MONASH
UNIVERSITY

AUSTRALIA

A COMPUTERIZED SYSTEM FOR FORECASTING SPARE PARTS SALES:

A CASE STUDY

R. D. Snyder

Working Paper No. 11/89

December 1989

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 976 5

A COMPUTERIZED SYSTEM FOR FORECASTING SPARE PARTS SALES: A CASE STUDY

R. D. Snyder

Working Paper No. 11/89

December 1989

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS AND POLITICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Abstract

The focus of this paper is a computerized sales forecasting system for the control of automotive spare parts. The logic of the forecasting method, a refinement of exponential smoothing, is outlined together with a method for monitoring forecast errors. Experiences in developing, implementing and operating the system are also described.

Keywords: Sales forecasting, inventory control, exponential smoothing, Kalman filtering.

1. INTRODUCTION

A spare parts distributor in Victoria owned by a major Japanese car manufacturer is the focus of this paper. Its function was the maintenance of adequate stocks for about 60,000 automobile parts to meet demand from local dealers and traders. Its situation was complicated by long replenishment lead times and uncertainty in demand. Replenishment orders, placed once a month with the parent company, took about two months to arrive in Australia where industrial disputes at the docks could further cause delays. The local distributor was expected to maintain a high service to customers with a minimal investment in stock under these rather difficult circumstances.

In 1985 the Parts Development Manager, who had overall responsibility for the inventory system, decided that performance was far from satisfactory. There was an excess supply of many inventories coupled with widespread stockouts of others. Practices current at the time amounted to little more than rules of thumb developed in the distant past by experienced employees using intuition and personal judgement rather than formal analysis. These rules appeared to be failing quite badly in the conditions current at the time.

Having obtained an MBA in England, the Parts Development Manager was conversant with basic scientific inventory theory (Plossl, 1985) and the possibilities it held for improving the situation. He therefore assembled a small team of analysts and programmers from the data processing department to develop and implement a computer system for controlling inventories based on the theory that he had learnt during his studies. Part of the task required the development of a subsystem for forecasting sales and estimating associated standard deviations for the purpose of determining appropriate levels of safety stock. Feeling the need for external assistance in this matter, the author was commissioned to investigate the situation and to suggest an appropriate forecasting strategy. Over time this arrangement evolved into one involving the development of the forecasting subsystem and the provision of advice during its implementation. The purpose of this paper is to describe the experience and document the associated methods.

2. SYSTEM OVERVIEW

A critical factor in the design of the forecasting subsystem was the question of scale. The sheer size of the task of generating up to 60,000 forecasts on a monthly basis precluded the possibility of extensive management involvement. No matter how desirable it is to bring management experience and knowledge of market conditions to bear in the forecasting process, economic considerations dictated that this was impossible on any extensive scale. The only cost effective approach was a computerized forecasting system based on methods from statistical time series analysis.

It was recognized, however, that a total reliance on automated forecasts was not entirely satisfactory. Structural changes, which can be quite common in the market place, diminish the information value of historical data upon which time series approaches are based and hence can create situations where automated forecasts are quite out of kilter with market conditions current at the time. It was also recognized that there was a need to detect those annoying data entry errors which inevitably occur in the daily operations of order entry systems so as to avoid the distorting effects on the time series forecasts. Thus the system depicted in Figure 1 was established with a subsystem to monitor forecast errors at the end of each month of operation and to indicate the need for manual intervention in cases with unduly large errors. The basic idea is that an executive can investigate the causes of a large error and take what is seen to be the best action in the circumstance. Data entry errors can be detected and corrected, returning control to the computer once this is done. Manual forecasts can be generated where appropriate, taking account of prevailing market intelligence, and bypassing the automatic forecasting subsystem in the process. Either way, forecasts are then fed to the inventory control subsystem for use in determining re-order quantities prior to the start of a new month of operation.

3. FORECASTING SUBSYSTEM

3.1 Statistical Framework

For a system like this there are, potentially, a wide range of statistical time series methods from which to choose. Seasonal methods were rejected because most automobile parts fail to display significant seasonal effects. Growth curve approaches based on exponential, Gompertz and logistics functions and their ilk were rejected because the particular path followed by the sales of a part over its life cycle of 10 years is not known a priori and is difficult to determine with only partial sales histories. Leading indicators based on the sales of similar parts for earlier models of cars also proved to be unreliable because of the unpredictable effects of design changes. Thus methods based on structural considerations (Harvey and Todd, 1983; Harvey, 1984, 1985; Harvey and Durbin, 1985) seemed to be out of the question.

Given that a particular life cycle pattern will be observed in most cases by the end of 10 years but that its form, particularly in the early stages, is difficult to detect, it seems sensible to adopt a statistical model which assumes that the underlying level of the series changes over time, albeit in an uncertain way. This leads to the concept of the following local level model, expressed for convenience in a form of pseudocode:

sales[now] = level[now] + irreg[now] (1.1)

level[next] = level[now] + alpha * irreg[now] (1.2)

where **next = now + 1** and where the primary source of randomness, namely the irregular components, are normally and independently distributed random variables with a common mean and standard deviation of 0 and **sigma** respectively. Here **alpha** is a parameter which determines the rate of change in the underlying level. When **alpha** = 0 there is no change so that the associated series is stationary. In the case where **alpha** = 1, the model reduces to the random walk **sales[now] = sales[last] + irreg[now]** where **last = now - 1.** Generally, larger values of **alpha** are required for series with a more pronounced pattern.

The local level model has its counterpart in the Box and Jenkins (1976) framework. By differencing (1.1) and eliminating the level variables with (1.2), the following ARIMA(0,1,1) model is obtained

sales[now] - sales[last] = - delta * irreg[last] + irreg[now] (2.1)

where

delta = 1 - alpha. (2.2)

This model, as demonstrated by Muth (1960) and Box and Jenkins (1976), underpins simple exponential smoothing (Brown, 1959; Holt, 1957), a technique that has been quite widely used in inventory control applications with considerable success (Gardner, 1985). Simple exponential smoothing relies on the repetitive application of the following recursive scheme:

pred.error[now] = sales[now] - pred.sales[now] (3.1)

est.level[next] = est.level[now] + alpha * pred.error[now] (3.2)

pred.sales[next] = est.level[next] (3.3)

Because **alpha** determines the extent to which the data is smoothed, it is commonly referred to as the smoothing parameter.

The exponential smoothing predictor has the closed form solution

$$\mathbf{pred.sales[next] = alpha} * \sum_{\mathbf{age=0}}^{\mathbf{oldest}} \mathbf{(delta) **age * sales[now - age]}$$

$$+\ \mathbf{(delta)**(oldest + 1) * pred.sales[now - oldest]} \qquad (4)$$

This indicates that the prediction is a geometrically weighted linear function of the observations and ~~the~~ a seed prediction. When - 1 < **delta** < 1 the weights decline with age and the impact of the seed prediction declines. For this reason **delta** is sometimes referred to as the discount factor. In semi-infinite samples (i.e. when **oldest** → ∞) the influence of the seed prediction disappears and the current prediction becomes a genuine average of the data (i.e. the weights sum to one).

Exponential smoothing is an adaptive method for estimating the local level model. A comparison of the closed form solutions for the level and its estimate suggests that any error in the seed estimate perpetuates itself, albeit in discounted form, in later estimates. Apart from this, both closed form expressions are identical, suggesting that in semi-infinite samples the level and its estimate become identical. No other estimator can do better than this and so in semi-infinite samples exponential smoothing, despite its adaptive nature, can be viewed as optimal for the local level model.

3.2 Kalman Filter

In reality, samples are finite. This was particularly true for the company concerned which, because of the usual financial restrictions on the availability of storage media, could only at most maintain three years of data in its data bases. Added to this, new cars were released annually, necessitating the introduction of new parts with no sales histories. In the year that this work was undertaken, the release of its new range of cars was expected to result in the introduction of 20,000 new parts to its system. In these circumstances it is necessary to employ a technique which extracts as much information as possible from the limited available data. This rules out exponential smoothing because, despite its satisfactory large sample properties, it is not optimal in small samples. As indicated by (4), the predictions generated by exponential smoothing ultimately depend on an arbitrary seed value, the influence of which declines exponentially with larger sample sizes. In small samples, however, the seed value can dominate the sample observations in forming the prediction and as a consequence a poor seed value can distort the predictions quite markedly.

An alternative strategy is to seek a technique which minimizes the mean squared one step ahead prediction error assuming that the data is generated by the local level model. It turns out that this local level model is a special case of the state space framework outlined in Snyder (1985) and that it can therefore be estimated in an optimal fashion by the Kalman filter. There is a general perception in many circles that Kalman filtering is not a particularly easy technique and its use has consequently been mainly restricted to technical applications. However, in the case of the local level model it takes a particularly simple form not all that dissimilar from exponential smoothing - see Snyder (1988) for derivation. The major change is that, in effect, **alpha** becomes time dependent, and is generated by a formula which ensures that the mean squared one step ahead forecast error is minimized. More specifically:

pred.error[now] = sales[now] - pred.sales[now] (5.1)

est.level[next] = est.level[now] + alpha[now] * pred.error[now] (5.2)

pred.sales[next] = est.level[next] (5.3)

delta[next] = 1/(delta + 1/delta - delta[now]) (5.4)

alpha[next] = 1 - delta[next] (5.5)

where **delta[1] = 0**. In effect **alpha[now]** and **delta[now]** are short-run smoothing and discount parameters and it can be established that they converge to long run values of **alpha** and **delta** respectively. Thus in large samples there is effectively no difference between this simple Kalman filter and simple exponential smoothing. It is only in small samples that they differ and they do so because this simple Kalman filter makes better use of the limited information in this circumstance.

Examples

1. When **alpha = 0** the local level model collapses to a stationary global level model. Here **delta = 1** so that (5.4) and (5.5) between them yield successive values of 1, 1/2, 1/3, 1/4, ... for the short-run smoothing parameter **alpha[now]**. In this circumstance the Kalman filter becomes a recursive method for calculating a simple average. Compare this with

simple exponential smoothing which fails to properly accommodate this special but important case.

2. When **alpha** = 1 the local level model converts to a random walk. It can be established that (3.4) and (3.5) then yield successive values 1,1,1,... for **alpha[now]**. The filter reduces to the naive, but often powerful, forecast method **pred.sales(next) = sales(now)**.

3. When **alpha** = 0.1 the successive values of **alpha[now]** are 1, 0.5028, 0.3395, ..., 0.1, 0.1, ... This illustrates the convergence of the Kalman gain to its long run value. In the early stages it appears that the Kalman filter places an approximately equal weight on each observation and in this sense it can be viewed as a refinement of a method proposed by Taylor (1981) for initializing exponential smoothing.

3.3 Standard Deviations

The inventory module not only required forecasts of sales but also estimates of standard deviations for safety stock calculations. In applications of exponential smoothing the one-step ahead forecast errors are often used as the basis for estimating this parameter. However, it should be recognized that one-step ahead forecast errors form a heteroskedastic series because forecasts are less reliable and hence errors tend to be larger in small samples. It can be established that the mean square error of a forecast is given by

$$\mathbf{mse[now] = delta * sigma/delta[now]} \qquad (6)$$

and so the best estimate of sigma must be

$$\mathbf{est.\ sigma} = \Big(\sum_{\mathbf{period=2}}^{\mathbf{now}} \mathbf{delta[period] * error[period] **2/delta)/last} \qquad (7)$$

It was this estimate which was fed to the inventory module.

3.4 Choice of Long Run Smoothing Parameter

The value of the parameter **alpha** or equivalently **delta** must be specified before employing the Kalman filter. The choice of **alpha** affects the quality of the associated forecasts and is therefore an issue also requiring attention. It is pertinent to note, along similar lines to Burridge and Wallis (1988), that if |**delta**| > 1 then 1/**delta** can be used in place of **delta** in the Kalman filter to yield exactly the same forecasts. Accordingly, it is sufficient to restrict the choice of **delta** and hence **alpha** to the equivalent ranges |**delta**| $\leq$ 1 and 0 $\leq$ **alpha** $\leq$ 2 respectively. (Note the strong similarity with the invertibility condition for an ARIMA(0,1,1) model in Box and Jenkins (1976).)

Even within this range, however, there are further difficulties in finding a sensible value for **alpha**. Evaluation criteria such as the likelihood function are highly nonlinear functions of **alpha** and may often possess local minima. Accordingly, to find the best value of **alpha**, it is necessary to use grid search methods where the Kalman filter is applied repetitively for a range of trial values of **alpha**. Since the computational loads of a procedure like this are rather large for a single series, they are prohibitive for 60,000 lines. In these circumstances the decision was made to expedite matters by using a single global value for **alpha** across the entire range of inventory. To keep computational loads within acceptable bounds, the process of finding a value for this global **alpha** was restricted to a small random sample of their inventories. The method involved the estimated value of sigma and the simple average of each sales series to compute a measure of relative risk reminiscent of a coefficient of variation i.e.

rel.risk = est.sigma/av.sales (8)

This measure was calculated for all items in the sample and aggregated for each trial value of **alpha**. The associated program then selected the value of **alpha** with minimal aggregate relative risk across the sample.

4. MONITORING FORECAST QUALITY

Tracking signals (Trigg 1964; McClain, 1988) are often used in inventory applications of exponential smoothing to monitor forecast quality. However, the distribution theory is almost non-existent for this approach and it is therefore not clear how the critical values of these tracking signals should be selected. Instead an alternative approach, based on the current one-step forecast error, was developed.

A consequence of the optimality properties of the Kalman filter is that the prediction errors are not only normally distributed with a common mean of zero, but they are also statistically independent. Furthermore, given that the estimated mean squared error last period, based on (6) and (7), depends on only **error[1]...error[last]**, it is statistically independent of **error[now]**. Hence, for given **alpha**, the ratio

t = error[now]/est.mse[last] (9)

has a t-distribution with **last** degrees of freedom. The absolute value of **t** in (9) was used as the basis of a statistical test for monitoring the forecast quality. When the process is under control this statistic should be close to zero. However, beyond a critical value determined from a t-distribution table after specifying the significance level of the test, the item forecasts are deemed to be out of control. This idea, or at least a simplified variation of it to expedite matters, was incorporated into a computer program to generate an exception report of all those items deemed to be out of control. The items were ranked in this list according to the size of the absolute value of **t**. This was designed to enable the executive charged with the monitoring function to rationally allocate his effort, given the scarcity and value of his time, so that priority could be given to those items most out of control.

5. IMPLEMENTATION AND PERFORMANCE

The implementation of the forecasting modules was carried out with few hitches. The main problem that arose occurred because of the use of different computer languages. The entire system of the company was programmed in COBOL, a language quite unsuitable for scientific applications, whereas FORTRAN 77 was used for the forecasting modules. Given that COBOL does not contain conventional floating point arithmetic, linking became a problem. Many of the difficulties that arose occurred because COBOL and FORTRAN programmers have considerable difficulty talking to and understanding each other. However, eventually, after many unnecessary mishaps, a successful linkage was achieved. With hindsight, most of the difficulties encountered along the way could have been avoided by employing the services of one of those scarce individuals with a working knowledge of both languages.

Budget limitations meant that no formal means were used to evaluate the performance of the system once it became operational except for the compilation of some aggregate statistics on stock levels and stockouts. By this time the development work had ceased, the contract had run out, and so the contact with the company ceased. Matters had gone reasonably smoothly and there was no apparent need for further advice.

Two years later contact was renewed when the company took over some major distributors in other States, effectively doubling the size of their market overnight. They now had to control inventories in a number of geographically dispersed locations and were seeking advice on how to best blend the new operations. It emerged that they had been quite satisfied with the performance of the system prior to the takeover but now felt that a reappraisal was required in the light of their new circumstances.

On investigation it transpired that the company did have a serious problem. Soon after the takeover, the exception report from the monitoring subsystem had suddenly exploded in size. Instead of a few pages of printout containing a list of a few hundred lines for investigation, the executive concerned was suddenly confronted with a massive printout the thickness of which could be measured in centimetres. Something was clearly wrong. The executive concerned understood this but feared that it might be perceived as a reflection of his own performance. Furthermore, given his other duties, he quite simply did not have the time available to investigate all the items in the report on an individual basis. His solution at the time was to shelve the printout and stay silent about the situation in the hope that the problem would automatically resolve itself. Unfortunately, the same thing happened in the second month, the executive continued to ignore the warning signs and senior management remained ignorant of the impending disaster. Then about two weeks later the moment of truth arrived. Stocks right across their range began to run out. It was only then, for the first time, that the Parts Development Manager had become aware that the company had a problem. He was unfortunately unable to identify its cause.

During the subsequent investigation by the author it transpired that no attempt had been made to modify the forecasts to allow for their new situation. Ideally, sales histories should have been obtained from the new subsidiaries, combined with the company's existing sales records, and the forecasts should have been regenerated with the revised data. Instead the system had operated with only the sales histories of Victoria to predict sales for the expanded market. Sales levels were consequently being seriously underestimated. In fairness to the company's management, differences in computer systems may have made it quite difficult and expensive to collect and process the old sales records of the new subsidiaries. But it would have been a relatively simple matter to double the size of the old forecasts across the range of their inventory to obtain more plausible results. Unfortunately management did not understand the need for this at the time because they were under the reasonable but false allusion that the forecasts would adjust automatically after a few months of operations in their new environment. Exponential smoothing and its refined counterpart in this paper are to some extent adaptive in nature. However, the rate at which forecasts adapt depends on the size of **alpha**. It transpired that the process used to select the value of the smoothing parameter described earlier in the paper had yielded a value of zero. In other words, the company had in effect been using a classical simple average to compute their forecasts. Little adaption was possible with such a low smoothing parameter.

This outcome highlighted a problem with the method for selecting the smoothing parameter. As stated earlier, over the 10 year life of a part, it is possible to detect an underlying usage pattern which may, in many cases, be quite pronounced. However, only a few years of data was available at most, where such patterns in these restricted samples are less discernable. Accordingly, there was a serious downward bias in the procedure for selecting the value of the smoothing parameter. Given that there was still insufficient data at the time to do anything differently, a recommendation was made to set the value of alpha to 0.1 or 0.2, a recipe recommended by Brown (1959) from many years of experience with exponential smoothing in the field. Possibly even higher values should have been contemplated.

With hindsight it was also possible to identify another potential problem with their system in the area of safety stock determination. The author, at the time, had little involvement in the development of the inventory module but became aware of the fact that they were determining their estimate of the standard deviation of lead time demand by taking the value of **est_sigma** from the forecast routine and multiplying it by the factor **sqrt(leadtime)**. This method depends on the assumption that monthly demands are statistically independent, an assumption which is incompatible with the local level model except when **alpha = 0.** It can be shown that the standard deviation of lead time demand can be severely underestimated when the dependence between monthly demand is ignored - see Johnston and Harrison (1986) and Harvey and Snyder (1989). Consequently, since implementation, the company has probably operated with safety stocks below those required to achieve its service level objectives.

CONCLUSIONS

In this forecasting project we encountered problems associated with the scale of the operations and paucity of data. The response was to implement a simple but robust Kalman filter with low computational requirements and good statistical properties. Despite some difficulties along the way the resulting system proved to be a success and highlighted the point that Kalman filtering has a viable and useful role in business forecasting. At the same time the project indicated the need for an adaptive approach to forecasting and that a total reliance on manual approaches to this problem can break down when users have an incomplete understanding of their system. Accordingly, some automation of the adaptation process seems to be desirable. Just how this could be done remains an open question but preliminary results with a multi-series approach in Snyder, Shah and Lehmer (1988) is suggestive of one possible response to this problem. Finally, we had succeeded in developing and implementing a more statistically satisfactory monitoring system based on the classical t-distribution, exploiting the orthogonality of the one-step ahead forecast errors from the Kalman filter. Overall then, the experience proved to be of value in the sense that it stimulated new responses to an old problem and established the viability of the developments, despite the limitations of budgets and consequent short-cuts during the process.

REFERENCES

Box, G.E.P. and Jenkins, G.M., 1976, *Time Series Analysis : Forecasting and Control*, Holden Day, San Francisco.

Brown, R.G., 1963, *Statistical Forecasting for Inventory Control*, McGraw-Hill, New York.

Burridge, P. and Wallis, K.F., 1988, Prediction Theory for Autoregressive-Moving Average Processes, *Econometric Review*, 7, 65-95.

Gardner, E.S., 1985, Exponential Smoothing : The State of the Art, *Journal of Forecasting*, 4, 1-28.

Harvey, A.C. and Todd, P.J.H., 1983, Forecasting Economic Time Series with Structural and Box-Jenkins Models : A Case Study, *Journal Of Business and Economic Statistics*, 1, 299-315.

Harvey, A.C., 1984, Time Series Forecasting Based on the Logistics Curve, *Journal of the Operational Research Society*, 35, 641-646.

Harvey, A.C., 1985, Trends and Cycles in Macroeconomic Time Series, *Journal of Business and Economic Statistics*, 3, 216-227.

Harvey, A.C. and Durbin, J., 1986, The Effects of Seat Belt Legislation on British Road Casualties : A Case Study in Structural Time Series Modelling, *Journal of the Royal Statistical Society*, A, 149, 187-227.

Harvey, A.C. and Snyder, R.D., 1989, Structural Time Series Models in Inventory Control, *International Journal of Forecasting* (forthcoming).

Holt, C.E., 1957, Forecasting Trends and Seasonals by Exponentially Weighted Averages, *ONR Memorandum No. 52*, Carnegie Institute of Technology.

Johnston, F.R. and Harrison, P.J., 1986, The Variance of Lead Time Demand, *Journal of the Operational Research Society*, 37, 303-308.

McClain, J.O., 1988, Dominant Tracking Signals, *International Journal of Forecasting*, 4, 563-572.

Muth, J.F., 1960, Optimal Properties of Exponentially Weighted Forecasts, *Journal of the American Statistical Association*, 55, 299-306.

Plossl, G.W., 1985, *Production and Inventory Control : Principles and Techniques,* Second Edition, Prentice-Hall, Englewood Cliffs.

Snyder, R.D., 1985, Recursive Estimation of Dynamic, Linear Statistical Models, *Journal of the Royal Statistical Society*, B, 47, 272-276.

Snyder, R.D., 1988, Progressive Turning of Simple Exponential Smoothing Forecasts, *Journal of the Operational Research Society*, 39, 393-400.

Snyder, R.D., Shah, C. and Lehmer, C., 1988, Multi-series Heuristic for Exponential Smoothing, Working Paper No. 9/88, Department of Information Systems, Monash University.

Taylor, S.G., 1981, Initialization of Exponential Smoothing Forecasts, *AIIE Transactions*, 13, 199-205.

Trigg, D.W., 1964, Monitoring a Forecasting System, *Operational Research Quarterly*, 15, 271-274.

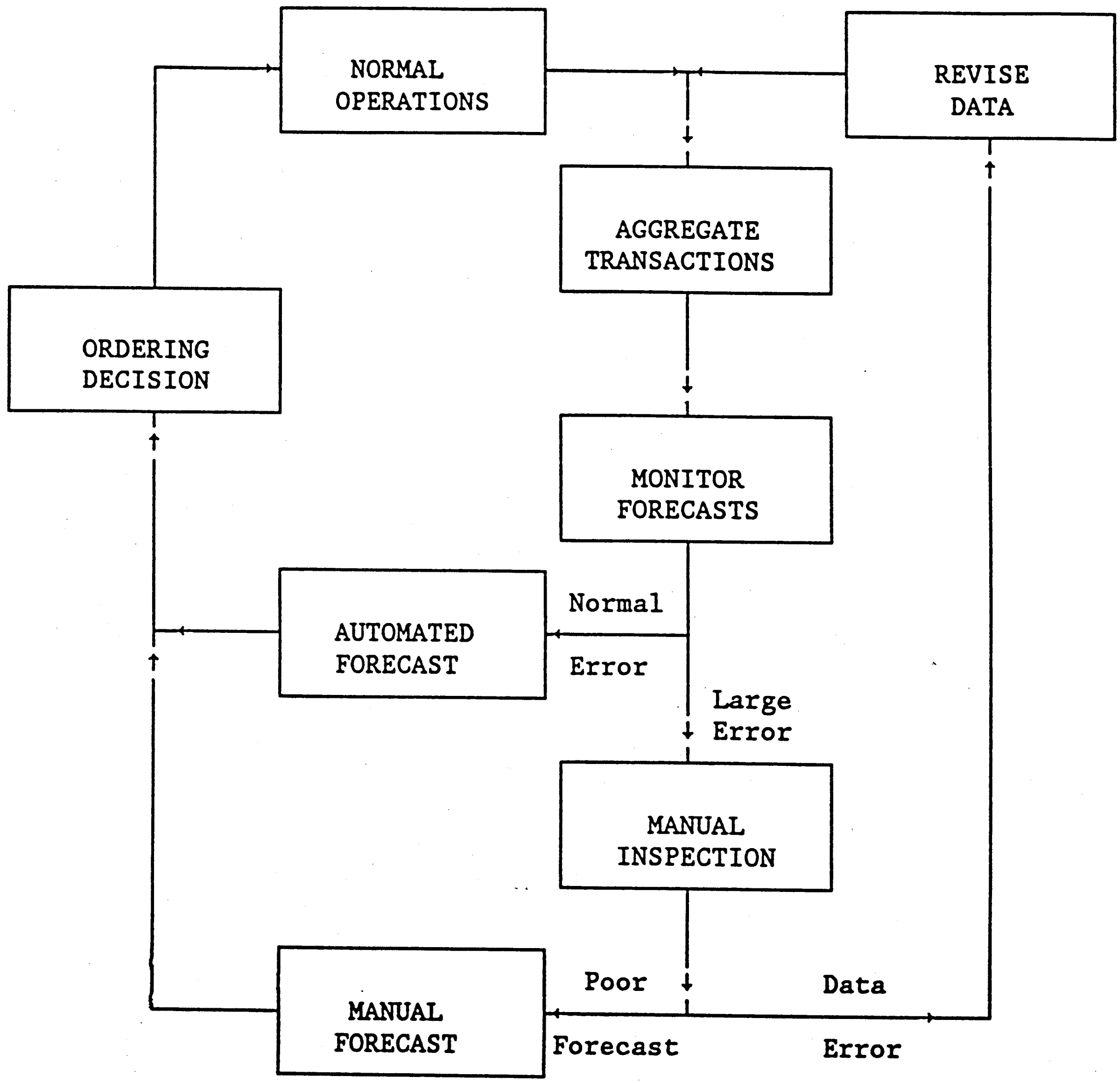

Figure 1. Monthly Cycle for Each Item

MONASH UNIVERSITY

DEPARTMENT OF ECONOMETRICS

WORKING PAPERS

1989

1/89 R. D. Snyder, C. Chah & C. Lehmer, "Multi-series Heuristics for Exponential Smoothing.

2/89 Maxwell L. King and Phillip M. Edwards, "Transformations for an Exact Goodness-of-Fit Test of Structural Change in the Linear Regression Model".

3/89 Muhammad I. Bhatti and Maxwell L. King, "A Beta-Optimal Test of the Equicorrelation Coefficient".

4/89 Muhammad I. Bhatti, "Null Distribution of the Small Sample Mean Correlation Coefficient: An Application to Medical Research."

5/89 Grant H. Hillier, "On the Normalisation of Structural Equations: Properties of Direction Estimators".

6/89 Brett A. Inder, "A New Test for Autocorrelation in the Disturbances of the Dynamic Linear Regression Model".

7/89 Russel J. Cooper, Dilip B. Madan and Keith R. McLaren, "A 'Gormanesque' Approach to the Solution of Intertemporal Consumption Models".

8/89 Keith R. McLaren, "A Variant on the Arguments for the Invariance of Estimators in a Singular System of Equations".

9/89 T.R.L. Fry, "Univariate and Multivariate Burr Distributions: A Survey".

10/89 Merran Evans, "Robustness and Size of Tests of Autocorrelation and Heteroscedasticity to Non-normality".

11/89 R.D. Snyder. "A Computerized System for Forecasting Spare Parts Sales: A Case Study".

12/89 Maxwell L. King and Ping X. Wu. "Small-Disturbance Asymptotics and the Durbin-Watson and Related Tests in the Dynamic Regression Model".

13/89 W.T.M. Dunsmuir and R.D. Snyder. "ABC Analysis in Inventory Control - The Issue of Stability".