Retail Demand for Greenhouse Tomatoes: Differentiated Fresh Produce

Gary D. Thompson
Dept. of Ag. & Resource Economics
PO Box 210023
University of Arizona
Tucson, AZ 85721-0023
Email: garyt@ag.arizona.edu

Paper prepared for presentation at the
American Agricultural Economics Association Annual Meeting,
Montreal, Canada, July 27-30, 2003

Retail Demand for Greenhouse Tomatoes: Differentiated Fresh Produce

The growing availability of a variety of fresh tomatoes at retail in the United States reflects significant product differentiation in fresh food items. Fresh tomatoes are differentiated in numerous ways: by size, shape, color, packaging, production method, country of origin, and to a limited extent branding. Not all tomatoes are red; yellow, orange, and more exotic colors such as “zebra” tomatoes are becoming more widely available. Among types of red tomatoes, beefsteak, roma, cherry, and on-the-vine are some of the more readily recognizable types. Some of these types are often available in multiple sizes at any given time. Production methods further differentiate quality because production conditions can be controlled more closely in greenhouses than in fields, resulting in more uniform quality. Packaging also differentiates tomatoes with products available in plastic containers, net and plastic bags, and flats with individual cups for protecting tomatoes in transit. In addition to domestically produced tomatoes, imports are often available from Canada, Mexico, Holland, Belgium, Spain, and Israel. Although few brands of tomatoes are recognized by shoppers, many tomatoes are marketed with price lookup (PLU) stickers sporting brand names, and packaged tomatoes typically have brands prominently displayed.

Understanding retail demand of differentiated food products is important because substantial investments in research and development, production capacity, and marketing are necessary for making such products available at retail. Seed companies throughout the world expend substantial sums to develop varieties with traits which may or may not be accepted by final consumers. Production of highly perishable products such as tomatoes requires significant investments in procuring land, labor, and other inputs in the appropriate microclimates throughout the year, potentially in various countries, to assure availability of high quality field-grown products year round. Establishment of greenhouse production facilities requires sizeable, irreversible investments which may take years to amortize. Marketing of fresh tomatoes also requires commitment of capital: field-grown tomatoes pass through intermediaries known as repackers who typically are located near large retail markets. Repackers sort field-grown tomatoes, which do not mature uniformly, and, in turn, sell directly to supermarket chains and others. Greenhouse tomatoes, by contrast, are not repacked so that greenhouse firms must maintain sales staff who deal directly with supermarket chains and others. Without an understanding of the nature of retail demand for differentiated types of fresh tomatoes, these sizable investments in product development, production, and marketing may not yield future profits.

Trade publications have documented the growing market shares during the late 1990's of greenhouse and, more recently, on-the-vine tomatoes. Yet little is known about the magnitude of this growth because data on retail sales of such products are very limited. The purpose of this study is to describe and analyze the nature of retail demand in the United States for various types of fresh tomatoes. In particular, the growth of greenhouse vis-à-vis field-grown tomatoes is examined. Retail data from major supermarket chains in six metropolitan are used to track market shares, dollar sales, and prices of various types of tomatoes from 1997 through 1999. An econometric model is also employed to estimate own- and cross-price elasticities for five types of tomatoes: regular, greenhouse, on-the-vine, roma, and cherry tomatoes. The implications of the elasticities for future growth in consumer demand for various types of tomatoes are developed.

The Nature of the Retail Data

Retail data for this study were purchased from Fresh Look LLC. Fresh Look assembles sales data from major supermarket chains. Data are collected in supermarkets from two sources: price lookup (PLU) codes and universal product codes (UPC). UPC's are contained on a product's bar code which is scanned at checkout. In contrast, PLU's are numerical codes which checkout personnel must enter on a keypad. Scanned UPC's generally contain very few errors whereas human error in entering PLU's may occur more frequently. An added complication with PLU's is that some produce items are random weight items, meaning that the unit price does not reflect a common weight or volume base. A typical random-weight item in produce sections is head lettuce; prices are usually expressed per head of lettuce, not per unit weight. Nearly all tomatoes possessing PLU's are not random-weight items. In the Fresh Look data employed in this study, all tomatoes had a price per unit weight.

Data for major chains in each of six metropolitan areas were purchased. The metropolitan areas are Albany, NY, Atlanta, GA, Chicago, IL, Dallas, TX, Los Angeles, CA, and Miami, FL. Collectively, these six metropolitan areas account almost 14% of the U.S. population (see Table 1). They represent geographically dispersed markets, some with rapidly growing populations—Atlanta and Dallas—and others with population growth rates less than the national average—Albany and Chicago. Whether measured in terms of households or market size for supermarkets, these six areas also represent just under 14% of the national market.

For reasons of anonymity, the supermarket chains reporting data to Fresh Look cannot be divulged. In general, the aforementioned metropolitan areas were chosen because there were at least two supermarket chains in each area reporting data since the beginning of 1997 . Many supermarket chains have been reticent to collect PLU data because of the problems associated with human error and random weights. In the six metropolitan areas chosen, there were generally enough chains collecting data since 1997 to represent half to two thirds of supermarket sales in the respective areas.[1] As noted in table 1, the number of chains participating range from a low of two in Albany to a high of five in Dallas.

Fresh Look data contain the following information by week for each chain: product identifiers such as PLU or UPC accompanied by brief descriptive text; total dollar sales; total volume sales in pounds; unit prices; and the percentage of stores within the chain in which the product was sold. It is worth mentioning that unit price multiplied times volume sales exactly equal dollar sales. No documentation of advertised sale prices, sales made with coupons, or sales at discounted prices owing to use of frequent purchaser cards is available. Nor is information available on in-store promotion activity of any sort. It is possible some portion of the tomato sales in the data occurred at prices other than those reported in the data; however, no information exists to identify those sales.

The variety of types of tomatoes reported included field-grown vine ripe and mature green, greenhouse, on the vine, roma, cherry, teardrop, grape, yellow, orange, zebra, tomatillo, dried, and organic tomatoes. Many types were further differentiated by size as well as packaging and brands. In order to make the number of tomato categories manageable, six generic categories were chosen: regular, greenhouse, on the vine,[2] roma, cherry, and organic tomatoes. Tomatillos and dried tomatoes were excluded as were all colors of tomatoes other than red. Further, tomatoes of all sizes were aggregated into a single category. For example, PLU's 4798 and 4799, small and large greenhouse or hydroponic tomatoes, were aggregated. Unit prices

---

[1] Exact sales figures for the participating chains versus sales for all supermarket chains in each metropolitan area are not available.

were calculated as weighted averages of the component unit prices while dollar and volume sales were summed. For purposes of this study, tomato data were further aggregated across supermarket chains within each metropolitan area to obtain time series observations for 156 weeks from 1997 through 1999 for five tomato categories in six metropolitan areas.

Descriptive Statistics

Average weekly prices, volume sales, and market shares for each metropolitan market are displayed in table 2. Several tendencies in retail prices are evident. Greenhouse, on-the-vine, and organic tomatoes tend to the most expensive of all types of tomatoes. The sole exception is Atlanta where greenhouse tomatoes had a lower average prices than all other types of tomatoes except roma tomatoes. Conversely, regular and roma tomatoes tend to be cheapest types of tomatoes.

In statistical terms, average prices for different types of tomatoes tend to be different from each other: of the 90 pairs of prices compared, only 10 were statistically indistinguishable from each other (Table 3). In most cases, regular tomato prices tend to be lower than those of most other types of tomatoes; however, average prices for regular and cherry tomatoes in Atlanta and Chicago are not statistically different from one another. Greenhouse tomato prices tend to be higher than those of many other types but are not statistically higher than on-the-vine prices in Atlanta or Chicago. In Dallas, average weekly prices for greenhouse, on-the-vine, and organic tomatoes are all statistically the same. The differences observed in average retail prices suggests that some of the patterns observed in market shares may be a result of relative price differences. For example, relatively higher prices for greenhouse, on-the-vine, and organic tomatoes may be partly responsible for relatively smaller market shares.

---

[2] All on-the-vine tomatoes are produced in greenhouses. Yet, because each tomato in a cluster is still attached to the vine, this type of tomato is sold separately. In some cases, it is not apparent to the shopper on-the-vine tomatoes are

The sizes of the six metropolitan markets can be judged by looking at the average weekly volume sales in table 2. Average weekly tomato sales for the supermarkets in Los Angeles are more than double those for the supermarkets in Chicago, and more than triple those for the supermarkets in Atlanta, Dallas, and Miami. Albany has average weekly sales of less than 1 million pounds. However, in per capita terms, Albany displays the highest consumption. The wide range in per capita consumption figures occurs for numerous reasons. First, the share of supermarket sales represented by those supermarkets included in the Fresh Look data varies from area to area. The two supermarkets included in for Albany may account for a large portion of supermarket sales in Albany. By contrast, the three supermarkets included for Atlanta seem to represent a smaller portion of total supermarket sales in Atlanta. Another source of the variation in per capita sales figures likely occurs because population figures are collected on a county basis. The density of stores in counties likely differs so that some chains' sales are likely concentrated in a smaller part of a given metropolitan area. Given these provisos about the differences in per capita consumption, for the sample data, it is clear Los Angeles represents a relatively large market and even Chicago accounts for a large share of tomato sales.

Market shares of different types of tomatoes vary significantly across metropolitan areas. Los Angeles clearly has the most diversified share of tomato types: regular tomatoes only accounted for one third of all sales while roma tomatoes have one quarter of sales and greenhouse tomatoes over 20 percent. Atlanta and Miami, both cities in the South, still have over half their market shares composed of regular tomatoes. Greenhouse tomatoes have made the least inroads in these two Southern cities, with share of less than 10 percent. Perhaps surprisingly, on-the-vine tomatoes have relatively high markets shares in Atlanta and Miami. Roma tomatoes have the highest shares, more than one quarter of sales, in Dallas and Los

---

grown in greenhouses.

Angeles, both cities with nearly 30 percent of their populations identified as Mexican(U.S Census Bureau, QT-P3).[3] In all metropolitan areas, cherry tomatoes appear to be a “residual” category with never more than a 7 percent market share. Organic tomatoes were clearly a niche item with less than 1% market share in all areas except Dallas.

The time series data give some appreciation for how prices, volume sold, and market shares have changed in the last three years of the 1990’s. From table 4, it is clear that in all areas but Atlanta, the declining market share of regular tomatoes was accompanied by a decline in volume sold at retail. Put differently, growth in the volume of tomato sales was not sufficient in these markets to offset the decline in market share of regular tomatoes. These declines in share and volume of regular tomatoes mean that growers, shippers and repackers of regular tomatoes in these five metropolitan areas have likely experienced significant reductions in revenues associated with regular tomatoes.

In most cases, when the market share of regular tomatoes has declined, the decline has occurred while nominal prices have increased, though Miami is an exception. The nominal price increases are not always large compared to the rate of inflation for foods and beverages nationally.[4] For most types of tomatoes, average prices either declined or increased at less than the rate of inflation. A notable exception to this pattern are tomato prices in Chicago, which increased for all types of tomatoes except greenhouse tomatoes.

Greenhouse and on-the-vine tomatoes have garnered additional market share, in many cases more than 5 percent. And even when growth in market share of greenhouse and on-the-

[3] The percentage of Mexican residents in each metropolitan area ranged from a low of 0.4 in Albany to a high of 29.5 in both Dallas and Los Angeles. The simple correlation coefficient between percentage of Mexican population and percent market share of roma tomatoes was 0.756. The positive correlation between the two provides auxiliary evidence that roma tomato sales are apparently higher the larger is the population of Mexican residents.

[4] The U.S. city average consumer price indices for food and beverages were 157.7 and 164.6 in 1997 and 1999, respectively (U.S. Department of Labor). This means the average inflation rate for food products over the sample period was 6.9%.

vine tomatoes has been modest, volume sold has doubled or more. In every metropolitan area, average prices of greenhouse tomatoes fell, from a low of –1% in Albany to a high of –18% in Atlanta. These formidable increases in greenhouse and on-the-vine market shares suggest many consumers have either switched permanently to purchasing greenhouse and on-the-vine tomatoes or at least prefer to buy them when their prices are closer to those of regular tomatoes.

The last three types of tomatoes—roma, cherry, and organic—showed mixed performance at retail. In general, roma and cherry tomatoes have exhibited little change in market share. The sole exception seems to be Miami where the share of cherry tomatoes jumped by 9 percent. Shares of organic tomatoes have remained static or shrunk appreciably. The volumes sold are so small compared to other types of tomatoes, that even a 93 percent increase in volume sold in Los Angeles, for example, represents a miniscule increase in market share. The only market to register appreciable growth is Dallas where market shares of organic tomato grew by just under 2 percent, despite more than a ten fold increase in sales volume.

The foregoing descriptive statistics suggest greenhouse and on-the-vine tomatoes have gained considerable market share, at least in the six metropolitan markets considered. Their gain in market share has come largely at the expense of falling market shares and sales volumes for regular tomatoes. While declining prices of greenhouse tomatoes have apparently stimulated this growth in market share, prices of on-the-vine have generally not decreased and yet the their market shares have grown. Roma tomatoes have relatively large market shares in Dallas (26%) and Los Angeles (31%) but those shares have grown only about 1.5 percent during the sample period. With few exceptions, market shares of cherry and organic tomatoes are small and have not grown appreciably. Thus the majority of the dynamism in product differentiation in the fresh tomato market appears to have occurred because of innovations in greenhouse production and distribution. While greenhouse and on-the-vine tomatoes have enjoyed growing consumer acceptance at retail, field-grown tomatoes have apparently lost significant market share, in many cases despite reductions in retail prices. Given these dynamic changes in the retail tomato market, it is important to understand how responsive retail demand has been to changes relative prices and expenditures. To that end, an econometric model of consumer demand is introduced.

Econometric Model

The quadratic almost ideal demand model (QAIDS) was chosen for estimating demand elasticties because of its flexibility (Banks, Blundell, and Lewbel). It is a rank three demand model capable of generating nonlinear Engle curves. QAIDS nests the more widely known almost ideal demand model (AIDS) of Deaton and Muellbauer in a parsimonious way with the addition of a single additional parameter per good in the demand system. Expenditure shares on the i^{th} good in the QAIDS model are given by:

$$(1) \qquad w_{it} = \alpha_i + \sum_j \gamma_{ij} \ln p_{jt} + \beta_i \ln\left[\frac{x_t}{a(p_t)}\right] + \frac{\delta_i}{b(p_t)}\left\{\ln\left[\frac{x_t}{a(p_t)}\right]\right\}^2 + \varepsilon_{it} \quad i = 1,\ldots,m; t = 1,\ldots,T$$

Unit price and quantity of the i^{th} good are denoted as p_{it} and q_{it}, respectively, in time period *t*. Expenditure on all *m* goods in the system is $x_t \equiv \sum_{i=1}^{m} p_{it} q_{it}$, which yields the expenditure share of the i^{th} good denoted as $w_{it} \equiv \frac{p_{it} q_{it}}{x_t}$. The terms in the denominators of (1) are defined as follows: $\ln a(p_t) = \alpha_0 + \sum_j \gamma_{ij} \ln p_{jt} + 1/2 \sum_i \sum_j \gamma_{ij} \ln p_{it} \ln p_{jt}$, and $b(p_t) = \prod_j p_{jt}^{\beta_j}$, the translog and Cobb-Douglas price aggregators, respectively. The stochastic disturbance in each share equations is denoted ε_{it}.

The system of goods in (1) requires nonlinear estimation of $m(m+3) + 1$ parameters $\alpha_i, \beta_i, \gamma_{ij}$, and δ_i when no restrictions from consumer theory are imposed. The parameter α_0 customarily is set at a predetermined level. The value chosen for α_0 was slightly less than the minimum sample value of the natural logarithm of expenditure (Banks, Blundell, and Lewbel, p. 534). Imposition of restrictions from consumer theory reduces the number of parameters to be estimated. Adding up restrictions, given by $\sum_i \alpha_i = 1,\ \sum_i \gamma_{ij} = \sum_i \beta_i = \sum_i \delta_i = 0$, restrict $(m+3)$ of the parameters. Homogeneity of degree zero in prices, $\sum_j \gamma_{ij} = 0\ \forall i$, provides m additional restrictions. And symmetry, $\gamma_{ij} = \gamma_{ji}\ \forall i, j$, results in $m(m-1)/2$ restrictions. The QAIDS model nests the AIDS model if $\delta_i = 0\ \forall i$. These additional m restrictions may be applied if statistical tests cannot reject the null of $\delta_i = 0\ \forall i$.

The model in (1) was estimated using per capita quantities. Details for how weekly population figures were generated for each metropolitan area are given in Appendix A. No demographic variables were included in (1) because with weekly data spanning three years, the variation in demographic variables would have been nearly imperceptible.

An individual QAIDS model for each of the six metropolitan areas was specified to reflect the decisions consumers in each of the areas must make. Further, weak separability of fresh tomatoes from other food, non-durable, and durable goods was assumed so that estimation could focus on the demand system for fresh tomatoes in each metropolitan area. Finally, organic tomatoes were excluded from the demand systems because there were significant numbers of weeks in some areas in which organic tomatoes simply were not available. With no observations on price or quantity of organic tomatoes in those particular weeks, the only option would have been to reduce the sample size by omitting observations for other types of tomatoes in those same weeks.[5] Given the foregoing, six demand systems were estimated—one for each metropolitan area—for five types of tomatoes: regular, greenhouse, on-the-vine, roma, and cherry tomatoes.

Estimation and Inference

Because of the large number of systems to estimate in each of the metropolitan areas, adding-up, symmetry, and homogeneity were imposed in all cases. Hypothesis testing focused on three issues: the possible presence of serial correlation with high frequency time series data; potential endogeneity of group expenditure; and nesting of the AIDS within the QAIDS model.

Tests for serial correlation were performed because the models were estimated in the levels and weekly data were used. The QAIDS model is nonlinear in the parameters, thereby making tests statistics based on linear models inappropriate. As a diagnostic tool, White's approximate Durbin-Watson statistic for single-equation, nonlinear models was used equation by equation to test for presence of autocorrelation.[6] For demand systems estimated in share form, the restrictions on autoregressive process are well known (Berndt and Savin). In the interest of parsimony, a single autoregressive parameter for all equations was estimated with up to four lags considered. Four lags would indicate a lag of one month, which is just beyond the maximum shelf life of extra-long shelf life varieties of fresh tomatoes. The particular parametric form for the autoregressive process considered was:

$$\mathbf{u}_t = \sum_{j=1}^{4} \mathbf{R}_j \mathbf{u}_{t-j} + \boldsymbol{\varepsilon}_t \quad t = 1, \ldots, T \tag{2}$$

[5] Limited dependent variable approaches to handling zero expenditure shares could not be employed because not only were zero quantities observed but also no observations on price exist when organic tomatoes were not sold in those weeks.

[6] Although White's test uses parameter estimates from single-equation nonlinear least squares, for diagnostic purposes, the parameter vector from the system estimates was use for each equation.

where $\mathbf{R}_j$ is an (m x m) matrix of autocorrelation coefficients associated with the j$^{\text{th}}$ lag, $\mathbf{u}_t' = [u_{1t} \quad u_{2t} \quad \cdots \quad u_{mt}]$, and $\boldsymbol{\varepsilon}_t' = [\varepsilon_{1t} \quad \varepsilon_{2t} \quad \cdots \quad \varepsilon_{mt}]$.[7] However, only diagonal $\mathbf{R}_j$ were considered in which all m autoregression coefficients were considered equal. With up to four-period lags considered, this meant only a modest number of additional parameters had to be estimated in an already highly nonlinear model. More elaborate specifications of the autocorrelation parameters could have been used (Moschini and Moro) but with the possibility of up to four period lags, such specifications would require estimating more parameters. A multivariate version of Godfrey's locally equivalent alternative test for nonlinear models (Godfrey, p.117) was employed.

Group expenditure in demand systems is often postulated as not being weakly exogenous (LaFrance; Edgerton). The method used to test for endogeneity of expenditure was that of Banks, Blundell, and Lewbel, in which the residuals from an ordinary least squares regression of the logarithm of expenditure on the instruments are included as an extra regressor in the QAIDS or AIDS models. The instruments employed are described in Appendix B. Whether the estimated coefficients associated with these residuals are distinguishable from zero was judged using likelihood ratio tests.

Also, tests for the QAIDS versus the AIDS model were implemented. The null hypothesis of $\delta_i = 0 \; \forall i$ was tested using a likelihood ratio test.

Testing the three sets of hypotheses jointly requires some choice of the order of performing the tests. The most general model—QAIDS with endogenous expenditure and up to an AR(4) process—nests the other models. At the other end of the spectrum, the least general

[7] Although m denotes the total number of equations, one equation is omitted because the contemporaneous covariance matrix is singular. Hence, only (m – 1) equations are estimated. To preserve simple notation, the correction in notation to account for the omitted share equation is left tacit.

model is the AIDS model with exogenous expenditure and no AR process. The order in which the hypotheses were tested is as follows:

a) maintaining endogenous group expenditure and the AR(4) specification, QAIDS vs. AIDS was tested.

b) Given the results in a), endogeneity of expenditure was tested.

c) Given the results in a) and b), the order of the AR process was tested.

Varying the order of the testing had no impact on the choice of the order of the AR process. Given the order of tests performed in each area, the resulting models were chosen

Metropolitan Area	**Preferred Specification**
Albany	AIDS, Endogenous Expenditure, AR(3)
Atlanta	QAIDS, Endogenous Expenditure, AR(2)
Chicago	AIDS, Exogenous Expenditure, AR(2)
Dallas	QAIDS, Endogenous Expenditure, AR(3)
Los Angeles	AIDS, Endogenous Expenditure, AR(3)
Miami	QAIDS, Exogenous Expenditure, AR(2)

In very loose terms, the QAIDS model seems preferred for metropolitan areas in the southern United States—Atlanta, Dallas, and Miami—whereas the AIDS model appears adequate elsewhere. Exogeneity of group expenditure could not be rejected in Chicago (p-value=0.586) or Miami (p-value=0.456). In the remaining four areas, exogeneity was rejected with p-values ranging from 0.036 to 0.002. Even if size corrections for the tendency of the test statistics to over-reject (De Boer and Harkema) are applied inferences are not affected.

Due to the large number of uncompensated elasticities calculated—25 own- and cross-price as well as 5 expenditure elasticitiesfor each area—only selected elasticities will be discussed. Table 5 reports own-price elasticities for each of the five types of tomatoes. Note that the elasticities are evaluated at 1999 averages because consumer behavior in the most recent one third of the sample is likely more indicative of market developments since 1999.

Most own-price elasticities display considerable price responsiveness. Regular tomatoes (predominantly field grown, beefsteak varieties) have own-price elasticities larger in absolute value than unity in four of the six metropolitan areas. Even in Atlanta and Dallas, own-price elasticities for regular tomatoes are relatively large, -0.64 and –0.79, respectively. Greenhouse tomatoes also tend to display considerable price responsiveness: their values range from –0.48 (Dallas) to –3.08 (Atlanta). Curiously, however, tomatoes on the vine, which are also grown in greenhouses, show relatively less response to changes in own price than do greenhouse tomatoes.[8] At the extreme, the response to own-price change is indistinguishable from zero in Dallas. The own-price responsiveness of roma and cherry tomatoes tends to vary considerably across metropolitan areas. Consumer behavior across metropolitan areas seems to vary highly, particularly as concerns responses to roma and cherry tomato prices. And certain metropolitan areas tend to exhibit more price response: Dallas shows relatively less response to own-price changes in all tomato categories except roma tomatoes. Albany, by contrast, exhibits high responses to own-price changes in most categories.

Without considering cross-price responses, the own-price elasticities suggest retail volume sold of regular field grown and greenhouse tomatoes are most responsive to own-price changes. In many cases, a one percent reduction in the weekly retail price would induce more than a one percent increase in volume sold. In areas where the share of greenhouse tomatoes is small—Atlanta (3%), Miami (9%), and Albany (12%)—reductions in own price could increase

[8] The difference is own-price elasticities is all the more curious because, as indicated in table 3, prices for greenhouse and on-the-vine tomatoes do not always differ significantly.

volume sold considerably. Even in Los Angeles with a relatively high share of 21%, reducing greenhouse prices could evoke substantial increases in volume sold of greenhouse tomatoes.

Cross-price elasticities for regular, greenhouse, and on-the-vine tomatoes are displayed in table 6. In contrast to the absolute magnitude of own-price elasticities, nearly all the cross-price elasticities are quite small. In Albany, Los Angeles, and Miami, cross-price elasticities unequivocally indicate substitution between regular field-grown tomatoes and greenhouse tomatoes (beefsteak and on-the-vine-varieties). Regardless of the particular price changed, an increase on the price of one type of tomato induces increased volume sold of a competing type of tomato. In Albany, where consumers appeared to be relatively more sensitive to own-price changes, the cross-price elasticities tend to be relatively larger than those in other markets. Nonetheless, the substitution relationships indicated are small in magnitude, ranging from 0.06 (Miami) to 0.35 (Albany). Although small in magnitude, these elasticities suggest reductions in the prices of greenhouse tomatoes can induce a reduction in the sales of regular field grown tomatoes.

The negative cross-price elasticities, on the other hand, suggest a complementary relationship between various types of tomatoes. In most cases, the magnitudes of these negative cross-price elasticities are quite small in Chicago and Dallas but the notable exceptions are found in Atlanta. These apparently anomalous cross-price elasticities in Atlanta may partially be a reflection of the anomalous pricing of greenhouse tomatoes in Atlanta where greenhouse tomato prices average considerably lower than in any other area.

Conclusions

Fresh tomatoes sold at retail in the United States are now differentiated products. One of the key ways in which fresh tomatoes are differentiated is by production method: field grown vs.

greenhouse produced. Even greenhouse-grown tomatoes are differentiated by beefsteak varieties and on-the-vine products. Packaging and labeling of field- vs. greenhouse-grown tomatoes further differentiates these fresh produce items at retail. Field-grown varieties are not all beefsteak tomatoes; roma (small pear shaped) and cherry tomatoes are also produced in the field. Despite the significant differentiation of fresh tomatoes in recent years, little is know about consumer preferences for these various types of tomatoes.

The retail-level data employed in this study permit both a descriptive and an econometric analysis of product differentiation in fresh tomatoes. Data availability limits the analysis to six major metropolitan areas dispersed throughout the Untied States: Albany, Atlanta, Chicago, Dallas, Los Angeles, and Miami. Jointly these six areas account for just under 15 percent of the U.S. population and supermarket sales. Weekly data were available only beginning in 1997 because scanner data companies previously did not assemble data for products without universal product codes. Despite those limitations, weekly observations for 156 consecutive weeks from the beginning of 1997 to the end of 1999 afford ample variation for statistical analysis.

The introduction of greenhouse tomatoes—beefsteak and on-the-vine varieties—has changed the composition of fresh tomato sales at retail. In markets like Los Angeles where greenhouse tomatoes have made considerable inroads, market share of regular field-grown tomatoes has declined substantially to as low as 20 percent of weekly sales. Shares of roma and cherry tomatoes remained stable, while greenhouse and on-the-vine tomatoes have gained market share. Acceptance of greenhouse tomatoes is far from uniform. Consumers in Atlanta have been reticent to buy more beefsteak greenhouse tomatoes although sales of on-the-vine greenhouse tomatoes have increased markedly. Average share of all types of greenhouse tomatoes over the

1997-99 period ranged from a low of 17% in Chicago to a high of 36% in Los Angeles. Perhaps as important, shares of field grown products have declined continuously over the same period.

The econometric analysis indicates reductions in the price of greenhouse tomatoes could further increase volume sold at retail. Both regular field-grown tomatoes and greenhouse varieties display considerable price responsiveness with own-price elasticities exceeding one in absolute value. Given the high frequency of the data—weekly observations—at specific retail markets, the magnitudes of these elasticities are probably not surprising. The own-price elasticities suggest that week-to-week changes in retail prices are likely to induce larger than proportional changes in volume sold at retail.

Estimates of cross-price relationships between field-grown and greenhouse tomatoes suggest competition between these two types of tomatoes in Albany, Los Angeles, and Miami. Yet a price increase for, say, a field-grown tomato tends to have a modest effect on volume sold of greenhouse products: the largest cross-price elasticity was estimated at 0.35.

Seed companies, grower-shippers, repackers, and greenhouse growers can make use of these results derived from retail data. In broad terms, the analysis suggests there is still ample room for growth in greenhouse tomatoes of all types. Field-grown products, on the other hand, will face continued competition in the form of greenhouse tomatoes. In most cases, however, consumers are sensitive to retail prices, suggesting the any reductions in the costs of production, shipping, repacking, and promotion that are passed on to consumers will have substantial effects on retail sales.

References

Berndt, Ernst R. and N.E. Savin. “Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances,” *Econometrica* 43(1975):937-957.

Banks, James, Richard Blundell, and Arthur Lewbel. “Quadratic Engel Curves and Consumer Demand,” *Review of Economics and Statistics* 79:4(November 1997):527-539.

Deaton, Angus and John Muellbauer. “An Almost Ideal Demand System,” *American Economic Review* 70(June 1980):312-326.

De Boer, P.M.C. and R. Harkema, “Some Evidence on the Performance of Size Correction Factors in Testing Consumer Demand Models,” *Economics Letters* Vol. 29(1989):311-315.

Edgerton, David L. “On the Estimation of Separable Demand Models,” *Journal of Agricultural and Resource Economics* 18:2(December 1993):141-146.

Godfrey, L.G. *Misspecification Tests in Econometrics: The Lagrange Multiplier Principle and Other Approaches*, Cambridge: Cambridge University Press, 1988.

LaFrance, Jeffrey T. “When is Expenditure “Exogenous” in Separable Demand Models?,” *Western Journal of Agricultural Economics* 16:1 (July 1991):49-62.

Moschini, Giancarlo and Daniele Moro. “Autocorrelation Specification in Singular Equation Systems,” *Economics Letter* 46(1994):303-309.

United States Department of Commerce, U.S. Census Bureau, *QT-P3. Race and Hispanic or Latino: 2000*, by city, accessed on July 8, 2002 at http://factfinder.census.gov/bf/lang=en.

United States Department of Commerce, U.S. Census Bureau, (MA-99-3b) Population Estimates for Metropolitan Areas and Components, Annual Time Series, accessed at http://eire.census.gov/popest/archives/metro/ma99-03b.txt on April, 26, 2002

United States Department of Labor, Bureau of Labor Statistics, *Consumer Price Index, Food and Beverages Home, U.S. City Average, Seasonally Unadjusted (CUSR0000SAF11)* , accessed on July 6, 2002 at: http://146.142.4.24/labjava/outside.jsp?survey=cu.

White, K.J. "The Durbin-Watson Test for Autocorrelation in Nonlinear Models," *Review of Economics and Statistics*, 74:2(May 1992):370-373,

Table 1. Six U.S. Metropolitan Areas, 1997-1999.

Areas	(1) Average Population (Millions)	(2) Population Growth Rate	(3) Number of Households (Millions)	(4) Grocery ACV (Million Nominal $)	(5) Number of Supermarket Chains Reporting
Albany	0.96	-0.3%	0.41	1.43	2
Atlanta	3.57	6.2%	1.18	3.72	3
Chicago	8.40	1.4%	2.92	8.78	3
Dallas	4.62	5.1%	1.58	5.65	5
Los Angeles	15.07	3.0%	4.85	17.03	4
Miami	4.69	3.0%	1.68	5.68	3
U.S., Total	270.24	1.8%	95.30	309.62	
U.S.,Metropolitan	216.42	2.0%			
% of Total	13.8%		13.2%	13.7%	
% of Metropolitan	17.2%				

Sources: For columns (1) and (2), Population Estimates Program, Population Division, U.S. Census Bureau accessed at http://eire.census.gov/popest/archives/metro/ma99-03b.txt on April, 26, 2002. For columns (3) and (4), InfoScan: 1995 Market & Region Profiles, Information Resources, Inc.,1995. For column (5), Fresh Look LLC.

Table 2. Sample Means, Weekly Observations, Jan. 5, 1997 – Dec. 26, 1999.

	Albany	Atlanta	Chicago	Dallas	Los Angeles	Miami
Retail Price (Nominal $/lb.)						
Regular	1.84	1.99	2.25	1.29	1.73	1.83
Greenhouse	2.13	1.56	2.90	2.32	2.50	2.27
On the Vine	2.16	2.64	2.73	2.39	2.96	2.78
Roma	1.64	1.53	1.80	1.42	1.58	2.35
Cherry	2.70	2.00	2.22	2.20	2.36	2.08
Organic	2.85	2.04	2.92	2.38	2.63	2.12
Volume Sales (lbs.)						
Regular	109,385	66,111	243,576	261,733	437,711	289,830
Greenhouse	27,343	5,283	46,805	66,733	204,985	40,167
On the Vine	47,805	20,749	12,165	13,933	116,344	45,731
Roma	32,857	27,647	70,397	174,499	370,772	82,990
Cherry	11,495	3,490	46,226	5,950	49,653	34,401
Organic	360	26	778	4,681	1,040	6
Total	229,244	123,306	419,946	527,530	1,180,505	493,124
Per Capita	12.4	1.8	2.6	5.9	4.1	5.5
Market Share (Value)						
Regular	46%	54%	57%	43%	32%	53%
Greenhouse	12%	3%	13%	19%	21%	9%
On the Vine	23%	22%	4%	4%	15%	12%
Roma	12%	18%	14%	31%	26%	19%
Cherry	7%	3%	11%	2%	5%	7%
Organic	0.2%	0.02%	0.2%	1.3%	0.1%	0.002%

Table 3. T Tests for Differences in Mean Prices, Jan. 5, 1997 – Dec. 26, 1999.

Pairs of Mean Prices	Albany	Atlanta	Chicago	Dallas	Los Angeles	Miami
Regular-Greenhouse	-12.73	12.71	-15.06	-38.95	-25.55	-18.79
Regular-OTV	-15.28	-18.49	-10.68	-26.00	-24.32	-19.23
Regular-Roma	10.39	19.66	15.13	-9.58	7.98	-20.78
Regular-Cherry	-38.84	**-0.48**[a]	**1.14**	-36.18	-12.30	-4.46
Regular-Organic	-32.39	8.55	-32.13	-33.29	-20.36	47.06
Greenhouse-OTV[b]	**-2.04**	-46.85	2.71	**-1.41**	-10.69	-11.26
Greenhouse-Roma	24.73	**-2.31**	25.21	32.19	28.67	-4.26
Greenhouse-Cherry	-32.81	-11.29	20.60	3.70	**2.47**	4.26
Greenhouse-Organic	-29.62	2.86	**-0.41**	**-1.44**	**-2.80**	61.98
OTV-Roma	34.65	43.76	19.73	21.54	26.88	9.86
OTV-Cherry	-33.51	16.85	11.48	3.66	9.12	13.76
OTV-Organic	-36.69	18.54	-4.78	**0.08**	5.97	56.74
Roma-Cherry	-56.08	-16.67	-12.78	-27.51	-17.48	6.97
Roma-Organic	-50.71	4.73	-39.21	-25.91	-26.03	79.47
Cherry-Organic	-5.75	10.75	-30.73	-4.39	-5.20	43.85

[a] Negative values indicate the second price in a price pair is larger. Numbers of boldface indicate mean prices are *not* statistically different for α=0.01, one-tailed test.

[b] OTV denotes on-the-vine tomatoes.

Table 4. Changes in Market Shares and Volume Sales, 1997 vs. 1999[a]

	Regular	Greenhouse	On the Vine	Roma	Cherry	Organic
Albany						
Price	3%	-1%	-2%	-7%	4%	-7%
Volume	-33%	84%	61%	16%	39%	-50%
Shares	-17%	5%	9%	1%	2%	-0.22%
Atlanta						
Price	-24%	-18%	1%	0%	22%	
Volume	24%	705%	123%	-2%	-46%	
Shares	-11%	3%	13%	-3%	-2%	0.02%
Chicago						
Price	28%	-13%	59%	33%	15%	24%
Volume	-32%	143%	565%	-17%	27%	-38%
Shares	-15%	6%	7%	0%	2%	-0.14%
Dallas						
Price	4%	-16%	6%	7%	1%	20%
Volume	-17%	130%	187%	20%	16%	1232%
Shares	-15%	8%	4%	2%	0%	1.8%
Los Angeles						
Price	10%	-6%	3%	8%	28%	8%
Volume	-67%	129%	124%	-6%	9%	93%
Shares	-32%	16%	13%	2%	2%	0.05%
Miami						
Price	-28%	-7%	8%	3%	41%	
Volume	-13%	96%	24%	-8%	121%	
Shares	-20%	5%	5%	0%	9%	0.005%

[a] Percentages represent differences in averages over the 52 weeks of 1997 versus the 52 weeks of 1999.

Table 5. Uncompensated Own-Price Elasticities at 1999 Mean Values

Category	Albany	Atlanta	Chicago	Dallas	Los Angeles	Miami
Regular	-1.417	-0.64	-1.417	-0.789	-1.418	-1.23
	(.000)[a]	(.000)	(.000)	(.000)	(.000)	(.000)
Greenhouse	-1.86	-3.08	-0.94	-0.48	-1.38	-1.34
	(.000)	(.000)	(.025)	(.015)	(.000)	(.008)
On the Vine	-1.63	-0.68	-0.65	-0.01	-1.14	-0.91
	(.000)	(.095)	(.020)	(.983)	(.000)	(.042)
Roma	-0.91	-0.67	-1.43	-1.54	-0.85	-0.10
	(.002)	(.009)	(.000)	(.000)	(.000)	(.796)
Cherry	-2.53	0.52	-1.74	0.05	0.45	-0.89
	(.000)	(.000)	(.000)	(.123)	(.015)	(.000)

[a] Numbers in parentheses are p-values derived from the approximate standard errors calculated using the delta method.

Table 5. Selected Uncompensated Cross-Price Elasticities at 1999 Mean Values

Quantity/Price	Albany	Atlanta	Chicago	Dallas	Los Angeles	Miami
Regular/Greenhouse	0.23	0.05	-0.03	-0.05	0.14	0.13
Regular/OTV[a]	0.32	-0.24	-0.01	0.01	0.08	0.06
Greenhouse/Regular	0.12	-2.62	0.03	-0.03	0.22	1.69
OTV/Regular	0.35	-0.70	-0.03	0.22	0.26	0.49

[a] OTV denotes on-the-vine tomatoes

Appendix A. Population Calculations

In order to calculate per capita quantities of fresh tomatoes in each of the six metropolitan areas, population estimates by week for 156 consecutive weeks from 1/5/1997 to 12/26/1999 were needed. The U.S. Bureau of the Census provides estimates of the mid-year population of each county in the United States. For the six metropolitan areas considered here, the relevant counties are displayed in the table below.

Table A1. Counties Included in Selected Metropolitan Areas

Metropolitan Area & Counties	*Metropolitan Area & Counties*
Albany	**Chicago**
Albany	Cook
Montgomery	DuPage
Rensselaer	Kane
Saratoga	Lake
Schenectady	McHenry
Warren	Will
Washington	Lake
Atlanta	Porter
Barrow	**Dallas**
Cherokee	Collin
Clayton	Dallas
Cobb	Denton
Coweta	Ellis
DeKalb	Kaufman
Douglas	Rockwall
Fayette	Johnson
Forsyth	Parker
Fulton	Tarrant
Gwinnett	**Los Angeles**
Henry	Los Angeles
Newton	Orange
Paulding	Riverside
Rockdale	San Bernardino
Spalding	**Miami/Fort Lauderdale**
Walton	Broward
	Miami-Dade
	Palm Beach

Source: U.S. Census Bureau, (MA-99-3b) Population Estimates for Metropolitan Areas and Components, Annual Time Series, accessed at http://eire.census.gov/popest/archives/metro/ma99-03b.txt on April, 26, 2002

Mid-year estimates for 7/1/1990 through 7/1/1999 were extracted for each county above. This provided ten annual estimates of each county's population, which were summed appropriately to obtain each metropolitan area's population.

Table A2. Metropolitan Population, Mid-Year, (millions)

Mid-Year	**Albany**	**Atlanta**	**Chicago**	**Dallas**	**Los Angeles**	**Miami**
1990	0.950	2.835	7.880	3.906	13.924	4.075
1991	0.957	2.907	7.952	3.987	14.116	4.156
1992	0.963	2.984	8.032	4.052	14.328	4.226
1993	0.967	3.075	8.105	4.119	14.444	4.286
1994	0.970	3.175	8.170	4.199	14.498	4.373
1995	0.970	3.272	8.233	4.283	14.555	4.457
1996	0.966	3.368	8.292	4.384	14.658	4.543
1997	0.962	3.463	8.345	4.499	14.841	4.624
1998	0.960	3.569	8.402	4.618	15.064	4.692
1999	0.950	2.835	7.880	3.906	13.924	4.075

Once estimates of mid-year population for each metropolitan area were obtained, weekly estimates were obtained by fitting 4th-degree polynomials to the ten mid-year observations. The 4th-degree polynomial was relatively parsimonious and provided good fit—R^2 from 0.9969 to 0.999—for generating the weekly observations.

This approach to interpolating the 156 weekly population values requires at least two assumptions for use in subsequent econometric analysis. First, population should be exogenously determined. This seems a reasonable assumption in the context of demand for fresh tomatoes. Second, the data generating process for population must be relatively stable for weekly observations. Although weekly population is unobserved, one would suspect changes in population from year to year are not subject to sharp fluctuations. Hence, filling in the missing weekly values of population does not require unreasonable assumptions.

Appendix B. Instruments for Endogenous Group Expenditure

The weekly data employed in the econometric analysis calls for instrumental variables with sufficient weekly variation to be of use. The instruments chosen for this analysis are all observed daily, thereby affording aggregation of daily observations to weekly observations.

The focus of the analysis is retail consumption in selected U.S. markets. Most tomatoes consumed in these markets are produced in the United States, Mexico, Canada, and Holland. Accordingly, instruments reflecting economic conditions in these countries were selected. Daily measures of stock exchange activity for the United States, Mexico, and Canada were chosen. Specifically, the three measures were:

1. Dow Jones Industrial Average: Dow 30 Industrial Stock Price Index

2. Mexican IPC (*Indice de Precios y Cotizaciones*) Index – Daily Close

3. Toronto Composite Index – Daily Close

Besides the stock market indices, exchange rates were also selected as instruments. The rates used were:

1. Mexico, (Peso per US Dollar)

2. Canada, (Canadian Dollar per US Dollar)

3. Netherlands, (Guilder per US Dollar)

4. U.S. Dollar: Major Currencies Index, (March 1973=100)

All the daily observations on stock market indices and currency exchange rates were obtained from http://www.economy.com/freelunch.

For purposes of testing exogeneity, the 7 instruments above as well as the logarithm of prices were regressed on the logarithm of expenditure. The residuals from the regression were used as auxiliary regressors in the QAIDS and AIDS models to perform a Wu-Hausman type test for the exogeneity of tomato group expenditure.