Chib, S., and Greenberg, E. (1995), "Hierarchical Analysis of SUR Models With Extensions to Correlated Serial Errors and Time-varying Parameter Models," *Journal of Econometrics*, 68, 339–360.

Garcia-Ferrer, A., Highfield, R. A., Palm, F., and Zellner, A. (1987), "Macroeconomic Forecasting Using Pooled International Data," *Journal of Business & Economic Statistics*, 5, 53–67.

Goodnight, J., and Wallace, T. D. (1972), "Operational Techniques and Tables for Making Weak MSE Tests for Restrictions in Regressions," *Econometrica*, 40, 699–709.

Haytovsky, Y. (1990), "Seemingly Unrelated Regression With Linear Hierarchical Structure," in *Bayesian and Likelihood Methods in Statistics and Econometrics*, eds. S. Geisser, J. S. Hodges, J. S. Press, and A. Zellner, Amsterdam: North-Holland, pp. 163–177.

Lindley, D. V., and Smith, A. F. M. (1972), "Bayesian Estimates for the Linear Model," *Journal of the Royal Statistical Society*, Ser. B, 34, 1–41.

Maddala, G. S. (1991), "To Pool or not to Pool: That Is the Question," *Journal of Quantitative Economics*, 7, 255–264.

Maddala, G. S., Trost, R. P., and Li, H. (1994), "Estimation of Short Run and Long Run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators," paper presented at the Econometric Society European Meeting, Maastricht, August 29–September 2.

McElroy, M. B. (1977), "Weaker MSE Criteria and Tests for Linear Restrictions in Regression Models With Nonspherical Disturbances," *Journal of Econometrics*, 6, 389–394.

Min, C., and Zellner, A. (1993), "Bayesian and Non-Bayesian Methods for Combining Models and Forecasts With Applications to Forecasting International Growth Rates," *Journal of Econometrics*, 56, 89–118.

Mittnik, S. (1990), "Macroeconomic Forecasting Using Pooled International Data," *Journal of Business & Economic Statistics*, 8, 205–208.

Nandram, B., and Petruccelli, J. D. (1997), "A Bayesian Analysis of Autoregressive Time Series Panel Data," *Journal of Business & Economic Statistics*, 15, 328–334.

Palm, F. C., and Sneek, J. M. (1984), "Significance Tests and Spurious Correlation in Regression Models With Autocorrelated Errors," *Statistical Papers*, 25, 87–105.

Stock, J. H., and Watson, M. W. (1989), "New Indexes of Coincident and Leading Economic Indicators," in *NBER Macroeconomics Annual 1989*, eds. O. J. Blanchard and S. Fischer, Cambridge, MA, MIT Press, pp. 351–394.

Wallace, T. D. (1972), "Weaker Criteria and Tests for Linear Restrictions in Regressions," *Econometrica*, 40, 689–709.

Wallace, T. D., and Toro-Vizcarrondo, C. E. (1969), "Tables for the Mean Square Error Test for Exact Linear Restrictions in Regression," *Journal of the American Statistical Association*, 64, 1649–1663.

Zellner, A. (1971), *An Introduction to Bayesian Econometrics*, New York, Wiley.

——— (1984), *Basic Issues in Econometrics*, Chicago: University of Chicago Press.

——— (1986), "On Assessing Prior Distributions and Bayesian Regression Analysis with g-prior Distributions," in *Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti*, eds. P. K. Goel and A. Zellner, Amsterdam: North-Holland, pp. 233–243.

——— (1994), "Time Series Analysis, Forecasting and Econometric Modelling: The Structural Econometric Modelling, Time Series Analysis (SEMTSA) Approach," *Journal of Forecasting*, 13, 215–233.

Zellner, A., and Hong, C. (1989), "Forecasting International Growth Rates Using Bayesian Shrinkage and Other Procedures," *Journal of Econometrics*, 40, 183–202.

Zellner, A., Hong, C., and Min, C. (1991), "Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time-varying Parameter, and Pooling Techniques," *Journal of Econometrics*, 49, 275–304.