Publications and Research   **Hunter College**

1976

Aggregate Versus Subaggregate Models in Local Area Forecasting

D. M. Dunn

William (Bill) H. Williams
CUNY Hunter College

T. L. DeChaine

Aggregate Versus Subaggregate Models in Local Area Forecasting

D. M. DUNN, W. H. WILLIAMS and T. L. DeCHAINE*

Should statistical forecasts be constructed by aggregating *data* to each level for which forecasts are required or aggregating the *forecasts* from the lower levels? The relevant literature suggests no general answer. In this study using actual data, forecasts aggregated from lower-level modeling were found best.

1. INTRODUCTION

In this article the effects of aggregation on forecast accuracy are investigated. The application of interest is the forecasting of telephone demand in certain metropolitan areas. In the telephone industry, forecasts are required for many different hierarchical levels of aggregation. This situation makes it natural to ask whether forecasts should be based on data aggregated to each level for which forecasts are required, or constructed by aggregating forecasts from the lower levels.

This problem does not have a general mathematical solution. Unfortunately, except under very stringent and (generally) unrealistic assumptions, the precise distribution of forecast errors is unknown. Therefore, any practical solution to this problem will involve empirical comparisons of alternative formulations. Since more parameters are being estimated, one might expect that metropolitan area forecasts derived from the subaggregate forecasts would have less variance than the forecasts based directly on the overall metropolitan data. But which forecasts are more accurate?

In this article results from studies in two medium-sized U.S. cities are presented. The generality of these results lies in the fact that the forecasts are *not* always most accurately developed by using data which are aggregated to the level for which forecasts are needed. Indeed, for this particular example, forecasts based on *sub*aggregate data and models were more accurate than forecasts derived directly from aggregated data. Hence, in some problems it may be worthwhile to collect data and construct separate forecasts for subaggregate regions.

The literature on aggregation effects is increasing. Theil [13, 14] derived detailed conditions for the bias which can be introduced when aggregating from microeconomic relations to macroeconomic relations, and Green [10] produced a survey of work before 1964. In 1960 Grunfeld and Griliches [11] published an important article in which they argued that it may well be easier to specify aggregate equations, and hence, that aggregate models could be more accurate. Later, Edwards and Orcutt [7] and Orcutt, Watts and Edwards [12] argued that loss due to aggregation could be great and generally supported disaggregated models. So did Zellner. He notes in [17, p. 366] that "... aggregating data involves *an important loss of information*, information which can be used to discover new economic hypotheses, and to measure effects which are beclouded by aggregation." More recently, Aigner and Goldfeld [1, 2] studied some models in which "aggregates are measured more accurately than their components" (see [2, p. 114]). At this point a number of things can be said. First, models can be constructed to support either aggregate or disaggregate analysis. Taken as a whole, the research so far gives the impression that one is likely to be better off in practice with disaggregated models than aggregate ones. However, case studies appear to be lacking. Data presented in the literature are simulated and reflect mostly the assumed model characteristics. This led to the present analysis of a real set of data. In addition, nonlinear models are included here, while the models studied in the literature are linear.

The algorithms employed to form the aggregate groups are not a concern as they are, for example, in [8]. In this study and in the class of studies for which the results of this article will be useful, the different levels of aggregation are defined by the problem at hand.

2. DESCRIPTION OF THE PROBLEM

The allocation of construction funds in the Bell System is a major planning function. Successful allocation requires accurate forecasts for many different levels of aggregation (see, e.g., [4, 6, and 15]). Forecasts at high levels of aggregation are necessary to determine the gross financial needs of the individual companies and the System. However, subaggregate local area forecasts are also needed to insure that the additional capacity will be installed in appropriate locations.

For this study, an aggregate level is defined as an *exchange* area (basically, an area where one can make a telephone call at no additional charge). Within each

exchange there may be one or more *wire centers*, each being identified by the physical facility which houses the switching equipment serving a particular geographic area. The subaggregate level is defined by the wire centers.

This study utilizes data supplied by the Michigan Bell Telephone Company from the exchanges of Flint and Grand Rapids. These exchanges were previously subjected to extensive research in [6], and, therefore, offered direct comparisons with the new aggregated exchange forecasts. The Flint exchange has five wire centers, while Grand Rapids has eight.

The data for each wire center (number of telephones in service each month) were chosen to reflect the residential telephone demand for the period from January 1954 through June 1970.

3. CHARACTERISTICS OF THE DATA

An examination of the number of main telephones in service (adjusted to the current physical boundaries of the wire centers) at the end of each month reveals a wide diversity of behavior in the wire center demand for residential telephones. Not only are the patterns of the behavior different across exchanges, but they also seem to be different within each exchange.

For example, in the large wire centers such as Grand Rapids Monroe (see Figure A), there is erratic behavior with an evident seasonal pattern. In comparison, wire centers such as Grand Rapids Empire (see Figure B) show remarkably stable growth, with little seasonality. These differences in behavior, both in seasonality and trend, suggest that improved aggregate forecasts may indeed be achieved by considering the subaggregate wire

A. Grand Rapids Monroe Residence Telephones in Service
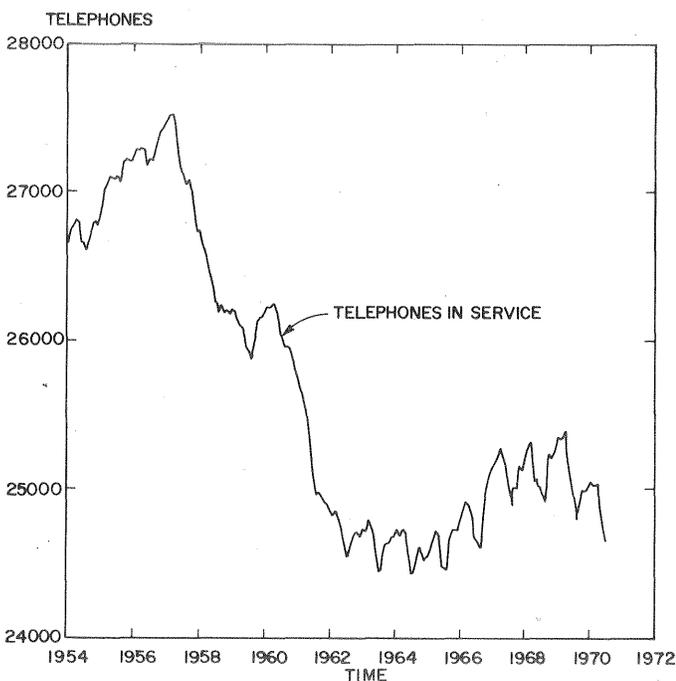
B. Grand Rapids Empire Residence Telephones in Service
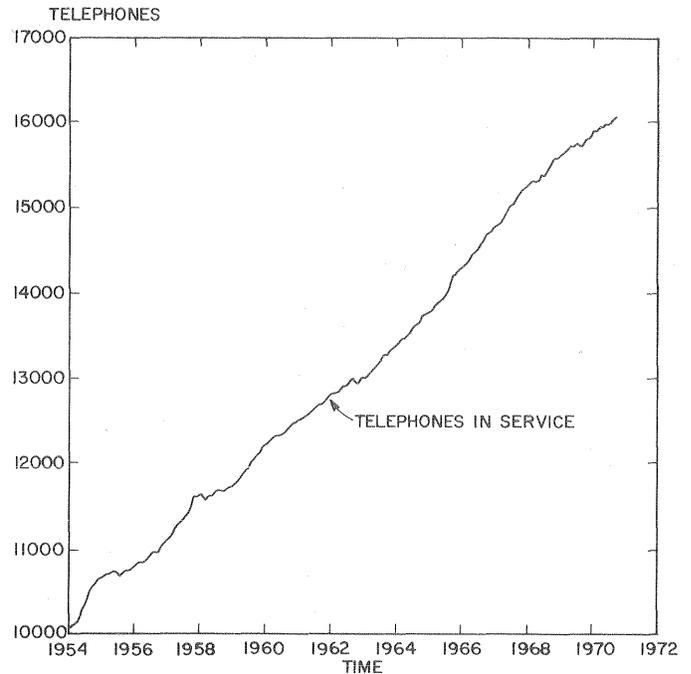
centers separately. It appears that individual wire centers (particularly those which exhibit stable growth patterns) may be more easily forecasted when examined separately.

To gain further insights into the structure of the wire center series, standard analysis techniques, such as plotting, differencing, autocorrelation and spectral analysis were employed. In summary, the wire center series seem to exhibit no strong autoregression or moving average behavior. Some of the series are seasonal, but this is effectively removed by a 12-month difference. Using this analysis as a basis, forecasting models were developed.

4. FORECASTING AT SUBAGGREGATE LEVELS

The development of appropriate forecasting procedures requires both a careful analysis of the series to be forecast and an understanding of the forecast environment. Consequently, this section is divided into three subsections: motivation, methodology, and description of results.

4.1 Motivation

Forecasting at the subaggregate level often requires forecasts for *many* individual areas or submarkets. For example, in Michigan Bell over 300 wire center forecasts are needed. Hence, it would be very useful to have a class of forecasting models which would work well for many of the demand patterns observed at subaggregate levels. Furthermore, these forecasts must be updated regularly to keep abreast of changing demand patterns, often many times a year. This requires that updated forecasts be easy to obtain, without a long and detailed analysis. Finally, demand patterns for local areas or specific markets can often be very dynamic. Our data analysis

supports this conjecture for the wire center data, and necessitates models which respond quickly to changes in the pattern of demand.

4.2 Forecasting Methodology

Since exogenous data are difficult and costly to obtain on a continuing basis for specific subaggregate areas, we will focus our attention on models which use only the history of the series.

The autoregressive-integrated-moving average (ARIMA) models described in [3] form a general class of forecasting models which are useful for a wide variety of time series. In the present context, a major problem with this class of models is that it requires a nonroutine process to choose a specific model from the class. Furthermore, the fitting is nonlinear and requires the maintenance of a substantial data history for updating of parameter estimates. Consequently, these general models are usually too complex for the subaggregate forecasting problem.

However, two subclasses of the ARIMA models hold promise. The autoregressive models for differenced series can be estimated using current regression packages and allow very accurate updating (see [5]). Furthermore, to perform the updating one need not retain the entire data history, yielding a considerable saving in storage space.

The autoregressive models are basically of the form

$$\Phi(B)\nabla^d z_t = a_t ,$$

where $\Phi(B)$ is a polynomial in B, the backward shift operator (e.g., $Bz_t = z_{t-1}$), ∇^d is the difference operator of order d, z_t is the series to be forecast, and a_t is a white-noise process. In particular, three models tended to be most useful:

Model	Notation
$(1 - \phi_1 B - \phi_2 B^2)\nabla z_t = a_t$	AR3
$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^{12})\nabla z_t = \theta_0 + a_t$	AR4
$(1 - \phi_1 B - \phi_2 B^{12} - \phi_3 B^{13})\nabla z_t = a_t$	AR3N

Another useful subclass of models are the IMA (k, k) and extensions. Previous work (see [6]) has indicated that these models (which are equivalent to exponential smoothing models; see [9]) are useful in forecasting series with dynamic behavior similar to that seen at the subaggregate level. They have the added advantage of being easy to fit, easy to update, and require minimal storage of summary measures and not the entire data history. The three most useful smoothing models, described in [6], were: first-order exponential smoothing (SMPL); adaptive exponential smoothing (ADP); and adaptive exponential smoothing modified for seasonal series (ADPS).

4.3 Model Results

As in previous research [6], forecasts are evaluated in terms of the mean absolute deviation (MAD) and root mean square (RMS) of the sequence of forecast errors. These errors are computed by moving the forecast model over the data to create the series of errors that would have occurred if the model had been actually used during that time period (for details, see [6]). In addition to the MAD and RMS evaluation of the forecasts, empirical prediction intervals (developed by Williams and Goodman [16]) are also computed.

The results of these various forecasting techniques are summarized in Table 1. Data are presented for the nine major wire centers; the four smallest wire centers are not included because of their very short data history. For each of the nine wire centers, the MAD and the RMS are presented for the model which minimized the MAD forecast error in that particular wire center for a lead time of 12 months. Table 1 also contains the MAD and RMS forecast errors scaled, respectively, by the mean absolute deviation and the root mean square of the actual changes with the same lead time. These annual changes are computed in the same way and for the same time intervals for which forecasts are generated. In this form, smaller numbers are better, and we see that in six of the nine offices there has been a worthwhile reduction in variability.

1. Best Wire Center Forecast Models (12 Month Lead Time)

Wire center	Lead time	Forecast model	Forecast error		Scaled forecast error	
			MAD	RMS	MAD	RMS
Flint						
Cedar	12	AR4	585.2	716.1	0.84	0.84
Sunset	12	AR4	378.8	459.4	0.69	0.70
Pilgrim	12	ADPS	102.5	124.3	0.20	0.24
Grand Rapids						
Monroe	12	AR3N	242.4	340.0	0.94	0.92
Empire	12	ADP	91.1	116.7	0.24	0.28
South Hall	12	SMPL	141.6	180.9	0.42	0.48
Lenox	12	ADPS	128.3	162.4	0.16	0.21
West	12	ADPS	33.2	46.1	0.18	0.24
East	12	AR3	48.1	62.7	0.17	0.22

Model parameter estimates are not included because of the forecasting procedures used. By moving the models over the data, new parameters are estimated as each new observation is included. Hence, for the autoregressive models, the data predicted are not used in estimating the prediction models. For the smoothing models, the single parameter was selected to minimize the moving overall MAD forecast error (using the same procedure of moving the model over the data).

5. COMPARISON OF ALTERNATIVE AGGREGATE EXCHANGE FORECASTS

In this section the best exchange forecasts obtained by examining the area as a whole are compared with those generated by aggregating the "best" individual wire center forecasts. These may be compared with the previous best exchange forecasts by Dunn et al. [6].

A comparison of the exchange forecast criteria for a fixed 12-month lead time (see Table 2) shows an improvement in the forecasting accuracy in every series and for both error criteria. More important is the character of the improvements. Note that for Flint, which has always been very difficult to forecast, the improvements due to disaggregation are greater. As such, this gives promise that for the "difficult" areas, improved forecast accuracy may be obtained by looking at subaggregate data directly.

2. Comparison of Exchange Forecasts
(12 Month Lead Time)

Exchange	Error criterion	Method		Percent improvement
		Aggregate of major [a] wire center forecasts	Single overall forecasts	
Flint	MAD	889.5	1145.3	22.3
	RMS	1081.5	1454.3	25.6
Grand Rapids	MAD	409.5	498.3	17.8
	RMS	540.3	567.1	4.7

[a] Omitted wire centers have short data histories, exhibit very stable growth, and amount to a very small percentage of the total demand.

6. SUMMARY AND CONCLUSIONS

In this article the development of forecasts of telephone demand for the exchange areas of Flint and Grand Rapids, Michigan, was described. Aggregated forecasts were based on forecasts from the geographically smaller wire center areas and were then compared with those developed directly from data aggregated to the exchange level area (see [6]). The aggregated forecasts were found to be more accurate.

The general conclusion is that to obtain the most benefit from statistical forecasting models, every effort should be made to obtain and analyze subaggregate data, especially if the subaggregate areas are expected to vary. The natural tendency in most studies is to base the forecasts on data already aggregated to the level for which forecasts are required. This may not always be the best procedure and certainly was not in the project discussed here.

[Received February 1974. Revised July 1975.]

REFERENCES

[1] Aigner, D.J. and Goldfeld, S.M., "Simulation and Aggregation: A Reconsideration," *Review of Economics and Statistics*, 55, No. 1 (February 1973), 114–8.

[2] —— and Goldfeld, S.M., "Estimation and Prediction from Aggregate Data When Aggregates Are Measured More Accurately Than Their Components," *Econometrica*, 42, No. 1 (January 1974), 113–34.

[3] Box, G.E.P. and Jenkins, G.M., *Time Series Analysis*, San Francisco: Holden-Day, Inc., 1970.

[4] Chaddha, R.L. and Chitgopeckar, S.S., "A 'Generalization' of the Logistic Curves and Long-Range Forecasts (1966–91) of Residence Telephones," *Bell Journal of Economics and Management Science*, 2, No. 2 (Autumn 1971), 542–60.

[5] Chambers, J.M., "Regression Updating," *Journal of the American Statistical Association*, 66, No. 336 (December 1971), 744–8.

[6] Dunn, D.M., Williams, W.H. and Spivey, W.A., "Analysis and Prediction of Telephone Demand in Local Geographic Areas," *Bell Journal of Economics and Management Science*, 2, No. 2 (Autumn 1971), 561–76.

[7] Edwards, John B. and Orcutt, Guy H., "Should Aggregation Prior to Estimation Be the Rule?" *Review of Economics and Statistics*, 51, No. 4 (November 1969), 409–20.

[8] Fisher, Walter D., *Clustering and Aggregation in Economics*, Baltimore: The Johns Hopkins Press, 1969.

[9] Goodman, M.L., "A New Look at Higher-Order Exponential Smoothing for Forecasting," *Operations Research*, 22, No. 4 (July–August 1974), 880–8.

[10] Green, John H.A., *Aggregation In Economic Analysis, An Introductory Survey*, Princeton, N.J.: Princeton University Press, 1970.

[11] Grunfeld, Yehuda and Griliches, Zvi, "Is Aggregation Necessarily Bad?" *The Review of Economics and Statistics*, 42, No. 1 (February 1960), 1–13.

[12] Orcutt, Guy H., Watts, Harold W. and Edwards, John B., "Data Aggregation and Information Loss," *American Economic Review*, 58, No. 4 (September 1968), 773–87.

[13] Theil, Henri, *Linear Aggregation of Economic Relations*, Amsterdam: North-Holland Publishing, Co., 1954.

[14] ——, *Principles of Econometrics*, New York: John Wiley & Sons, Inc., 1971.

[15] Thompson, H.E. and Tiao, G.C., "Analysis of Telephone Data: A Case Study of Forecasting Seasonal Time Series," *Bell Journal of Economics and Management Science*, 2, No. 2 (Autumn 1971), 515–41.

[16] Williams, W.H. and Goodman, M.L., "Constructing Empirical Confidence Limits for Economic Forecasts," *Journal of the American Statistical Association*, 66, No. 336 (December 1971), 752–4.

[17] Zellner, Arnold, "On the Aggregation Problem: A New Approach to a Troublesome Problem," in K.A. Fox *et al.*, eds., *Economics Models, Estimation and Risk Programming*, New York: Springer-Verlag, 1969, 365–74.