A Validation of the Bass New Product Diffusion Model in New Zealand

Malcolm Wright, Clinton Upritchard and Tony Lewis

The Bass model is a popular diffusion model that has been extensively tested on American and European time series data, with promising results. This study attempts to extend the generalisability of the model by examining its performance using time series data from high technology New Zealand innovations. The results demonstrate that the Bass model can reproduce the diffusion of New Zealand innovations. However, some limitations on its use are noted, particularly the effects of early fluctuations in the adoption process and consequent problems in forecasting from early data.

Keywords: Bass Model, diffusion curve, adoption process, time series

Introduction

Diffusion theory's special focus is on the processes by which an innovation "is communicated through certain channels over time among members of a social system" (Rogers 1983, p5).

In marketing, the diffusion of innovations occurs with every launch of a new type of product, and is widely thought to be influenced by both inter-personal and mass media communication. Potential adopters are regarded as having different propensities for relying on these two forms of communication when seeking information about an innovation. The effects of interpersonal communication in particular are thought to be a key factor in accounting for the speed and shape of the diffusion of an innovation (Rogers 1983; Gatignon & Robertson 1985; Mahajan, Muller & Bass 1990).

In general, diffusion models assume that the timing of first time purchases of an innovation are distributed in some fashion over the population. These models are expressed in a general category of models of diffusion rate at time t (Sultan, Farley, & Lehmann 1990), that is:

\[
\frac{dN(t)}{dt} = g(t)[N^* - N(t)]
\]

where \( \frac{dN(t)}{dt} \) is the rate of diffusion at time t,
\( N(t) \) is the cumulative number of adopters at time t,
\( N^* \) is the total number of potential adopters in the population, and \( g(t) \) is the probability of adoption for individuals who have not yet adopted

Different diffusion processes are embodied in different specifications of \( g(t) \). For example, the Fourt & Woodlock (1960) model assumes that the diffusion process is influenced solely by external factors, i.e., \( g(t)=p \), where \( p \) is the coefficient of external (mass media) influence. By contrast, Mansfield (1961) suggests a model in which the diffusion process is influenced solely by internal factors, i.e., \( g(t)=qF(t) \), where \( q \) is the coefficient of internal (or interpersonal) influence. The effect of internal influence is a function of the number of previous adopters, which increases with time; hence internal influence is not just \( q \), but is \( qF(t) \). Bass (1969) combined these models in a mixed influence model, the Bass model, where
\[ g(t) = p + qF(t) \]

Bass labelled those who adopt due to external influences innovators, and those who adopt due to internal influences imitators.

Mathematically the model can be expressed as:

\[ P(t) = p(0) + \frac{q}{m} Y(t) \]

where \( P(t) \) is the probability of purchase at time \( t \), given that no purchase has yet been made. The second term, \( p(0) \), represents the probability of initial purchase and reflects the tendency to trial by innovators. The parameter \( q \) represents the coefficient of imitation, and \( m \) the total number of initial purchases. Therefore, \( \frac{q}{m} \) is the proportion of adopters influenced by interpersonal communication, which is magnified by increases in the total number who have already adopted at time \( t \) (i.e., \( Y(t) \)).

Estimating the diffusion curve requires the parameters \( p \), \( q \) and \( m \) to be identified. Bass (1969) developed a method of estimating these parameters using ordinary least squares (OLS) multiple regression (this procedure is set out in Appendix A). The Bass model can then be used to predict the timing and magnitude of the sales peak, and the shape of the diffusion curve (again, see Appendix A for details). Thus, for marketers, the Bass model has significant implications for forecasting, decisions about new product viability, and product launch performance tracking.

Bass (1969) tested this approach on early sales data for eleven consumer durables. He showed that the model had a good fit to the sales curves (presumed to represent first time purchases, or adoptions) for each of the eleven product categories in his study. Further evaluative studies of the Bass model have provided strong empirical support for the structural soundness of the model in a variety of disparate circumstances. Jeuland (1994) compared the performance of the Bass model across 32 data sets, including the 11 durables in the original Bass study and 17 VCR markets in Europe and the United States, and reported a very good fit for OLS estimation model, with \( R^2 \) values of about 0.9. Similarly, in applying the Bass model to cocoa-spraying chemicals among Nigerian farmers over a 25 year period, Akinola (1986) reported an \( R^2 \) of 0.96. Lawton & Lawton (1979) reported an \( R^2 \) of 0.98 for the diffusion of an educational innovation in the US over an 11 year period, and Kalish & Lilien (1986) reported an \( R^2 \) value of 0.89 when testing the Bass model against the diffusion of the photovoltaic home energy systems in South West US over a 10 year period. (These results may be slightly over stated, as some authors do not report adjusted \( R^2 \), despite having often estimated the model on relatively few data points.)

The Bass model has also been successful in forecasting the number of adoptions at the peak of the sales curve using early sales data (see Dodds 1973; Lawton & Lawton 1979; Tigert & Farivar, 1981) and in estimating long term patterns of diffusion (see Akinola, 1986; Mahajan & Peterson 1978). Companies that have used the model include Eastman Kodak, RCA, IBM, Sears and AT&T (Mahajan et al. 1990).

A number of attempts have also been made to extend or improve the application of the Bass model. For example, market potential (\( m \)) has been estimated using exogenous sources of information (see Heeler & Hustad 1980; Souder & Quaddus, 1982; Mesak & Mikhail 1988). Maximum likelihood and non-linear estimation procedures have been used to estimate the model parameters directly from the time series data (Mahajan et al. 1990).
Extensions of the Bass model have also attempted to account for a dynamic ceiling on the number of potential adopters (m), the effect of competition within a market (Mahajan & Peterson 1978), and the influence of a variety of marketing variables on the adoption process (see Mahajan et al. 1990 for a review). Bass has also proposed a generalised version of the model to take account of unusual changes in price or advertising during the adoption process (Bass, Krishnan & Jain 1994). Few of these extensions have been subject to extensive replication, and the original version of the Bass model remains in widespread use.

The evidence of the Bass model’s success in fitting historical data lends credibility to the model’s basic structural soundness and its ability as a forecasting device for marketers of new products (Heeler & Hustad 1980). However, most of the data sets have been obtained from the US or from European countries (Sultan, Farley & Lehmann 1990). Attempts to fit the model to other environments have generally not been entirely successful; when applying the Bass model to a number of products from many international settings Heeler & Hustad (1980) found that the fit of the data to the model was poor in about one third of the cases (eg., air conditioners in Israel, black and white televisions in Algeria and Yugoslavia, and refrigerators in India). After using exogenous sources of information to estimate the parameter m, the prediction of the magnitude of the peak improved in Heeler & Hustad's analysis, but the timing of the sales peak was frequently underestimated in many international settings (air conditioners in Australia, black and white televisions in Yugoslavia, colour televisions in Japan, refrigerators in Hungry and food mixers in Hong Kong).

For this reason, the body of knowledge encompassing the Bass model provides little evidence for determining the extent to which the Bass model is reliable as a forecasting device in New Zealand. Thus the objective of this study was to replicate the original Bass study in an attempt to extend the generalisability of the model to a New Zealand environment. In particular the fit of the OLS estimation model and resulting Bass diffusion curve to New Zealand time series data was investigated at varying levels of aggregation. The ability of the Bass model to forecast the remainder of the diffusion curve using early data was also examined.

**Method**

**The data**

Time series data from four innovations in New Zealand were obtained. Two data sets were for technology based service products and two for telecommunication products. Consequently, this study extends the application of the Bass Model to high technology innovations, as well as to New Zealand.

The Bass model predicts the diffusion of product categories, so the data was obtained from monopolists, enabling company sales to be treated as category sales. The data was also assumed to consist of first time purchases only, which were taken to be equivalent to adoptions. This was justified because the inter-purchase time for the telecommunications products was assumed to be longer than the times series, while for one high technology service product no repeat purchase was possible, and for the other the data specifically identified first time users. Further details of the data follow.
Telecommunication Sector

Time series data were obtained for two telecommunication products: Telecom Mobile Cellular Communication subscriptions, and a product which we refer to as an "unnamed" telecommunication product. Some details on the unnamed product are withheld because of the commercially sensitive nature of this product. The cellular subscription data exhibited multiple peaks, as new generations of cellular phones were released and market potential increased; only the data associated with the first peak is examined here. The unnamed product data exhibited an initial period of stability during a trial, or soft launch, followed by sales take-off after a full commercial launch.

University Sector

A time series of first time users of Univoice was obtained from Massey University. This was an automatic telephone enrolment service, offered in New Zealand for the first time by Massey University, for students intending to enrol or re-enrol for the subsequent academic year. Univoice offered students a simple and convenient way to order enrolment, allowance and loan packs, as well as access to other related registration activities. Advertising for this product relied heavily on print material, with promotional material sent direct to current students and included in enrolment packs.

Supermarket Sector

A time series of adopters of a retail shopping card was obtained from Foodtown, one of the main supermarkets in New Zealand. Shoppers who applied for a Foodtown card became members of a "club" that gave them discounts on selected in-store products. Foodtown cards were not extensively advertised, with only in-store advertising having been employed.

Procedure

The time series data for these products was analysed using the original Bass model (Bass 1969) to test the applicability of the model to New Zealand data. The OLS estimation model (equation (4) in Appendix A) was applied to the data, and the resulting regression parameters were input into a spreadsheet which used equations (3), (5), (6), and (7) in Appendix A, to produce estimates of the Bass models parameters, the timing and magnitude of the peak, and a graph of the diffusion process.

The performance of the model was then measured in four ways:

1. The fit (adjusted R^2) of the OLS estimation model.

2. The plausibility of the Bass model parameters p, q and m. It was expected that p and q would be positive, with p<q, that the parameter m should be approximately equal to the total number of first purchases in each time series, and that pm should approximately equal the beginning of the diffusion curve.

(In fact the plausibility of m could only be directly tested when diffusion was completed within the period covered by the time series data, which only occurred in the Univoice time series. If diffusion was not completed, m should be somewhat greater than the total number of first purchases in the time series data.)
3. The fit (adjusted $R^2$) of the Bass diffusion curve to the actual time series data.

4. The predictive ability of the Bass model when estimated from early data; in particular, the fit (adjusted $R^2$) of the Bass diffusion curve to the time series data, and a comparison of the timing and magnitude of the predicted and actual sales peaks.

Results and Discussion

The results suggested that the Bass model could accurately reproduce the diffusion of New Zealand innovations. However, the model's ability to do so was reduced when fluctuations in the data existed, especially early in the adoption process, and when the data was highly aggregated. The model was also less successful in reproducing the actual diffusion curve when it was estimated from only the first few periods of the time series data.

The goodness of fit of the Model

The model was first applied to the complete set of historical data. The results, shown in Table 1 and Figure 1, appear to support the extension of the Bass model to New Zealand telecommunication and Univoice products, but not to Foodtown cards. (The last two columns in Table 1 are percentage values; a value of 100 indicates perfect agreement between the Bass estimate of $pm$ or $m$, and the time series data.)

Table 1. Overall Model results

<table>
<thead>
<tr>
<th>Product</th>
<th>Aggregation Level</th>
<th>Estimation Adj. $R^2$</th>
<th>$p$</th>
<th>$q$</th>
<th>$m$</th>
<th>Bass Adj. $R^2$</th>
<th>$pm$/1st period (%)</th>
<th>$m$/total sales (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univoice</td>
<td>fortnightly$^1$</td>
<td>.87</td>
<td>.10</td>
<td>.47</td>
<td>5,788</td>
<td>.85</td>
<td>110</td>
<td>100</td>
</tr>
<tr>
<td>Foodtown</td>
<td>monthly</td>
<td>.31</td>
<td>.07</td>
<td>.07</td>
<td>291,701</td>
<td>.41</td>
<td>100</td>
<td>140</td>
</tr>
<tr>
<td>Cellular</td>
<td>quarterly</td>
<td>.99</td>
<td>.00</td>
<td>.47</td>
<td>61,543</td>
<td>.97</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td>Unnamed</td>
<td>quarterly$^2$</td>
<td>.49</td>
<td>.04</td>
<td>.38</td>
<td>38,749</td>
<td>.45</td>
<td>480</td>
<td>110</td>
</tr>
</tbody>
</table>

1. The Xmas and New Year periods were removed by evenly dividing these periods among their adjacent periods.
2. The initial drop in sales was removed to avoid erroneous results. This reflects the standard approach of modelling from sales take-off, rather than from initial availability.

The amount of variance explained (adjusted $R^2$) by the OLS estimation model was high, with the average being 0.67. After removing Foodtown data, on which the model performed poorly, the average variance explained by the estimation model was 0.78 (or, for comparison with other research, 0.82 unadjusted).

The estimation procedure generally yielded plausible parameter estimates, with two exceptions. The Foodtown model gave an implausible relationship between $p$ and $q$, and a considerable overestimate of $m$. The telecommunication products also provided poor estimates of $pm$, but this seems to be due to fluctuations associated with the trial, launch, or improvement of the products in the early stages of the diffusion process.
While the telecommunication products appeared to slightly overestimate $m$, an inspection of Figure 1 shows that this is not a problem. The adoption process had not quite finished by the end of the time series, so we would expect $m$ to be slightly higher than total sales in the time series.

The Bass diffusion curves also appeared to fit the times series data relatively well. This was seen in the adjusted $R^2$ values, which were comparable to those of the OLS estimation model.

![Figure 1. Fit of the Bass Model to the Time Series Data](image)

The close relationship in Figure 1 between the predicted and actual sales peaks lends further credibility to the model’s claim to reproduce the adoption process. Figure 1 also shows that the Bass model reproduced the general shape of the time series. These results, combined with the excellent adjusted $R^2$ values for the Cellular and Univoice time series, suggest that the Bass model can accurately reproduce the adoption of New Zealand innovations.

When applying the Bass model to the Foodtown time series, however, the results were generally poor. The model was applied to weekly, fortnightly, and monthly data (see Table 2)
with very little success. At the weekly and fortnightly level of aggregation, the resulting parameters were even more implausible than at the monthly level of aggregation.

An explanation for the poor performance of the model using the Foodtown time series can be found in the actual pattern of adoptions for this product. By inspecting Figure 1, we can see that the overall pattern of the Foodtown time series was unlike the other products. Instead of exhibiting roughly a bell shaped pattern, we find the first period adoptions were close to the magnitude of the peak, the peak then occurred in the next period, and was followed by a pattern of steady decline in adopters. This suggests that a pure innovation model, like that of Fourt & Woodlock (1960) may be a more accurate representation of this adoption process. There was, however, a great deal of fluctuation in the data, and this fluctuation is likely to be equally problematic for both the Bass model and the Fourt & Woodlock model. It is possible that greater temporal aggregation would reduce the severity of these fluctuations, but this was not possible as too few data points were available to accommodate meaningful quarterly or yearly aggregation, and adoption had almost finished in the time period under analysis.

Level of aggregation and number of periods

The level of aggregation is also capable of affecting the shape of the time series diffusion data. Heeler & Hustad (1980) and Tigert & Farivar (1981) suggested that quarterly aggregation is preferable as it eliminates some of the seasonality in time series, while providing more data points than annual aggregation. Aggregation also reduces the effect of random fluctuations. Despite Tiger & Farivar’s comments, the best level of aggregation will not necessarily be quarterly, but will be the one which produces a smooth and regular diffusion pattern, without reducing the degrees of freedom too far.

Table 2 details some interesting results concerning the fit of the OLS estimation model and the Bass diffusion curve at different levels of aggregation.

First, the levels of aggregation which produced the highest adjusted \( R^2 \) in the OLS estimation model did not necessarily produce the highest adjusted \( R^2 \) for the Bass diffusion curve. Second, over-aggregation clearly occurred for monthly Univoice and annual Cellular time series; the unadjusted estimation \( R^2 \) for these time series were 0.95 and 1.00 respectively, but the adjusted Bass model \( R^2 \) were only 0.11 and 0.10. Previous research has sometimes used the unadjusted \( R^2 \) for the OLS estimation model as a measure of model performance; the results in Table 2 suggest that this is likely to lead to an erroneous assessment of model performance, especially when the data is highly aggregated. A better measure is the adjusted \( R^2 \) of the Bass diffusion curve.

Third, while the adjusted \( R^2 \) of the Bass diffusion curve for the Unnamed telecommunications products was marginally higher at the monthly level, quarterly aggregation produced more plausible parameter estimates. The difference in adjusted \( R^2 \) between monthly and quarterly aggregation was negligible for this product, so the quarterly level of aggregation was treated as "best" due to its greater plausibility. This result emphasises the need to examine the parameter estimates as well as the fit of the Bass model.
### Table 2. Model results with different aggregation levels

<table>
<thead>
<tr>
<th>Product</th>
<th>Aggregation Level</th>
<th>Est. Adj. R^2</th>
<th>p</th>
<th>q</th>
<th>m m</th>
<th>Bass Adj. R^2</th>
<th>pm/1st period (%)</th>
<th>m/total sales (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univoice</td>
<td>weekly</td>
<td>.57</td>
<td>.05</td>
<td>.14</td>
<td>6.06</td>
<td>.45</td>
<td>2170</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>fortnightly</td>
<td>.66</td>
<td>.08</td>
<td>.37</td>
<td>5.905</td>
<td>.64</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>fortnightly^1</td>
<td>.87</td>
<td>.10</td>
<td>.47</td>
<td>5.788</td>
<td>.85</td>
<td>110</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>monthly</td>
<td>.90</td>
<td>.06</td>
<td>1.01</td>
<td>5.725</td>
<td>.11</td>
<td>30</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>weekly</td>
<td>.43</td>
<td>.02</td>
<td>.00</td>
<td>340,738</td>
<td>.28</td>
<td>580</td>
<td>160</td>
</tr>
<tr>
<td>Foodtown</td>
<td>fortnightly</td>
<td>.39</td>
<td>.02</td>
<td>-.01</td>
<td>515,636</td>
<td>.28</td>
<td>170</td>
<td>260</td>
</tr>
<tr>
<td></td>
<td>monthly</td>
<td>.31</td>
<td>.07</td>
<td>.07</td>
<td>291,701</td>
<td>.41</td>
<td>100</td>
<td>140</td>
</tr>
<tr>
<td>Cellular</td>
<td>quarterly</td>
<td>.99</td>
<td>.00</td>
<td>.47</td>
<td>61,543</td>
<td>.97</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>quarterly^2</td>
<td>.99</td>
<td>.01</td>
<td>.47</td>
<td>60,930</td>
<td>.81</td>
<td>140</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>annual</td>
<td>.00</td>
<td>.03</td>
<td>2.58</td>
<td>47,213</td>
<td>.10</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>Unnamed Telecom.</td>
<td>monthly</td>
<td>.60</td>
<td>.00</td>
<td>.14</td>
<td>43,486</td>
<td>.07</td>
<td>00</td>
<td>120</td>
</tr>
<tr>
<td>Product</td>
<td>monthly^2</td>
<td>.44</td>
<td>.01</td>
<td>.13</td>
<td>39,811</td>
<td>.47</td>
<td>1010</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>quarterly</td>
<td>.71</td>
<td>.00</td>
<td>.44</td>
<td>41,458</td>
<td>.38</td>
<td>10</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>quarterly^2</td>
<td>.49</td>
<td>.04</td>
<td>.38</td>
<td>38,749</td>
<td>.45</td>
<td>480</td>
<td>110</td>
</tr>
</tbody>
</table>

1. The Xmas and New Year periods were removed by evenly dividing these periods among their adjacent periods.
2. The initial drop in sales was removed to avoid erroneous results. This reflects the standard approach of modelling from sales take-off, rather than from initial availability.

Fourth, there appears to be no direct generalisation across the data sets about the best level of aggregation. The results were best at the fortnightly level for Univoice, at the monthly level for Foodtown, and the quarterly level for the telecommunications products. However, the best levels of aggregation yielded 11, 12, 15, and 11 time periods, respectively. Thus, the optimum level of aggregation for these data sets was one which yielded about 12 periods over the expected timeframe of the diffusion. (The OLS estimation model effectively had one less period in each case, as period zero was not used in estimation, although it was used in calculating the R^2 between the diffusion curve and the time series data.)

### The effects of early fluctuations

A further observation relates to the starting period of analysis. Tigert & Farivar (1981) suggested that an arbitrary decision be made about the starting period of the diffusion process to remove the period when firms are still in an experimental mode as "the improvements in the new process or product may almost be as important as the new idea itself" (Nevers 1972). Thus, the application of the model to new products which change dramatically in the early periods of the adoption process could produce inaccurate parameter estimates.

There may also be early fluctuations for other reasons. For example, the Cellular product did not undergo a dramatic change, but there was a drop in new subscriptions after the first quarter. As the first period (treated as period zero) is not used in the OLS estimation model, this fluctuation did not actually degrade model performance.

A larger drop (over four quarters) in sales occurred in the initial diffusion of the unnamed Telecommunications product. As with the Cellular data this initial drop in sales was removed.
to see if the accuracy of the model would improve. Table 2 shows that, after removing these periods, the model’s performance did improve at both levels of aggregation. At the monthly level, the inclusion of this early dip in sales also led to an implausible negative diffusion curve.

Bass (1969) suggested that the starting period for analysis should be the first period when actual sales are greater than or equal to \( p_m \). This seems somewhat circular; before an estimate of \( p_m \) can be produced, a starting period must be chosen and the Bass model estimated. The Bass model can then be re-estimated after dropping any initial periods in which sales are less than \( p_m \), but this will result in a different estimate for \( p_m \). It is possible that the initial periods which were dropped will now be greater than the new estimate of \( p_m \). Consequently this seems to be an unsatisfactory method for choosing the starting period for analysis. Instead, it seems more sensible to follow the practice of choosing as the initial period the point from which an obvious sales "take off" occurs.

### The forecasting ability of the Model

The ability of the Bass model to replicate the adoption process lends some credibility to the model’s structural soundness, but is a rather weak test of the model’s forecasting ability. In order to test the model’s ability to accurately forecast the timing and magnitude of the sales peak, the period up to the peak was divided in half and, where feasible, the data from the first half of this pre-peak period was used to attempt to predict the remainder of the diffusion curve.

As it happened, this test of forecasting ability could only be applied to the Telecommunication products (see Table 3). The Bass model requires a minimum of three periods to estimate the diffusion curve and, as the peak for Univoice data was in period 2, it is unsurprising that the model accurately "predicted" peak, although the performance of the model in predicting the rest of the diffusion process is more impressive (with an \( R^2 \) of .47). As already shown, the model does not generally work for Foodtown card subscription data; this is also the case here, with the senseless result being produced of a peak in period -22. The Foodtown model also required at least ten periods of data to avoid producing a negative square root error when estimating \( m \).

### Table 3. Predictions using early data from the best aggregation level

<table>
<thead>
<tr>
<th>Product</th>
<th>No. of Periods</th>
<th>Estimation</th>
<th>Bass Adj.</th>
<th>Timing of the peak (periods)</th>
<th>Magnitude of the peak (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adj. ( R^2 )</td>
<td>( R^2 )</td>
<td>Actual</td>
<td>Bass</td>
</tr>
<tr>
<td>Univoice</td>
<td>3</td>
<td>.00</td>
<td>.47</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Foodtown</td>
<td>10</td>
<td>.07</td>
<td>.40</td>
<td>1</td>
<td>-22</td>
</tr>
<tr>
<td>Cellular</td>
<td>5</td>
<td>.97</td>
<td>.51</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Unnamed</td>
<td></td>
<td>.00</td>
<td>.27</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.66</td>
<td>-.02</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
For the Cellular data, the model only worked with a minimum of five periods; using three or four periods resulted in a negative diffusion curve. With five periods of data the overall shape of the diffusion was captured, but the timing and the magnitude of the sales peak were both underestimated.

For the Unnamed telecommunications product data, three periods of estimation data produced meaningful results (there were no problems with negative diffusion, negative peaks, or a negative square root error when estimating m), but the model seriously underestimated both the timing and the magnitude of the peak. Using four periods resulted in negative values for p and m, while the use of five periods gave a negative square root error in estimating m. When six periods of estimation data were used, the predicted peak was very close to the actual peak (although by this stage the peak was included in the estimation data), but the $R^2$ value was unimpressive.

These results suggest that the Bass model is not reliable for forecasting the peak of the adoptions because of the likelihood of fluctuations in the data early in the diffusion process. On the other hand, although the Bass model may not be successful in predicting the timing and magnitude of the sales peak, it is capable of predicting the shape of the overall diffusion curve.

**Conclusions**

For marketing managers, diffusion models have the potential to provide insights into the introduction and growth phases of new product categories. The possible benefits derived from models based on innovation diffusion theory lie in the capacity of these models to predict the timing and magnitude of the sales peak for an innovation, as well as the general shape of the diffusion curve, with relatively little input data.

This research has shown that in three out of four cases, the Bass model accurately reproduced the diffusion of a New Zealand innovation when the data was appropriately aggregated. It appears to be important to aggregate the data used to a level which yields about 12 periods over the expected timeframe of the diffusion.

Reproducing the diffusion of an innovation provides limited benefit for practitioners; the value of the Bass model lies in its claimed forecasting ability. However, using early data to predict the remainder of the adoption process was not very successful in this research. For three of the four data sets analysed the Bass model could not be used as a forecasting tool. For the remaining data set the Bass model was unable to predict the peak of the diffusion process. To a large extent, these problems appear to be due to fluctuations in the data early in the adoption process. It is, therefore, also important to be able to adjust data in some fashion to allow for such fluctuations. One possible solution would be to rely on managerial judgements about what represents a deviation from "normal" sales.

It is also possible that the forecasting performance of the Bass model in this study could have been improved by different estimation procedures, or by some of the model extensions which have been suggested in the literature. These are areas for further research.

Meanwhile, as Mahajan et al. (1990) suggest, diffusion models such as the Bass model can be used for descriptive and normative purposes. The Bass model can be used to describe the diffusion process, to test specific diffusion based hypotheses, to determine whether sales
targets represent a feasible adoption process, and to make limited forecasts to provide a basis for product marketing decision making.

References


Malcolm Wright is a lecturer at the University of Canberra, Clinton Upritchard was a graduate student in the Department of Marketing at Massey University, and Tony Lewis is an Associate Professor in the Department of Marketing at Massey University.
Appendix A. The Bass Model

The key behavioural and mathematical assumptions in the Bass model, as described by Dodds (1973) and Nevers (1972) are as follows:

1. Over the period of interest there are m initial purchases of the product and there are no repeat purchases.

2. The forces of innovative and imitative behaviour are assumed to operate in the market and exert different effects on the rate of initial purchase. These behavioural forces are represented by p and q respectively in the model. Imitators are influenced in the timing of their adoption by social system pressures. This social force is represented by the Y(T) variable (the number of previous adopters). Innovators are not influenced by the number of previous adopters in the timing of their purchase. The probability of a purchase at T, given that no purchase has yet been made is hypothesised to be:

\[ P(T) = p + \left( \frac{q}{m} \right) Y(T). \]  

3. Assuming sales are comprised entirely of initial purchases,

\[ S(T) = P(T) \left[ m - Y(T) \right] \]

or, using (1),

\[ S(T) = pm + (q - p) Y(T) - \left( \frac{q}{m} \right) Y(T)^2 \]  

Where:

- \( S(T) \) = initial sales (adoptions) at T,
- p = coefficient of innovation, corresponding to the probability of an initial purchase T = 0,
- q = coefficient of imitation
- m = the number of initial purchases (adoptions) of the product over the total period, and
- Y(T) = number of previous buyers at time T.

4. The assumptions of the theory are formulated in terms of a continuous model and a density function of time to initial purchase. The solution in which time is the only variable:

\[ S(T) = \left[ \frac{m (p + q)^2}{p} \right] \times e^{-(p+q)T} / \left[ 1 + \left( \frac{q}{p} \right) e^{-(p+q)T} \right]^2 \]  

5. To estimate the parameters p, q and m from the time series data the following analogue to (2) is employed in multiple regression where ordinary least squares estimates of a, b and c are obtained (this will be referred to as the OLS estimation model):

\[ S(T) = a + b Y(T - 1) + c \left[ Y(T - 1) \right]^2 \]  

\[ T = 2, 3, ... \]
(NB: This last part appears to be a mistake in the literature, due to an assumption that 
T=1 is the start of period 1, when it is in fact the end of period 1. Therefore this 
should be T = 1,2...)

6. The parameters of the basic model (m, p and q) are identified in terms of these 
regression coefficients and are:

\[ q = -mc, \quad p = a/m, \quad m = \frac{-b - \sqrt{b^2 - 4ac}}{2c} \] \hspace{1cm} (5)

7. Thus, to predict sales, the estimates for m, p and q are substituted into (3).

8. The peak value of S(T) and the predicted time of this peak are shown to be:

\[ S(T^*) = m (p + q)^2 \quad (6) \]
\[ T^* = (p + q)^{-1} \ln \left( \frac{q}{p} \right) \quad (7) \]