*Working Paper Series*

WORKING PAPER NO. 869

# A NOTE ON AGGREGATION, DISAGGREGATION AND FORECASTING PERFORMANCE

by

Arnold Zellner and Justin Tobias



DEPARTMENT OF AGRICULTURE AND RESOURCE ECONOMICS AND POLICY
BERKELEY

CALIFORNIA AGRICULTURAL EXPERIMENT STATION

*University of California*

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 869

# A NOTE ON AGGREGATION, DISAGGREGATION AND FORECASTING PERFORMANCE

by

Arnold Zellner and Justin Tobias

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
April 1998

# A Note on Aggregation, Disaggregation and Forecasting Performance

Arnold Zellner and Justin Tobias *

April 29,1998

In this note we report the results of an experiment to determine the effects of aggregation and disaggregation in forecasting the median growth rate of eighteen countries' annual output growth rates. In one approach, following Zellner and Hong (1989), we model the aggregative annual median growth rate, $w_t$ , as an autoregression of order three with lagged leading indicator input variables, denoted by ARLI, as follows:

$$w_t = \alpha + \beta_1 w_{t-1} + \beta_2 w_{t-2} + \beta_3 w_{t-3} + \beta_4 MGM_{t-1} + \beta_5 MSR_{t-1} + \epsilon_t, \tag{1}$$

where $MGM_t$ = median annual growth rate of real money in year $t$, $MGS_t$ = median annual growth rate in real stock prices in year $t$ and $\epsilon_t$ is a zero mean, non-autocorrelated, constant variance error term. Given data on 18 industrialized countries' annual output growth rates, it is possible to compute annual median growth rates, and use them and data on the other input variables appearing in (1) to obtain forecasts of future median annual growth rates of the 18 countries and turning points for them. The results of such calculations will be reported below after describing two alternative approaches to forecasting the median growth rate using disaggregated data and models for them. As an alternative to (1), we can employ the disaggregated ARLI relationships,

$$y_{it} = \gamma + \delta_1 y_{it-1} + \delta_2 y_{it-2} + \delta_3 y_{it-3} + \delta_4 GM_{it-1} + \delta_5 SR_{it-1} + \delta_6 SR_{it-2} + \delta_7 MSR_{t-1} + u_{it} \tag{2}$$

*Research financed in part by the National Science Foundation and by income from the H.G.B. Alexander Endowment Fund, Graduate School of Business, U. of Chicago.

where the subscript $i, t$ denotes the value of a variable for the $i^{th}$ country in the $t^{th}$ year, and $y, GM, SR$, and $MSR$ denote the annual growth rates of real GDP, real money, real stock prices and the median growth rate of real stock prices, respectively. See Garcia-Ferrer et al. (1987) and Zellner and Hong (1989) for discussions and uses of (2) in forecasting. Note that here we are assuming that coefficients have the same values for all countries, assumptions relaxed in these earlier papers. Then, given forecasts of 18 countries' annual output growth rates from (2), it is clearly possible to compute the median of the 18 forecasts and use it as a forecast of $w_t$, the median growth rate and compare it to the forecast that is obtained by use of equation (1), as will be done below.

Further, equation (2) can be expanded to include the current median output growth variable to obtain the following equation:

$$\begin{aligned} y_{it} &= \gamma + \phi w_t + \delta_1 y_{it-1} + \delta_2 y_{it-2} + \delta_3 y_{it-3} + \delta_4 GM_{it-1} + \\ &\quad \delta_5 SR_{it-1} + \delta_6 SR_{it-2} + \delta_7 MSR_{t-1} + u_{it} \end{aligned} \tag{3}$$

As Zellner and Hong (1989) do, we use equation (1) to obtain a forecast of $w_t$ and use it and equation (3) to forecast the output growth rates of the 18 countries. We use the median of these 18 forecasts as our forecast of the median growth rate of the 18 countries.

In the first experiment, we use annual data, 1948-73 to fit our three models and then employ them to forecast the median of the annual output growth rates year by year for the period, 1974-84, updating our parameter estimates as we move through the forecast period. The results are shown in Table 1, Part A. It is seen that use of the disaggregated equations in (3), which include the aggregate variable $w_t$ , perform the best with a RMSE and RMAE of prediction of 1.22 and 1.04, respectively. Second best is the performance of the aggregate relation in (1), with RMSE = 1.54 and RMAE = 1.20. Last in performance are the disaggregated relations in (2) that do not include the variable $w_t$ with RMSE = 1.78 and RMAE = 1.23. The empirical results for equation (1) using revised data are very similar to those reported in Zellner and Hong (1989) using unrevised data and which are better than those using just an AR(3) for $w_t$ without leading indicator variables.

As a second experiment, we employed annual data, 1948-79 to fit our models and forecasted the median growth rate of the 18 countries year by year for the period 1980-95 with results shown in Table 1, Part B. Again it is the case that use of the disaggregated equations in (3), which include the aggregate variable $w_t$, performed best with RMSE = 1.40 and RMAE= 1.10. The aggregate model in (1) performed next best with RMSE = 1.63 and RMAE = 1.21 and the model (2) slightly worse with RMSE = 1.70 and RMAE = 1.19.

Using the same data, we also performed calculations to determine which of the three models performed best in forecasting turning points in the median growth rate of the 18 countries over the period, 1981-1995. As in previous work, we define a downturn (DT) in period $T+1$ as occurring if the following sequence median output growth rates occurs:

$$w_{T-2}, w_{T-1} < w_T > w_{T+1}.$$

Also, by definition, no downturn (NDT) occurs when the following sequence is observed:

$$w_{T-2}, w_{T-1} < w_T \leq w_{T+1}.$$

Similarly, an upturn (UT) is said to occur in period $T+1$ if the following sequence of observations occurs:

$$w_{T-2}, w_{T-1} > w_T < w_{T+1}.$$

No upturn (NUT) occurs in period T+1 if the following sequence is observed:

$$w_{T-2}, w_{T-1} > w_T \geq w_{T+1}.$$

Given that we have a predictive density for $w_{T+1}$, we can easily compute the probability of DT and or NDT and use these probabilities along with a $2 \times 2$ loss structure to determine the forecast that minimizes expected loss. If the $2 \times 2$ loss structure is symmetric, a DT is the optimal forecast given that the probability of a DT is greater than 1/2. If the probability of a DT is less than 1/2, then the optimal forecast is NDT. Similar considerations relate to forecasting UTs and NUTs. See LeSage (1996), Zellner, Hong and Min (1991) and Zellner, Tobias and Ryu (1997) for applications of this turning point forecasting methodology.

In Table 2 are shown the number of DT, NDT, UT and NUT events that actually occurred in our sample and the number of correct forecasts using the procedure described above with each of our three models, shown in equations (1), (2) and (3). It is seen that use of the disaggregated AR(3)LI relations including the aggregate variable $w_t$ in equation (3) produced the best results, namely 4 of 4 correct DT forecasts, 2 of 2 correct NDT forecasts, 2 of 2 correct UT forecasts and 1 of 4 correct NUT forecasts. As regards the poor NUT forecasts, in Zellner,Tobias and Ryu (1998), it was found that use of "trend" add factors that represent inertia effects produced improved NUT forecasts for individual countries' annual output growth rates.

In summary, our forecasting experiment provides some evidence that improved forecasting results can be obtained by disaggregation given that an aggregate variable, $w_t$, appears in the disaggregated relations, as shown in (3). With disaggregation, there are more observations to estimate parameters and given that the dissagregated relations are reasonably specified, it is possible to obtain improved forecasts of an aggregate variable, here $w_t$, the median growth rate, a result that is in accord with some views expressed in the literature; see, *e.g.* Espassa (1994) and Palm and Zellner (1992).

## Table 1

**RMSE's and RMAE's of One-Year Ahead Aggregate and Disaggregate Forecasts of the Median of 18 countries Annual Real GDP Growth Rates** [1]

A: 1974-1984

| | Disaggregated AR(3)LI with $\hat{w}_t$ eqn. (3) | Aggregate AR(3)LI for $w_t$ eqn. (1) | Disaggregated AR(3)LI eqn. (2) |
|---|---|---|---|
| RMSE | 1.22 | 1.54 | 1.78 |
| RMAE | 1.04 | 1.20 | 1.23 |

B: 1981-1995

| | Disaggregated AR(3)LI with $\hat{w}_t$ eqn. (3) | Aggregate AR(3)LI for $w_t$ eqn. (1) | Disaggregated AR(3)LI eqn. (2) |
|---|---|---|---|
| RMSE | 1.40 | 1.63 | 1.70 |
| RMAE | 1.10 | 1.21 | 1.19 |

## Table 2

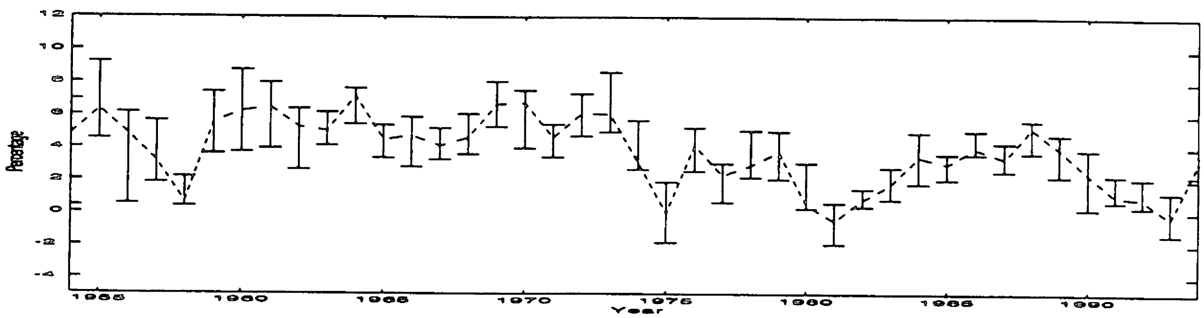
**Results of Forecasting Turning Points in the Median Output Growth Rate of 18 Countries: 1981-1995. Number of Correct Forecasts in Alternative Models**

| | DT | NDT | UT | NUT |
|---|---|---|---|---|
| | Observed Outcomes | | | |
| | 4 | 2 | 2 | 4 |
| Model | Number Correctly Forecasted | | | |
| Aggregate AR(3)LI for $w_t$. (Equation (1) ) | 4<br>— | 2<br>— | 2<br>— | 0<br>— |
| Disaggregated AR(3)LI (Equation (2) ) | 2<br>— | 2<br>— | 2<br>— | 0<br>— |
| Disaggregated AR(3)LI with $\hat{w}_t$. (Equation (3) ) | 4<br>— | 2<br>— | 2<br>— | 1<br>— |

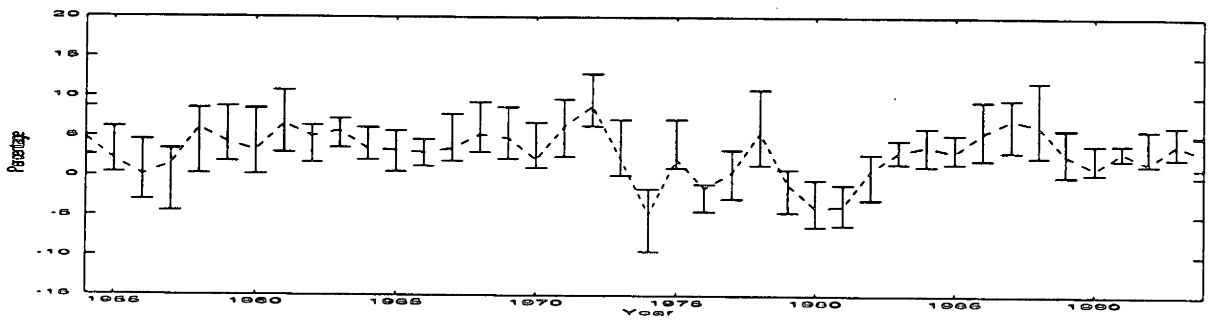
[1]The data are taken from the IMF computerized database at the University of Chicago. We use data from the following 18 countries: Australia, Austria, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, UK and the U.S. Observations are available from 1952-1995 for most countries, but begin in 1971 for Germany. We define $RMSE = \sqrt{(\sum_{t=1}^{T}(\hat{w}_t - w_t)^2)/T}$. and similarly, $RMAE = \sqrt{(\sum_{t=1}^{T} \mid \hat{w}_t - w_t \mid)/T}$.

Figure1: Medians and Interquartile Ranges for Growth Rates of Real Output (A), Real Money (B) and Real Stock Prices (C): 1954-1994.[2]

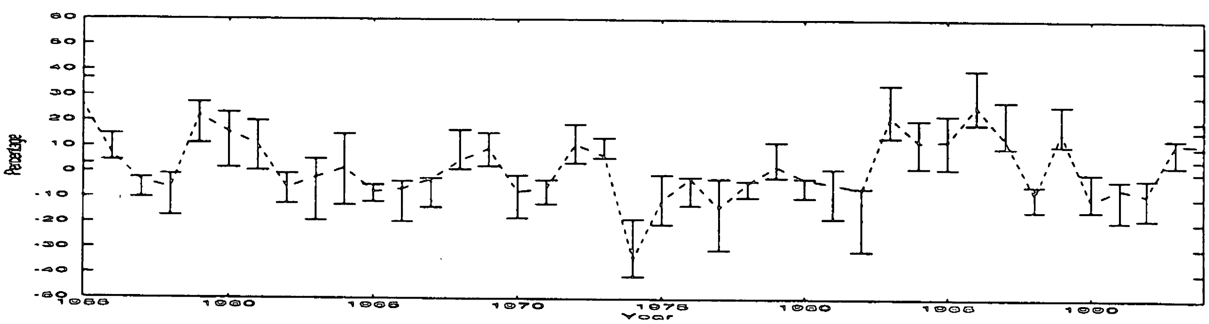
(A) Growth Rates of Real Output



(B) Growth Rates of Real Money



(C) Growth Rates of Real Stock Prices



[2]The dashed line connects the annual median growth rates and the vertical lines give the interquartile ranges.

## References


[1] Espassa, A., "Comment on 'Time Series Analysis, Forecasting and Econometric Modeling: The Structural Econometric Modeling, Time Series Analysis (SEMSTA) Approach'," *Journal of Forecasting* 13 (1994).

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

[3] Lesage, J. P., "Forecasting Turning Points in Metropolitan Employment Growth Rates Using Bayesian Exponentially Weighted Regression, Time-Varying Parameter and Pooling Techniques" *Journal of Regional Science* 30(1990), 533–548.

[4] Min, C-K. and A. Zellner, "Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates," *Journal of Econometrics* 56 (1993), 89–119.

[5] Palm, F. C. and A. Zellner, "To Combine or Not to Combine? Issues of Combining Forecasts," *Journal of Forecasting* 11 (1992), 687–701, reprinted in Zellner (1997).

[6] Zellner, A., *Bayesian Analysis in Econometrics and Statistics: The Zellner View and Papers,* Edward Elgar Publ. Ltd., U.K. and U.S., 1997.

[7] Zellner, A., "Time Series Analysis, Forecasting and Econometric Modeling: The Structural Econometric Modeling, Time Series Analysis (SEMSTA) Approach," *Journal of Forecasting* 13 (1994), 215-233.

[8] Zellner, A. and C. Hong, "Forecasting International Growth Rates Using Bayesian Shrinkage and Other Procedures," *Journal of Econometrics* 40 (1989), 183–202, reprinted in Zellner 1997.

[9] Zellner, A., C. Hong, and G.M. Gulati, "Turning Points in Economic Time Series, Loss Structures, and Bayesian Forecasting," in S. Geisser, J. Hodges, S.J. Press, and A. Zellner, eds., *Bayesian and Likelihood Methods in Statistics and Econometrics: Essays in Honor of George A. Barnard* Amsterdam: North-Holland 371–393.

[10] Zellner, A., C. Hong and C-k Min, "Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time- Varying Parmaeter, and Pooling Techniques," *Journal of Econometrics* 49 (1991), 275–304, reprinted in Zellner 1997.

[11] Zellner, A., J. Tobias and H. Ryu, "Bayesian Method of Moments Analysis of Time Series Models with an Application to Forecasting Turning Points in Output Growth Rates," H.G.B. Alexander Research Foundation, U. of Chicago, 1998.

DEPARTMENT OF AGRICULTURE AND
RESOURCE ECONOMICS AND POLICY

...RKELEY

CALIFORNIA AGRICULTURAL EXPERIMENT STATIO

University of California