Advertising and Retail Promotion of Washington Apples: A Structural Latent Variable Approach to Promotion Evaluation

Timothy J. Richards, X.M. Gao, and Paul M. Patterson

ABSTRACT

"Commodity promotion" consists of many activities, each designed to contribute to a consumer's product knowledge or influence tastes. However, both knowledge and tastes are unobservable, or latent, variables influencing demand. This paper specifies a dynamic structural model of fresh fruit demand that treats promotion and other socioeconomic variables as "causal" variables influencing these latent variables. Estimating this state-space model using a Kalman filter approach provides estimates of both the system parameters and a latent variable series. The results show that these latent effects contribute positively to apple and other fruit consumption, while reducing banana consumption.

Key Words: commodity promotion, demand system, DYMIMIC, fresh fruit, Kalman filter, LA/AIDS, latent variables.

Regional or state commodity promotion boards often hold market share in their particular commodity as a primary goal. However, increasing imports, a wider variety of available substitutes, or declining consumption of the commodity itself make targeting a share of a declining market somewhat misguided. Instead, commodity boards are becomingly increasingly interested in strategies to increase demand for their product, preferring to compete for a share of a growing market. This is particularly true of the U.S. apple market. Even though Washington apple growers' share of U.S. fresh apple consumption rose from 22.6% to 62.9% between 1951 and 1994, rising banana and other fresh fruit consumption (Figure 1) has caused promoters of Washington apples to question the viability of continuing a market-share strategy in the future. Although their interest in the past has centered on the factors that determine Washington market share, Washington growers must now turn to the determinants of apple-category share and an explanation for the rise in banana consumption.

Many factors can explain such a change in product-demand. While traditional demand systems have proven capable of estimating changes due to variation in relative prices and expenditure levels, the effects of changing tastes and information are less amenable to
system estimation. Many studies include a variety of socioeconomic factors in order to proxy trends in consumer tastes—principally towards more healthy diets and more convenience in meal preparation (Senauer, Asp, and Kinsey). Similarly, consumer preferences may change with the acquisition of more information about characteristics of the product.

This information can come from a variety of sources. While consumers can actively search for products with traits they desire, firms and trade associations provide the bulk of consumer information through advertising and promotion. Nelson argues that all promotion is informative to a certain extent, while Kotowitz and Mathewson refine this idea in arguing that providing consumers with better information about product attributes is the primary way in which promotion increases demand. Stigler and Becker ascribe a similar role to promotion, although through a different mechanism. As an input to a household production model, Stigler and Becker argue that the informative content of promotion increases demand through improving household productivity. Rejecting the notion that promotion merely provides information, Dixit and Norman argue that promotion is inherently manipulative and so shifts tastes directly. While providing information about a product or product class should expand the demand for both the promoted product and all close substitutes, or have a significant “generic” effect on the demand for the product-class, Dixit and Norman’s approach suggests that promotion has only an allocative, or brand-specific, role. These conceptual arguments, however, consider promotion activities as homogeneous.

Whereas mass advertising is more likely to convey price or quality information, retail promotions (such as coupons, in-store displays, product giveaways) are intended to initiate purchase within the store. Because mass, or direct, advertising contributes to consumers’ stock of knowledge or long-term tastes regarding a particular type of product, its effects are likely to linger beyond the dissemination of the message, while efforts to create impulse purchases may be effective only for a single shopping trip. This suggests that a dynamic model may be able to test for the relative strength of an activity’s generic as opposed to its brand effect. This paper develops a model designed to achieve this objective, where the dynamics arise through consumer tastes or their stock of information, rather than the promotion itself.

Whether promotion is intended to affect the stock of information or consumer tastes, both are unobservable, or latent, variables. The usual method of dealing with this latency, following Pollak and Wales, consists of including promotion expenditure as either a translating or scaling variable in a demand system as a proxy variable (see, for example, Goddard and Cozzarin, among many others). However, this approach can result in a different set of parameter estimates depending upon which proxies are used and—if an instrumental variable method is used—on the set of instruments. To address the problem of non-uniqueness, this paper develops and estimates a structural model for the latent information variable. As explained in more detail below, this approach uses the covariance structure among a set of “indicator” variables and direct relationships with a set of “cause” variables to identify the latent variable value. When several indicators and cause variables are employed, this approach is called the multiple-indicator and multiple-cause model (MIMIC) that has origins in Joreskog and Goldberger, and Goldberger (1972a,b; 1977). Not only do promotion-evaluation studies typically rely on proxies for these latent variables, but their methods of accounting for persistent effects of promotion are usually ad hoc. Cox reviews approaches to incorporating promotion dynamics into empirical models. In a structural latent variable framework, if the latent variable is autoregressive the model becomes a dynamic MIMIC, or DYMIMIC (Engle and Watson; Watson and Engle; Engle, Lilien, and Watson; Gao). As a result, the model is able to describe

---

1 Information, in this context, is defined broadly enough to include product characteristics, price, location, availability, quality, or any other factor that may help consumers arrive at a decision that more accurately reflects their true preferences.
the evolution of the latent variable itself, rather than a proxy.

A DYMIMIC model also provides estimates of the contribution of promotion and other socio-economic variables to the unobservable states of consumer knowledge and tastes and, in turn, the effect of these latent variables on the demand for fresh fruits. Further, by including both consumer advertising and retail promotion as observable “causes” of the latent variables in a demand system framework, this approach is able to differentiate between the effectiveness of each method in expanding the demand for specific fruits. Moreover, factors other than promotion can potentially influence latent demand variables. By including variables that capture consumer trends towards healthy eating or convenience in meal preparation among the set of cause variables, this approach is better able to provide estimates of the true effectiveness of promotion, independent of other factors that may be influencing the demand for fresh fruit.

The primary objective of this paper is thus to determine the effect of direct advertising, retail promotion, and other factors on the apple-share of fresh fruit demand. The first section of the paper provides a general description of the specification of the dynamic, structural latent variable model. The second section explains how promotion is incorporated into this model and offers a specific functional form to be used in the empirical application. More detail on the definition of specific variables and the methods used in estimating the DYMIMIC model is provided in the third section. The fourth section explains the estimation results. Drawing on these results, a final section offers suggestions for future research and draws implications for the promotion of other products.

**Econometric Model of Unobservable Factors Influencing Fruit Demand**

This section provides a general description of the DYMIMIC model and shows how promotion can be incorporated to help explain changes in the structure of demand. In its most general form, this model allows promotion to determine the dynamic evolution of a single-state variable with two possible interpretations—the state of consumer tastes or the state of product knowledge. Both are common in the commodity promotion literature. Although this approach belongs to a more general class of state-space models that are widely used in engineering and physical sciences, their application to problems in agricultural economics is becoming more widespread (Chavas; Tegene).

A state-space model consists of a set of measurement (or indicator) equations and a set of transition (or dynamic cause) equations. Measurement equations specify observable indicator variables as functions of the latent variable and a series of predetermined variables, plus a stochastic error term. Including this error term reflects that fact that the indicators are only imperfect measurements of the latent variable. The transition equations, on the other hand, describe the dynamic process of the latent variable. These equations, more often called *cause equations* in a static model, treat the current value of the latent variable as a function of its own past values, a vector of exogenous cause variables, and a stochastic disturbance term. Cause variables are selected based upon their hypothesized direct relationship to the unobservable variable. Another way of interpreting these two sets of variables is to think of cause variables as determining unobservable taste and information effects, while indicator variables provide the most direct observable evidence of changes in the latent variable. There is one measurement equation for each of the indicator variables, relating values of the indicators to the latent variable, exogenous factors, and a disturbance term. In general notation (Watson and Engle) the measurement equations are given by:

\[ y_t = \alpha_t x_t + \beta_t z_t + e_t. \]

In the empirical example of this paper the transition equations provide a parametric representation of the dynamics of the latent variable, as influenced by a set of cause variables, and a stochastic error term:
\[ x_t = \Phi x_{t-1} + \gamma F_t + G v_t, \]

where:

\[
\begin{bmatrix}
   v_t \\
   e_t
\end{bmatrix} = N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R_t \end{bmatrix} \right),
\]

and \( y \) is a vector of observable “measurements”, \( x \) is a vector of unobservable state variables, \( Z \) is a vector of exogenous determinants of \( y \), \( F \) is a vector of exogenous variables assumed to “cause” \( x, e, \) and \( v \) are random disturbances with covariance matrices given by \( R \) and \( Q \). Because identifying the latent variable value relies on using information contained on covariances between the cause and indicator variables, this parameter is identified. While \( x \) is potentially multi-valued, estimation in this case is considerably more complex, so the case of only one state variable will be considered (Aigner et al.).

There are, however, many cause variables used to explain this latent variable and many indicators used to identify its value. Indicators should represent manifestations of the economic factors that the latent variable is intended to represent. In this example, the first set of indicators consists simply of fitted fresh-fruit demand shares derived from estimates of an Almost Ideal Demand System (AIDS) (Gao and Shonkwiler). Also, the percentage of women in the workforce serves as an indicator of consumer tastes towards more convenient and easy-to-prepare foods. This variable may also represent a growing need for information on product characteristics—the less time spent in meal preparation, the less time people have to learn about the nutritional and preparatory aspects of different foods. Cause variables, on the other hand, can be thought of as explanatory variables of the larger trend or trends that the latent variable is intended to represent. In the fruit-demand model, the set of cause variables consists of the percentage of food expenditure allocated to fast food, the ratio of fat calories to carbohydrate calories available in the U.S. food supply, and the dollar amount of expenditure by the WAC on direct advertising and retail promotion. The first cause variable, “fast food,” is intended to capture consumers’ tastes towards greater convenience in meal preparation, while the ratio of fat to carbohydrate calories represents their taste for either more, or less, healthy diets or perhaps improved information on the health consequences of their diet.\(^2\) Direct advertising by the WAC is intended to represent efforts to provide broad nutritional, price, and availability information on Washington apples to consumers.\(^3\) Retail promotion, on the other hand, provides a measure of the WAC’s efforts to move apples into the retail channel through coupons, retail displays, promotion allowances, joint marketing efforts, or other “push” type strategies. Together, these cause and indicator variables form the DYMIMIC model structure.

Because the DYMIMIC is the most general of this class of structural latent variable models, it subsumes several other state-space specifications (Watson and Engle). For example, defining \( x \) as a vector of regression parameters, and \( \alpha_t \) as a set of exogenous variables produces the time-varying coefficients model described by Chow. When \( \phi = \beta = \gamma = 0 \), equations (1) and (2) describe a standard factor analysis model, but allowing for a non-zero \( \phi \) under these conditions defines a dynamic factor analysis framework. If \( \beta \) and \( \gamma \) are non-zero, but \( \phi = 0 \), the equations define a MIMIC framework used by Robins and West, Engle and Watson, Gao and Shonkwiler, and Brumm. Clearly, if the latent variables are thought to exhibit time-series properties, then the most general form of state-space model is to be used. The following section describes the demand system used to provide the majority of indicator variables in the general DYMIMIC model.

\(^2\) While this ratio fell throughout the 1980s, in recent years the relative amount of fat calories available for consumption has risen sharply (Center for Nutrition Policy and Promotion).

\(^3\) While other apple promotion agencies exist, they comprise only 31% of the market in 1994. Further, neither these agencies nor sellers of fresh oranges (Sunkist) or bananas (Dole, Chiquita, Pacific Fruit) would provide promotion data for the 1951–94 period.
Advertising Dynamics and the DYMIMIC Model

Whether promotion affects tastes directly through a persuasion effect or by reducing transactions costs through an information effect, this paper assumes that its influence on demand is indirect through a latent measure of information and taste. As Chang and Kinnucan suggest, demand studies typically assume the states of taste and knowledge are constant. Relaxing this assumption means that consumer utility can be written as a function of both quantity and the latent state variable:

\[(3) \quad \text{Max } U = U(q, \Xi) \text{ s.t. } p'q = m.\]

where \(q\) is an \(n\times1\) vector of consumption quantities, \(\Xi\) is the latent variable, \(p\) is an \(n\times1\) vector of goods prices, and \(m\) is total expenditure. Solving (4) yields demand equations:

\[(4) \quad q_i = q_i(p, m, \Xi).\]

Estimating (4) requires the specification of a theoretically plausible demand system. For the purposes of this study and many before it, a linear approximate version of the AIDS (LAIDS) of Deaton and Muellbauer has several advantages. Beyond the desirable properties they describe, the fact that their model is derived from an expenditure function of Gorman polar form means that it is affine in utility and, therefore, aggregates consistently across consumers (Green). Moreover, because the LAIDS model falls in the class of flexible functional forms, it allows testing of the theoretical restrictions of demand theory (Chang and Kinnucan). Finally, Blanciforti and Green cite advantages in using the LAIDS model to characterize changes in income and price elasticity over time.\(^4\)

Adapting the LAIDS model to include the effects of taste and knowledge uses a version of the translation method of Pollak and Wales or Rossi. The translation approach involves making the autonomous amount of expenditure a function of some explanatory variable. While this is a plausible way to model the information effects of promotion, the persuasive effects on taste are more likely to influence the slope parameters. However, modeling the effect both ways causes insurmountable estimation problems, so this study employs the simpler translating method to capture both the information and taste effects. Including the latent variable with this method results in a share system whose typical element is:

\[w_i = \tau_i \Xi + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln(m/P),\]

where \(\Xi\) is the latent taste-and-information variable, \(P\) is the Stone’s price index: \(P = \Sigma_i w_i \ln p_i\), and \(w_i\) is the budget share of good \(i\).\(^5\) Green shows that this method preserves all of the theoretical requirements of an empirical demand system. In general terms, translation essentially allows the shifting variables to alter the level of discretionary income. The following section provides specific definitions of each of the cause, indicator, and latent variables.

Data and Methods

This study uses annual data from 1951–1994 on retail-weight consumption of fresh apples (\(w_1\)), bananas (\(w_2\)), fresh oranges (\(w_3\)), other fresh fruit (\(w_4\)), and other food (\(w_5\)). Per-capita consumption of these products is from USDA's Fruit and Tree Nuts from 1970–1994 while data prior to 1970 are provided by USDA staff from archive. Retail price data are taken from the annual report of the Bureau of Labor Statistics’ Consumer Price Index: Monthly Summary. In order to keep the model as parsimonious as possible, the "other fresh fruit" variable is constructed as a composite

\(^4\) As a reviewer notes, Alston and Chalfant present results that favor a Rotterdam over a LAIDS specification in their data. Veeman and Xu, however, find that their Canadian meat demand data reject a Rotterdam in favor of an AIDS specification. Further, Alston and Chalfant show very little difference in demand elasticities between the two functional forms. Similar tests could not be performed here, however, as attempts to estimate a differential Rotterdam model within the DYMIMIC algorithm failed to converge.

\(^5\) Substituting the cause equation into this expression gives a specification very similar to Blanciforti and Green’s habit formation model.
of grape, peach, pear, and strawberry consumption. An average price for this group is constructed using a Stone's price index, which uses expenditure shares as weights. While each of these fruits is individually less than 10% of total fresh fruit expenditure, together they represent an average 24.6% of fresh fruit spending over the sample period. This compares to 15.5% for apples, 48.2% for bananas, and 11.7% for oranges. By including “other food” in the demand system as well, the resulting parameter estimates are interpreted as conditional on total food expenditure. Further, this specification assumes that food spending is weakly separable from all other consumer expenditure, a much weaker assumption than that used by other fruit demand studies (Green, Carman, and McManus). The “other food” variable is constructed by subtracting total fresh fruit expenditure from total food expenditure per capita reported in Food Consumption, Prices and Expenditure from the USDA and dividing the result by a food Consumer Price Index reported by the Bureau of Labor Statistics, after first adjusting the index for the contribution of fresh fruit.

Data for the cause and indicator variables come from a variety of sources. The information on the proportion of food expenditure consumed away from home is from various issues of the Food Marketing Review, the authors of which provide data on the proportion consumed as fast-food. Information on the ratio of fat to carbohydrate calories is provided by researchers at the Center for Nutrition Policy and Promotion in Washington, D.C. It is important to note that these data are in terms of food energy available per capita per day available in the U.S. food supply, and not necessarily calories consumed. Differences between these two values may be due to waste, food that is given away, or placed in storage. Both direct apple advertising and retail promotion are taken from WAC records. Advertising consists of expenditure on television, radio, billboards, magazines, and newspaper copy, but no breakdown of expenditures by media is available. Further, because no reliable cost measure nor expenditure share exists for these media, total expenditures are an imperfect measure of advertising intensity. Similarly, retail promotion expenditures cover a wide variety of activities from trade shows to cents-off coupons, so a reliable cost measure is also difficult to derive. Among the indicator variables, women’s participation rate in the work force is from the Bureau of Labor Statistics, while the fitted fruit-shares are taken from the demand system described above. With these data, estimates of the latent variable model are obtained using an algorithm that is becoming common in estimating state-space models.

In fact, one of the principal advantages in using the state-space specification of the DY-MIMIC model is that parameter estimates are found using standard maximum likelihood methods within the recursive Kalman filter algorithm (Watson). Many recent studies employ this technique in analyzing a variety of economic problems with structural latent variables (Burmeister and Wall; Engle et al.; and Harvey). In agricultural economics, Chavas and Tegene each use a Kalman filter to estimate time-varying regression parameters models, which are, in turn, used to test for structural changes in demand. While they define regression parameters as state variables, the current example treats the stock of information and tastes as a single latent state variable. Because interest lies in the evolution of information over time, and information is a latent variable, the estimation method must be able to provide estimates not only of the system parameters, but also of the latent variable itself. There are several methods for constructing the latent variable series from the structural model. Watson and Engle, Dempster et al., and Chen each describe an iterative estimation and minimization (EM) algorithm. Watson and Engle in particular compare the EM method favorably to the method of scoring. This study adopts a somewhat simpler approach than either in applying Harvey’s “smoothing” algorithm.

From the general state-space model description above, Engle and Watson define two sets of unknowns that must be estimated with the Kalman filter—the vector of parameters, \( \Phi = (\phi, \gamma, \alpha, \beta, Q, R) \), and the latent states, \( x_t \). Estimating both the parameters and the un-
observable variable requires a two-stage approach. In the first stage, initial values for the indicator equation parameters are found using standard regression methods. With these initial values, maximum likelihood estimates of the parameter vector $\Phi$ are found using the Kalman filter. Using the Kalman filter to reestimate the latent variable series based on the maximum likelihood parameters produces estimates that, while linear and unbiased, are not best. To find parameters that are indeed best, Harvey shows that smoothing produces the minimum MSE estimates of the latent variable series. Therefore, the second stage involves using this technique to recover the $x_t$ values. Essentially, smoothing is a backwards recursive method that begins with the Kalman filter estimates of $x_{t\mid T}$ and then proceeds to estimate values of $x_{t\mid t}$ for each observation. Because smoothing produces estimates that are based not just on information up to $t$ but from the whole sample, the resulting MSE must be at least as low as that obtained through filtering.

The log-likelihood function from the first stage estimation is quite simple. Define the innovations from estimating the indicator equations as:

$$\eta_t = y_t - E[y_t | \Psi_t],$$

where $\Psi_t$ contains all information up to time $t$, including the best estimates of $y$ through $(t-1)$ and any new information from the exogenous variables in time $t$ (Engle and Watson). Defining the covariance matrix of $\eta_t$ as $H_t$ produces the log-likelihood function:

$$\sum_t L_t(\Phi) = \sum_t -\frac{1}{2} \sum_t (\log|H_t| + \eta_t^\top H_t^{-1} \eta_t),$$

where $\Phi$ is a vector of parameters. Maximizing this log-likelihood function in GAUSS provides estimates of the DY-MIMIC model parameters, while the smoothing algorithm provides estimates of the $\Xi$ series.

**Results**

Given the greater complexity of a DY-MIMIC model compared to a conventional regression model, the first task is to prove that the DY-MIMIC approach represents an improvement over a more conventional—or even over a static—MIMIC model. Once this is established, the discussion returns to the objectives of the paper, namely evaluating the relative effects of direct advertising, retail promotion, and socioeconomic trends on the demand for fresh fruit. Tests of these hypotheses examine the role of the latent variable in each of the share equations and the parameters of each of the cause variables in the transition equation.

Three tests of the validity of the DY-MIMIC model exist. First, to test the superiority of the DY-MIMIC model over a proxy variable approach the null hypothesis is that the variance term ($G$) in the transition, or cause equation (2), is equal to zero. If the data fail to reject the null hypothesis, then the DY-MIMIC model has not been able to improve on the ability of a standard proxy variable model to explain the variation in fresh fruit demand remaining after all price and expenditure effects have been filtered out. Second, the null hypothesis in the transition equation is that the coefficient on the lagged latent variable value is equal to zero. Failure to reject this hypothesis supports a static MIMIC specification over the DY-MIMIC one. Third, if the coefficients of the latent variable in each of the indicator equations (also known as factor loadings) are equal to zero, then the model adds nothing to a demand model that includes only price and income effects. Table 1 presents the results pertaining to the transition equation and the disturbances from the measurement equations, while the fresh fruit LAIDS estimates, including the latent variable parameters, are in Table 2.

The results in Table 1 show that the DY-MIMIC specification is preferred to both a conventional regression model and a MIMIC approach. In this table the $G$ parameter is significantly different from zero, which suggests that the latent variable explains a significant proportion of the variability in fruit demand that is not already explained by the cause variables, which, in more conventional models, would be used as proxies. Further, the fact that the coefficient on the lagged-latent variable is
Table 1. DYMIMIC Cause Parameter and Indicator Variance Estimates

<table>
<thead>
<tr>
<th>Cause</th>
<th>Estimate</th>
<th>t-ratio</th>
<th>Se(e,)</th>
<th>Estimate</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertising</td>
<td>-0.0759*</td>
<td>-3.4313</td>
<td>e_a</td>
<td>0.0062*</td>
<td>5.9054</td>
</tr>
<tr>
<td>Retail Promotion</td>
<td>-0.1009</td>
<td>-1.1805</td>
<td>e_b</td>
<td>0.0128*</td>
<td>6.3112</td>
</tr>
<tr>
<td>Fast Food</td>
<td>0.0338</td>
<td>1.9592</td>
<td>e_o</td>
<td>0.0060*</td>
<td>6.1713</td>
</tr>
<tr>
<td>Fat: Carbo</td>
<td>3.6030</td>
<td>1.8165</td>
<td>e_of</td>
<td>0.0107*</td>
<td>6.4215</td>
</tr>
<tr>
<td>Q</td>
<td>1.0578*</td>
<td>5.8962</td>
<td>G</td>
<td>2.0552*</td>
<td>5.1283</td>
</tr>
<tr>
<td>( \Xi_{-1} )</td>
<td>0.6086*</td>
<td>2.5471</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A single asterisk indicates significance at a 5% level. The elements of the \( e \) vector reported here are standard errors for the disturbances in the indicator equations, where the disturbance for the “other food” equation is normalized to 1.0 for identification purposes. In the subscripts, a = apple, b = banana, o = orange, of = other fruit, \( \text{ww} \) = percentage of women in workforce. \( Q \) is the variance of the cause equation, and \( \Xi_{-1} \) is the coefficient on the lagged latent variable. The GAUSS Kalman filter algorithm provides no goodness-of-fit measures beyond the system log-likelihood function value.

significantly different from zero suggests a preference for the dynamic specification over the static model. Because this parameter is less than 1.0, the dynamic process of the taste and information variable is stationary, or—in more intuitive terms—shocks to the latent variable disappear over time so have little effect on demand more than a few periods in the future.

Table 2 provides the results necessary to determine whether tastes and information help explain the structure of fresh fruit demand. First note that the \( R^2 \) for four of the five equations is relatively high, indicating that this model provides a good fit to the data. Second, from these results, it is clear that the latent index is significant in some of the share equations. Specifically, this index has a significant and positive effect on apple, other fruit, and other food consumption, while it has a negative effect on banana consumption. While the positive effect of tastes on apple consumption is perhaps surprising, the response of other fruit is not. Consumers can now choose from a variety of fresh fruits available year-round.

Table 2. DYMIMIC LA/AIDS Fresh Fruit Parameter Estimates: 1951–1994

<table>
<thead>
<tr>
<th></th>
<th>Apples</th>
<th>Bananas</th>
<th>Oranges</th>
<th>Other Fruit</th>
<th>Other Food</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{\text{apples}} )</td>
<td>0.067*</td>
<td>-0.014</td>
<td>-0.015*</td>
<td>-0.025*</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(10.921)</td>
<td>(-1.103)</td>
<td>(-2.489)</td>
<td>(-2.376)</td>
<td>(-0.809)</td>
</tr>
<tr>
<td>( P_{\text{bananas}} )</td>
<td>-0.007</td>
<td>0.184*</td>
<td>-0.024**</td>
<td>-0.023</td>
<td>-0.109*</td>
</tr>
<tr>
<td></td>
<td>(-0.536)</td>
<td>(6.740)</td>
<td>(-1.874)</td>
<td>(-1.013)</td>
<td>(-6.092)</td>
</tr>
<tr>
<td>( P_{\text{oranges}} )</td>
<td>-0.014*</td>
<td>-0.001</td>
<td>0.014**</td>
<td>0.016</td>
<td>-0.032*</td>
</tr>
<tr>
<td></td>
<td>(-2.020)</td>
<td>(-0.032)</td>
<td>(1.993)</td>
<td>(1.335)</td>
<td>(-3.435)</td>
</tr>
<tr>
<td>( P_{\text{other fruit}} )</td>
<td>-0.006</td>
<td>-0.014</td>
<td>-0.009</td>
<td>0.049*</td>
<td>0.156*</td>
</tr>
<tr>
<td></td>
<td>(-0.578)</td>
<td>(-0.572)</td>
<td>(-0.824)</td>
<td>(2.342)</td>
<td>(5.131)</td>
</tr>
<tr>
<td>( P_{\text{other food}} )</td>
<td>-0.033</td>
<td>-0.131*</td>
<td>0.018</td>
<td>-0.019</td>
<td>-0.046*</td>
</tr>
<tr>
<td></td>
<td>(-1.515)</td>
<td>(-2.886)</td>
<td>(0.826)</td>
<td>(-0.503)</td>
<td>(-2.706)</td>
</tr>
<tr>
<td>( M )</td>
<td>-0.011</td>
<td>0.058*</td>
<td>0.052*</td>
<td>0.150*</td>
<td>-0.273*</td>
</tr>
<tr>
<td></td>
<td>(-0.771)</td>
<td>(2.008)</td>
<td>(3.825)</td>
<td>(6.222)</td>
<td>(-14.096)</td>
</tr>
<tr>
<td>( \Xi )</td>
<td>0.021**</td>
<td>-0.096*</td>
<td>0.003</td>
<td>0.034**</td>
<td>0.007*</td>
</tr>
<tr>
<td></td>
<td>(1.996)</td>
<td>(-4.298)</td>
<td>(0.276)</td>
<td>(1.809)</td>
<td>(4.784)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.863</td>
<td>0.946</td>
<td>0.766</td>
<td>0.596</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Variable definitions: \( M \) = total expenditure; \( \Xi \) = latent variable; \( P_i \) = retail price index of each product. T-ratios are in parentheses. A single asterisk indicates significance at a 5% level, a double asterisk at 10%. The value of the log-likelihood at its optimal value is 7.1366. The reported \( R^2 \) value is the square of the correlation between observed and predicted share values.
Table 3. DYMIMIC LA/AIDS Fresh Fruit Elasticities: 1951–1994

<table>
<thead>
<tr>
<th></th>
<th>Apples</th>
<th>Bananas</th>
<th>Oranges</th>
<th>Other Fruit</th>
<th>Other Food</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) apples</td>
<td>(-0.2418^*)</td>
<td>(-0.0689)</td>
<td>(-0.2889^*)</td>
<td>(-0.2711^*)</td>
<td>(0.0444^*)</td>
</tr>
<tr>
<td></td>
<td>((-3.6354))</td>
<td>((-1.4782))</td>
<td>((-3.3076))</td>
<td>((-3.7366))</td>
<td>((2.6259))</td>
</tr>
<tr>
<td>( P ) bananas</td>
<td>(-0.0446)</td>
<td>(-0.4015^*)</td>
<td>(-0.5665^*)</td>
<td>(-0.4549^*)</td>
<td>(-0.0797)</td>
</tr>
<tr>
<td></td>
<td>((-0.3201))</td>
<td>((-3.9498))</td>
<td>((-3.1454))</td>
<td>((-3.013))</td>
<td>((-1.9463))</td>
</tr>
<tr>
<td>( P ) oranges</td>
<td>(-0.1433)</td>
<td>(-0.0159)</td>
<td>(-0.8547^*)</td>
<td>(0.0373)</td>
<td>(-0.0306)</td>
</tr>
<tr>
<td></td>
<td>((-1.9402))</td>
<td>((-0.3138))</td>
<td>((-8.6803))</td>
<td>((0.4768))</td>
<td>((-1.4078))</td>
</tr>
<tr>
<td>( P ) other fruit</td>
<td>(-0.0598)</td>
<td>(-0.0807)</td>
<td>(-0.2519)</td>
<td>(-0.8088^*)</td>
<td>(-0.3553^*)</td>
</tr>
<tr>
<td></td>
<td>((-0.4537))</td>
<td>((-0.8917))</td>
<td>(-1.4809)</td>
<td>((-5.6389))</td>
<td>((-5.0956))</td>
</tr>
<tr>
<td>( P ) other food</td>
<td>(-0.3143)</td>
<td>(-0.5533^*)</td>
<td>(-0.0581)</td>
<td>(-0.5709^*)</td>
<td>(-0.0156)</td>
</tr>
<tr>
<td></td>
<td>((-1.3566))</td>
<td>((-3.3054))</td>
<td>((-0.1906))</td>
<td>((-2.2606))</td>
<td>((-0.3950))</td>
</tr>
<tr>
<td>( M )</td>
<td>(0.8808^*)</td>
<td>(1.2071)</td>
<td>(1.7601^*)</td>
<td>(2.0496^*)</td>
<td>(0.3477^*)</td>
</tr>
<tr>
<td></td>
<td>((5.6883))</td>
<td>((11.6998))</td>
<td>((8.8539))</td>
<td>((12.1620))</td>
<td>((7.5046))</td>
</tr>
<tr>
<td>( \Xi )</td>
<td>(0.2351^*)</td>
<td>(-0.3409)</td>
<td>(0.0409)</td>
<td>(0.2370^*)</td>
<td>(0.0169^*)</td>
</tr>
<tr>
<td></td>
<td>((1.9906))</td>
<td>((-4.3018))</td>
<td>((0.2772))</td>
<td>((1.8128))</td>
<td>((4.7333))</td>
</tr>
</tbody>
</table>

Variable definitions: \(M\) = total expenditure; \(\Xi\) = latent variable; \(P_i\) = retail price index of each product. A single asterisk indicates significance at a 5% level. The price elasticities in this table are uncompensated.

...through imports and controlled-atmosphere storage that were only available in-season during the 1950s and 1960s. Because consumers can now buy grapes year-round, due to imports from Chile and other Southern Hemisphere countries, total per-capita consumption of grapes from all sources has risen dramatically as consumers buy grapes habitually as a part of every shopping trip, rather than just as a special treat in the summer (Alston et al.). Further, the introduction of new fruits such as kiwifruit, seedless oranges, and melons has created products that did not exist at the start of the sample period. The negative banana-effect suggests that the observed trend towards banana consumption may be due more to price and income effects than to tastes for convenience or a greater amount of consumer information regarding the health benefits of banana consumption. Indeed, the elasticities reported in Table 3 show that this negative effect on banana consumption is not only statistically significant, but also economically significant, or not trivially small.

Specifically, these elasticities show that a 10% increase in the latent variable results in a 3.4% decrease in banana demand. Based on the index values shown in Figure 2, such an increase in the latent variable can occur in a relatively short time—from 1987 to 1992, for example, the index increased by almost exactly 10%. On the other hand, the elasticities of fresh apple and other fruit demand with respect to the latent index is nearly the same strength, but in the opposite direction. If changing tastes and information cannot explain the observed increase in demand, then price and expenditure effects must. The elasticities in Table 3 show that bananas are both more price elastic and income elastic than apples. Because the relative price of bananas to apples dropped by 1.4% over the sample period, this difference can perhaps explain some of the observed change in banana demand. However, other fruit, as a group, is more price elastic than bananas. The fact that banana prices dropped only 0.6% relative to the composite price of other fruits does little to explain the relative change in demand between these two products. Of more importance, however,
Table 4. DYMIMIC LA/AIDS Compensated Fresh Fruit Price Elasticities: 1951–1994

<table>
<thead>
<tr>
<th>Product</th>
<th>Apples</th>
<th>Bananas</th>
<th>Oranges</th>
<th>Other Fruit</th>
<th>Other Food</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{\text{apples}} )</td>
<td>-0.1632*</td>
<td>0.2021*</td>
<td>-0.0834*</td>
<td>0.0673</td>
<td>0.0546</td>
</tr>
<tr>
<td>( P_{\text{bananas}} )</td>
<td>0.0392</td>
<td>-0.0635</td>
<td>0.0678*</td>
<td>0.0921</td>
<td>-0.0484</td>
</tr>
<tr>
<td>( P_{\text{oranges}} )</td>
<td>-0.1314*</td>
<td>-0.0732</td>
<td>-0.7341*</td>
<td>0.0021</td>
<td>0.6798*</td>
</tr>
<tr>
<td>( P_{\text{other fruit}} )</td>
<td>0.0591</td>
<td>0.3857*</td>
<td>0.0712*</td>
<td>-0.7802*</td>
<td>0.2748*</td>
</tr>
<tr>
<td>( P_{\text{other food}} )</td>
<td>0.0766*</td>
<td>0.0183</td>
<td>-0.0489</td>
<td>0.0347</td>
<td>-0.0209*</td>
</tr>
</tbody>
</table>

See Table 3 for variable definitions. A single asterisk indicates significance at a 5% level. The compensated price elasticities are calculated at variable means using the formula:

\[
\frac{\Delta q_i}{q_i} = \frac{\Delta p_i}{p_i} + \frac{y_{ij}}{w_{ij}} + w_{ij}^j
\]

where \( i = j \), \( \tau_0 = 0 \) if \( i \neq j \).

is the expenditure elasticity. Total expenditure on fresh fruit increased by exactly 400% over the sample period. Other fruits, with an expenditure elasticity of 2.05, clearly benefit from increased fruit spending. Moreover, the expenditure elasticity of bananas is 50% higher than for apples, so this fact may be responsible for some of the observed change in demand. Notice also that virtually all pairs of products appear to be gross complements.

Although such results are often disregarded as evidence of a misbehaved demand system, their recurrence in the fruit demand literature suggests that they are more likely a reflection of actual consumer behavior (Lee, Brown, and Seale; Lee, Seale, and Jier wiriyapant). Gross complementarity will occur if a positive income (expenditure) effect resulting from a price reduction of one product outweighs the negative substitution effect. In other words, if consumers have a relatively fixed amount that they allocate to fresh fruit, then price reductions in one product will allow them to buy more of another product.7 The result is a negative cross-price elasticity. This effect is strongest in the case of banana consumption as a 1% decrease in the price of other food causes a 0.55% increase in the demand for bananas. Investigating whether or not these gross complements are indeed net complements as well may help to explain these seemingly counterintuitive results.

Table 4 shows the compensated, or Hicksian, price elasticities. As expected, many of the pairs found to be gross complements do not remain net complements. In fact, only apples and oranges remain statistically significant net complements. More importantly, both apples and other fruit become net substitutes for apples. Because the relative price of bananas falls with respect to both of these alternative products, the reasons for the observed changes in banana consumption (Figure 1) become more clear. Ultimately, however, the structure of demand among these fruits appears to be most sensitive to changes in the unobservable factors accounted for by the latent variable, so changes in its value over time are of considerable interest.

Figure 2 shows how the latent variable index changes over the sample period. Being able to place numerical values on changes in unobservable demand factors is unique to the DYMIMIC approach. While most studies of structural changes in demand are only able to estimate a single parameter or set of parameters, this model allows for a quantification of those effects that other studies only provide.
indirect evidence of. As an index, however, only changes in the latent variable value are meaningful. Normalizing the index at 100 in 1971, the index has increased by 26% over the sample period for an average annual gain of 0.59. If all changes in fresh fruit demand could be explained by variation in prices and expenditure, then this index would be constant. Therefore, finding such a sustained rise in the index suggests that there are other factors in the broader economy that are affecting the demand for fruit. Here we argue that these other factors are a composite of changing tastes and consumer information. The extent to which the index is affected by each of these factors is given by the parameters of the cause, or transition, equation.

These parameters are shown in Table 1. Surprisingly, both types of promotion have a negative effect on the latent variable, although only direct advertising is significant at conventional levels. With promotion program evaluation mandated by the Food and Agricultural Improvement and Reform Act in 1996, a number of studies now purport to demonstrate the effectiveness of generic fruit promotion (Erikson et al., Alston et al.). However, the results of this paper suggest that accounting for other factors that explain changes in demand may reduce the apparent effect of promotion. Given that the latent variable increases monotonically over the sample period, as have promotion budgets, it may be that such positive promotion effects are the result of spurious correlation. To the extent that promotion efforts are intended to change tastes and to add to consumer information, however, the fact that past values of the latent index are significant suggests that this variable has more of a generic than a product-specific effect us-
ing the logic developed in the introduction to
the paper. Consequently, the significance of di-
ext advertising supports the contention that it
has more of a persistent effect on demand,
whereas retail promotion is used more as a
"tactical" marketing tool with more short-
term effects.

Because the latent variable rises over the
sample period, the insignificance of promotion
suggests that the rise is largely determined by
the trends underlying changes in fast food
consumption and the aggregate fat:carbohy-
drate ratio. In the former case, a rise of 1% in
the proportion of food consumed as fast food
causes the index value to rise by 0.034, an
effect which is likely due to an increase in the
demand for convenience in meal consumption
and preparation. In terms of the parameters of
Table 1, the fat:carbohydrate coefficient sug-
gests that a similar rise in the percentage of
fats to carbohydrates in the diet causes the la-
tent variable to rise by over 3.6 points. While
this variable likely captures the trend towards
and then away from low-fat diets over the
sample period, it may also reflect the quality
of nutritional information available to con-
sumers. Whereas health officials argued for
strict lowfat diets during the 1980s, more re-
cent recommendations have advocated mod-
eration and dietary balance.

Conclusions and Implications

Dramatic changes have occurred in the struc-
ture of fresh fruit demand since the early
1950s. Whereas banana and specialty fruit
consumption have risen sharply, apple con-
sumption has been largely flat, and fresh or-
ge consumption has fallen markedly. These
changes can be due to either price and income
effects, socio-economic trends, or active pro-
motion by fruit marketers. Promotion itself
can also have different effects depending upon
whether it is oriented towards the consumer,
or to the distribution channel. This study con-
siders the role of each in a Dynamic Multiple
Indicator-Multiple Cause (DYMIMIC) model
of fresh fruit demand.

The DYMIMIC framework is the most
general form of structural latent variable mod-
el. As such, it defines a latent quantity mea-
suring the stock of knowledge or the state of
tastes. Indicators of this latent variable consist
of the proportion of women in the work force
and fresh fruit budget shares derived from a
LAIDS model. As a state-space model, equa-
tions relating these indicators to the latent var-
iable form the measurement equations, while
cause equations linking exogenous factors to
the latent variable constitute the transition
equations. This model is estimated using a
Kalman filter and smoothing algorithm. In-
cluding an aggregate fat:carbohydrate con-
sumption ratio, the percentage of food con-
sumed as fast food, and direct advertising and
retail promotion of Washington apples in the
set of cause variables allows the model to test
for the contribution of promotion expenditures
to changing tastes—thus changing demand—
for different fresh fruits.

The fresh fruit DYMIMIC model, includ-
ing apples, bananas, oranges, other fresh fruit,
and other food, is estimated with annual data
from 1951–1994. The results show that the la-
tent variable has a significant effect on the de-
mand for apples, bananas, other fruit, and oth-
er food. The latent index has a positive effect
on apple, other fruit, and other food demand,
while reducing the demand for bananas. In the
cause equation estimates, both fast food and
the fat:carbohydrate ratio have a positive ef-
fect on the latent variable value, while direct
advertising has a significant negative effect.
Retail promotion, on the other hand, has no
statistically significant effect. This suggests
that taste changes dominate the latent infor-
mation effect of advertising, and that advertis-
ing, in turn, has a stronger effect on consum-
ers' buying habits than does retail promotion.
The significance of the lagged latent variable
in this equation, however, implies that how-
ever consumer preferences are formed, they
tend to persist.

Future research in this area can extend
along many lines. First, a data set with more
observations would allow the definition of
more indicator variables, perhaps to more di-
rectly capture the influence of promotion on
consumers' stock of knowledge. Second, Aig-
ger et al. suggest that the use of multiple latent
variables will improve the performance of the model as a whole. Clearly, treating information and tastes as separate-state variables is a step in this direction. Third, specifying a two-stage demand system would allow the estimation of both product-specific and fruit-category price, expenditure, and latent variable elasticities. A fourth area for improvement concerns the nature of the USDA consumption data. Adjusting the published “per-capita consumption” figures, which are in fact per-capita utilization, for net exports and storage will more closely approximate the true amount that is consumed by U.S. households. Other variables may serve as better indicators of the value of leisure time in the household production model. Women’s wage rates, overtime hours, and the amount of time spent in leisure activity are but a few of the alternatives deserving consideration.

References


Goddard, E. W. and B. Cozzarin. “A Preliminary Look at Advertising Beef, Pork, Chicken, Tur-


