IIMB Management Review (2016) 28, 20–24

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/iimb

IMR-DCAL SPECIAL PAPER

A simple example for the teaching of demand theory: Aggregate demand estimation for onions in India

Devlina Chatterjee *

IME Department, IIT Kanpur, Kanpur, Uttar Pradesh, India

Received 10 March 2014; revised 24 March 2015; accepted 22 January 2016; available online 27 February 2016

KEYWORDS
Empirical;
Demand;
Estimation;
Regression;
Onions;
Indian;
Wholesale

Abstract Managerial economics textbooks rarely include empirical examples of demand estimation of any commodity from real data, perhaps because in reality one must consider coupled demand systems. We suggest that on a national level and over a short time, the price–volume data for onions provide a bona fide example of a single-commodity demand curve. Since the onion has no real substitutes and taste for onions does not fluctuate, the demand curve does not shift over time. Empirical analysis of aggregated national level data yields a demand curve with two regimes: constant consumption at low prices, and constant budget at high prices.

Introduction

The theory of demand and supply is a cornerstone of microeconomics. The idea that the price of a commodity lies at the intersection of its supply and demand curves is central to the teaching of managerial economics (Salvatore & Srivastava, 2008). As per theory, production decisions of a firm are based on the customers' demand curve, which is again dependent on the market structure and the price elasticity of demand for a particular product. But how can we empirically estimate the demand for a certain commodity?

The literature on demand estimation deals with systems of equations that try to include all the variables that might affect the demand for a commodity (Deaton & Muellbauer, 1980). At high levels of aggregation one estimates demand for food vs non-food commodities. When dealing with demand for food items, one accounts for different food categories such as cereals, meat and fish, eggs, dairy, vegetables, fruits and nuts, etc. (Green & Alston, 1990). Such regression models however are too complicated to be introduced to students encountering both economics and statistics for the first time. For example, Chen (1977) has dealt with 23 simultaneous equations in a complicated statistical model. Understanding the implications of such models or replicating their estimation procedure is beyond the scope of the typical introductory managerial economics course.

With this motivation, a teacher of managerial economics might ask, "Is there a straightforward case of demand

* Tel.: +91 512 2596960; fax: +91 512 2597553.
E-mail address: devlina@iitk.ac.in.
Peer-review under responsibility of Indian Institute of Management Bangalore.

http://dx.doi.org/10.1016/j.iimb.2016.01.002
0970-3896

estimation from real data, of a single commodity, that one can use as a classroom example?" In this paper we set out to estimate the aggregate national demand curve for onions. Our interest in onions was spurred by two factors: one was the tremendous rise in onion prices observed from time to time; the second was the fact that onions are a commodity for which there are no close substitutes and therefore they can possibly be studied as a single commodity demand system. The typical features of the onion market have been described in greater detail in a later section.

Demand estimation—theory

Identification problem

In empirically estimating the demand curve for any good, one encounters the well known "identification problem" (Salvatore & Srivastava, 2008), which is briefly described here for completeness. Historical data points on prices and quantities for any commodity represent, in principle, different equilibria where *different pairs* of demand and supply curves intersect, as illustrated in Fig. 1(a).

Observed price quantity combinations (P_1, Q_1), (P_2, Q_2), and (P_3, Q_3) could be the result of the intersection of different demand curves D_1, D_2, and D_3 and different supply curves S_1, S_2, and S_3. Hence the fitted line AB that passes through the observed points E_1, E_2, and E_3 would not represent accurately either a unique underlying demand curve or supply curve (Fig. 1(a)).

In estimation of a demand curve, various exogenous variables that might affect demand need to be included, namely (a) income, (b) prices of substitutes or complements, (c) changes in tastes or preferences, and (d) changes in technology. Each of these variables would cause the demand curve to shift up or down (rather than cause movement along the demand curve). Subsequently, the estimated coefficient for price in the regression model (with all these variables included) gives an indication of the own price elasticity of demand.

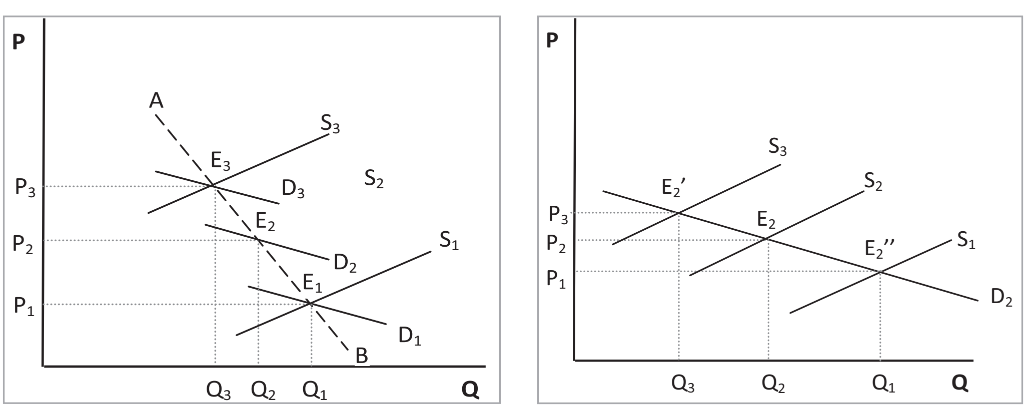
If it can be assumed that there have been no shifts in the demand curve in the period of observation, then the observed price–volume points must lie on *the same* demand curve (see points E_2', E_2 and E_2'' in Fig. 1(b)).

Onions

Seeking a simple example for empirical demand estimation, here we consider the demand for onions. Why onions? Onions have received attention from both the popular press as well as academia, mostly because onions prices have risen steeply compared to other food commodities. Prices of onions are believed to have toppled governments (Economist, 2013). Moreover, the four exogenous factors (a) through (d), discussed above, that affect demand of other commodities may not affect onion demand significantly as explained below.

(a) Income: Over a brief interval, the aggregate national income changes little, and onions constitute a relatively small portion of family expenditures in any case; hence changes in income have little effect on demand of onions.
(b) Prices of substitutes: Onions have no real substitutes in Indian cooking (contrast with cereals as a group; pulses; green vegetables; meat, fish and poultry).
(c) Taste: Onions are a popular ingredient in Indian cooking, and the Indian consumer's taste for onions varies little over time.
(d) Technology: Technology does not play an important role (for instance, dried onion powder is not a common substitute for fresh onions in Indian recipes).

All of these factors could in principle lead to a shift in the demand curve (as opposed to movement along the demand curve). Since none of these factors contributes significantly to the demand for onions, it can be assumed that the consumers' aggregate demand curve does not shift up or down (at least in the short run). Hence it would seem to follow that



Both demand and supply curves have shifted over time. Dotted line AB joining the observed price-volume points represents neither the true supply nor the true demand curve.

Only supply curves have shifted over time, and observed price-volume points lie on the true underlying demand curve.

Figure 1 The identification problem. (a): Both demand and supply curves have shifted over time. Dotted line AB joining the observed price–volume points represents neither the true supply nor the true demand curve. (b): Only supply curves have shifted over time, and observed price–volume points lie on the true underlying demand curve.

observed onion price and quantity data lie on the aggregate demand curve, with changes in prices predominantly caused by shifts in supply. Tentatively accepting this conclusion, we now attempt to estimate this demand curve.

Data and analysis

Description of data

We used data from two sources for the analysis.

1. Agricultural Marketing Information Network (http://agmarknet.nic.in/) has daily data on prices and arrivals for about 400 mandis across India (the number of mandis reported upon varied from day to day). The daily data from the agmarknet portal gives minimum, maximum, and modal prices and quantity of arrivals for about 400 mandis nationwide. Not every mandi has data for every day. Data were collected for 252 days from February 1, 2013 to October 24, 2013. These data were used for the price volume analysis that follows.
2. The National Horticulture Board (NHB) (www.nhb.gov.in) has data on both wholesale and retail prices. However, the data are reported for fewer (about 30) mandis across India than in the agmarknet portal.[1] We downloaded monthly averaged data for 22 months to look at the trends in prices. Daily data from the same source were used to look at the relationship between wholesale and retail prices.

Aggregation of data

Daily data of prices and quantities of arrivals of onions recorded at each of about 400 mandis were downloaded for the agmarknet portal. For each day the arrivals (or volumes) at each mandi, sorted in decreasing order, were added up. It was observed that, typically, less than 10% of all mandis contributed 80% of total volumes for each day. The sum of volumes at these main contributing mandis was then taken to represent the volume for the day. The price point for that day was taken to be the volume averaged modal price over these main mandis. Subsequently, daily price and volume data points were averaged (volume weighted) over non-overlapping periods of seven days each to get weekly data, thereby eliminating day-of-the-week effects. Thus we have national level price volume data for 36 weeks.[2]

In contrast to the agmarknet data, the data from the NHB required less processing. Aggregated monthly prices, both wholesale and retail, were available for 32 mandis across India. For each month, weighted average retail and wholesale prices were calculated (where the price reported in each mandi was weighted by its contribution to the total volume across all mandis reported). These prices were collected for 22 months. Daily data for 557 days over the same period were also obtained from this website. With regard to the daily data, two

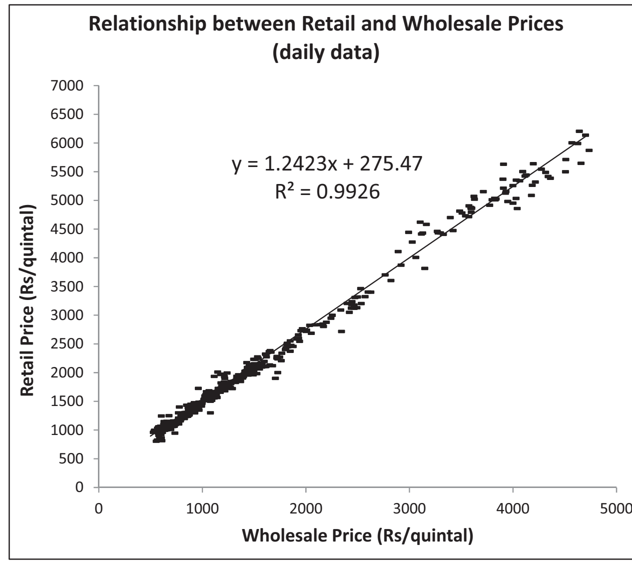


Figure 2 Retail vs wholesale prices (daily data from 32 mandis for 557 days).

points are noteworthy: first, data were not reported for Sundays and holidays; and second, there were a few obvious data entry errors that were manually corrected wherever detected.

Analysis and results

Retail vs wholesale prices

Fig. 2 shows the results of a linear regression analysis of daily retail prices against corresponding wholesale prices. The model fit is excellent ($R^2 = 0.993$). It is seen that the mark-up over wholesale prices is about 24% plus a constant of about Rs 275 per quintal (or Rs 2.75 per kg).

Statistically, the model shown in Fig. 2 is clearly robust. The R^2 is very high and the p-values are miniscule. However, visual examination of the data points suggests that there may be heteroskedasticity in the data. We performed a Breusch Pagan test[3] (see for example Stock & Watson, 2011) in R. The test statistics were BP = 100.5879, df = 1, p-value < 2.2e-16. That is, the data are indeed heteroskedastic. A possible economic interpretation of heteroskedasticity in the data is the following:

1. When prices are low and volumes are high, competition between retailers is greater, margins are smaller, and the scatter in the data seems to be smaller.
2. When volumes are low and prices are high, economies of scale suffer, inefficiencies are greater, absolute margins may be bigger, and day-to-day fluctuations in margins can be relatively larger. So the scatter in aggregated daily data is higher.

[1] Note that the agmarknet data, covering many more mandis, is more comprehensive than the NHB data. It therefore gives a more accurate view of the aggregated arrivals (volumes) on a national scale.
[2] These data are available in .csv format on the journal website (or by email from the author).

[3] Using the R package "lmtest" and the command "bptest".

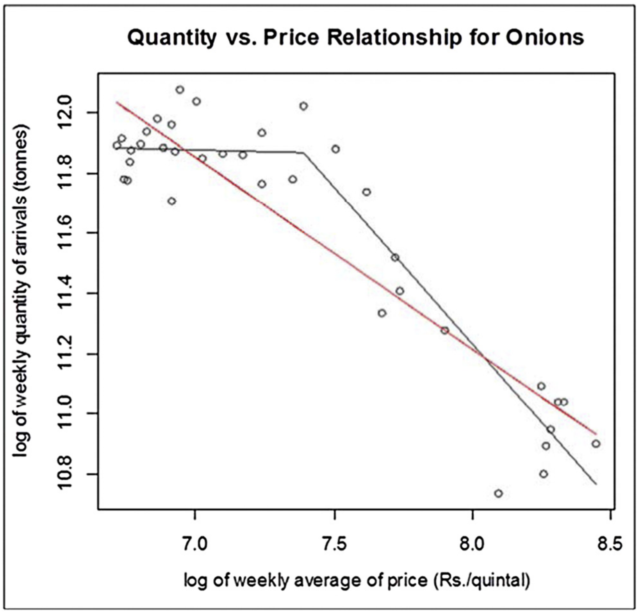


Figure 3 Total arrivals vs. weighted average price.

A model with heteroskedastic robust standard errors was built.[4] However, it was found that the fitted coefficients in the model only changed in the fourth decimal place. Therefore, the linear fit is deemed to be sufficiently robust and accurate, and it seems that *on average* there is not any undue profiteering between wholesale and retail markets, even at high price levels. As a result, it is permissible to look at aggregate volumes vs *wholesale* prices to study the demand curve (even though the individual customer sees the *retail* prices).

Aggregated weekly volume vs wholesale price

Now we consider Fig. 3, which shows the aggregated weekly volume vs wholesale price, on log–log axes, from the agmarknet data.

Two regression fits are shown in the curve (for detailed results see the Appendix). The first model is a simple linear regression for the entire data (adjusted $R^2 = 0.814$). However, the data suggest a breakpoint (i.e. a point where the slope changes suddenly). So we also fit a segmented model to the data using R (Vito, 2008). The result, plotted in the same figure, is clearly superior (adjusted $R^2 = 0.901$).

The logarithmic break point is estimated to be 7.39 with a standard error of about 0.1018 (highly statistically significant with a negligibly small p-value). This corresponds to a wholesale price of about Rs. 1620 per quintal. For log(price) <7.39, the slope of the regression line is estimated to be −0.04196 with a standard error of 0.15469 which shows that the slope there is not significantly different from zero (p value = 0.788). Accordingly, we conclude that if the wholesale price is lower than about Rs. 16 per kg, then the Indian consumers are insensitive to price.

For log(price) >7.39, the slope of the regression line is estimated to be −1.04200 with a standard error of 0.09958, i.e. the slope is not significantly different from −1. The slope of approximately −1 suggests that, in this regime, volume is inversely proportional to price. A doubling of the price, in this range, is accompanied by roughly halved consumption, and the total aggregate national expenditure remains approximately constant.

[4] Using the R package "car" and the command "coeftest".

Summary

To summarise, we have two main findings. The first is that there is a statistically robust linear relationship between wholesale and retail prices (albeit with greater variability in retail prices in high price regimes). The second, and perhaps more interesting, finding of this study is that there is an empirically observed break point in the volume–price relationship for onions, suggesting that volumes begin dropping when the price exceeds a certain threshold.

Going by the empirically estimated demand curve, we suggest that the price elasticity of onions is negligible below a threshold, and is about −1 for prices above this threshold. The low elasticity at low prices indicates Indians' preference for a fixed amount of onions in their diet. The zero elasticity regime is one of constant consumption, while the elasticity of −1 suggests constant aggregate expenditure on onions.

Thus, we have met the original goals of this study. We have found a commodity whose demand curve can reasonably be estimated from readily available data and which in turn leads to an interpretation that is both simple and at the same time interesting. We suggest that this may be used as a classroom example in the teaching of managerial economics.

Acknowledgements

I thank Mahima Mishra for downloading and processing some of the data, and Anindya Chatterjee for editorial comments.

Appendix A

Linear model

Call:
lm(formula = lwvol ~ lwprice)
Residuals:

Min	1Q	Median	3Q	Max
−0.41825	−0.11262	0.00615	0.08486	0.41929

Coefficients:

Estimate	Std.	Error	t value	Pr(>|t|)
(Intercept)	16.32462	0.38067	42.88	<2e-16 ***
lwpprice	−0.63852	0.05146	−12.41	3.55e-14 ***

—
Signif. codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1
Residual standard error: 0.1788 on 34 degrees of freedom
Multiple R-squared: 0.8191, Adjusted R-squared: 0.8138
F-statistic: 153.9 on 1 and 34 DF, p-value: 3.554e-14

Segmented model

```
Call:
segmented.lm(obj = lin.mod, seg.Z = ~lwprice, psi = 7.5)
Estimated Break-Point(s):
Est.                          St.Err
7.3900                        0.1018
t value for the gap-variable(s) V: 0
Meaningful coefficients of the linear terms:
Estimate     Std.      Error    t value   Pr(>|t|)
(Intercept)  12.17412  1.07346  11.341    9.57e-13 ***
lwprice      –0.04196  0.15469  –0.271    0.788
U1.lwprice   –1.00051  0.18397  –5.438    NA
—
Signif. codes: 0 “***” 0.001 “**” 0.01 “*” 0.05 “.” 0.1 “ ” 1
Residual standard error: 0.1305 on 32 degrees of freedom
Multiple R-Squared: 0.9092, Adjusted R-squared: 0.9007
> slope(segmented.mod)
$lwprice
             Est.      St.Err.  t value   CI(95%).l  CI(95%).u
slope1       –0.04196  0.15470  –0.2713   –0.3571    0.2731
slope2       –1.04200  0.09958  –10.4700  –1.2450    –0.8396
```

Appendix B:
Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.iimb.2016.01.002.

References

Chen, T. (1977). The Wharton agricultural model: structure, specification, and some simulation results. *American Journal of Agricultural Economics*, *59*(1), 107–116.

Deaton, A., & Muellbauer, J. (1980). An almost ideal demand system. *The American Economic Review*, *70*(3), 312–326.

Economist (2013). *India: The politics of rising onion prices*. The Economist Intelligence Unit. <http://www.eiu.com/industry/article/1931096777/india-the-politics-of-rising-onion-prices/2013-10-25> Accessed 02.11.13.

Green, R., & Alston, J. M. (1990). Elasticities in AIDS models. *American Journal of Agricultural Economics*, *72*(2), 442–445.

Salvatore, D., & Srivastava, R. (2008). *Managerial economics: principles and worldwide applications* (7th ed.). New York: Oxford University Press.

Stock, J., & Watson, M. (2011). *Introduction to econometrics* (3rd ed.). International Edition. Pearson Education.

Vito, M. R. M. (2008). Segmented: an R package to fit regression models with broken-line relationships. *R News*, *8/1*, 20–25. <http://cran.r-project.org/doc/Rnews/>.