Managerial Economics

Chapter 3

Empirical Methods for Demand Analysis

Table of Contents

Introduction

- Managerial Problem
- Estimating the Effect of an iTunes Price Change
- How can managers use the data to estimate the demand curve facing iTunes? How can managers determine if a price increase is likely to raise revenue, even though the quantity demanded will fall?
- Solution Approach
- Managers can use empirical methods to analyze economic relationships that affect a firm's demand.
- Empirical Methods
- Elasticity measures the responsiveness of one variable, such as quantity demanded, to a change in another variable, such a price.
- Regression analysis is a method used to estimate a mathematical relationship between a dependent variable, such as quantity demanded, and explanatory variables, such as price and income. This method requires identifying the properties and statistical significance of estimated coefficients, as well as model identification.
- Forecasting is the use of regression analysis to predict future values of important variables as sales or revenue.

3.1 Elasticity

- Price Elasticity of Demand

- The price elasticity of demand (or simply the elasticity of demand or the demand elasticity) is the percentage change in quantity demanded, Q, divided by the percentage change in price, p.

- Arc Price Elasticity: $\varepsilon = (\Delta Q/\text{Avg } Q)/(\Delta p/\text{Avg } p)$
  - It is an elasticity that uses the average price and average quantity as the denominator for percentage calculations.
  - In the formula $(\Delta Q/\text{Avg } Q)$ is the percentage change in quantity demanded and $(\Delta p/\text{Avg } p)$ is the percentage change in price.
  - Arc elasticity is based on a discrete change between two distinct price-quantity combinations on a demand curve.

- Point Elasticity: $\varepsilon = (\Delta Q / \Delta p)\ (p/Q)$
  - Point elasticity measures the effect of a small change in price on the quantity demanded.
  - In the formula, we are evaluating the elasticity at the point (Q, p) and $\Delta Q/\Delta p$ is the ratio of the change in quantity to the change in price.
  - Point elasticity is useful when the entire demand information is available.

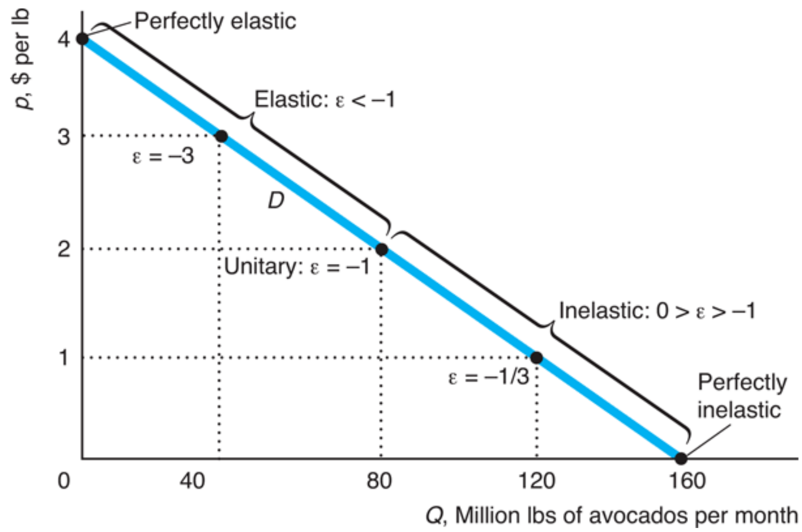
- Point Elasticity with Calculus: $\varepsilon = (\partial Q / \partial p)(p/Q)$
  - To use calculus, the change in price becomes very small.
  - $\Delta p \rightarrow 0$, the ratio $\Delta Q/\Delta p$ converges to the derivative $\partial Q/\partial p$

- Elasticity Along the Demand Curve
  - If the shape of the linear demand curve is downward sloping, elasticity varies along the demand curve.
  - The elasticity of demand is a more negative number the higher the price and hence the smaller the quantity.
  - In Figure 3.1, a 1% increase in price causes a larger percentage fall in quantity near the top (left) of the demand curve than near the bottom (right).

- Values of Elasticity Along a Linear Demand Curve
  - In Figure 3.1, the higher the price, the more elastic the demand curve.
  - The demand curve is perfectly inelastic ($\varepsilon = 0$) where the demand curve hits the horizontal axis.

– It is perfectly elastic where the demand curve hits the vertical axis, and has unitary elasticity at the midpoint of the demand curve.

- **Figure 3.1 The Elasticity of Demand Varies Along the Linear Avocado Demand Curve**



- Constant Elasticity Demand Form: $Q = Ap^{\varepsilon}$

  – Along a constant-elasticity demand curve, the elasticity of demand is the same at every price and equal to the exponent ε.

- Horizontal Demand Curves: ε = -∞ at every point

  – If the price increases even slightly, demand falls to zero.

  – The demand curve is perfectly elastic: a small increase in prices causes an infinite drop in quantity.

  – Why would a good's demand curve be horizontal? One reason is that consumers view this good as identical to another good and do not care which one they buy.

- Vertical Demand Curves: ε = 0 at every point

  – If the price goes up, the quantity demanded is unchanged, so $\Delta Q=0$.

  – The demand curve is perfectly inelastic.

  – A demand curve is vertical for essential goods—goods that people feel they must have and will pay anything to get.

- Other Elasticity: Income Elasticity, $(\Delta Q/Q)/(\Delta Y/Y)$

  – Income elasticity is the percentage change in the quantity demanded divided by the percentage change in income *Y*.

  – Normal goods have positive income elasticity, such as avocados.

- Inferior goods have negative income elasticity, such as instant soup.

- Other Elasticity: Cross-Price Elasticity, $(\Delta Q/Q)/(\Delta p_o/p_o)$
  - Cross-price elasticity is the percentage change in the quantity demanded divided by the percentage change in the price of another good, p_o
  - Complement goods have negative cross-price elasticity, such as cream and coffee.
  - Substitute goods have positive cross-price elasticity, such as avocados and tomatoes.
- Demand Elasticities over Time
  - The shape of a demand curve depends on the time period under consideration.
  - It is easy to substitute between products in the long run but not in the short run.
  - A survey of hundreds of estimates of gasoline demand elasticities across many countries (Espey, 1998) found that the average estimate of the short-run elasticity was –0.26, and the long-run elasticity was –0.58.
- Other Elasticities
  - The relationship between any two related variables can be summarized by an elasticity.
  - A manager might be interested in the price elasticity of supply—which indicates the percentage increase in quantity supplied arising from a 1% increase in price.
  - Or, the elasticity of cost with respect to output, which shows the percentage increase in cost arising from a 1% increase in output.
  - Or, during labor negotiations, the elasticity of output with respect to labor, which would show the percentage increase in output arising from a 1% increase in labor input, holding other inputs constant.
- Estimating Demand Elasticities
  - Managers use price, income, and cross-price elasticities to set prices.
  - However, managers might need an estimate of the entire demand curve to have demand elasticities before making any real price change. The tool needed is regression analysis.
- Calculation of Arc Elasticity
  - To calculate an arc elasticity, managers use data from before the price change and after the price change.
  - By comparing quantities just before and just after a price change, managers can be reasonably sure that other variables, such as income, have not changed appreciably.

- Calculation of Elasticity Using Regression Analysis
  - A manager might want an estimate of the demand elasticity before actually making a price change to avoid a potentially expensive mistake.
  - A manager may fear a reaction by a rival firm in response to a pricing experiment, so they would like to have demand elasticity in advance.
  - A manager would like to know the effect on demand of many possible price changes rather than focusing on just one price change.

3.2 Regression Analysis

A Demand Function Example

- Demand Function: $Q = a + bp + e$
  - Quantity is a function of price; quantity is on the left-hand side and the price on the right-hand side; Q is the dependent variable and p is the explanatory variable
  - e is the random error
  - It is a linear demand (straight line).
  - If a manager surveys customers about how many units they will buy at various prices, he is using data to estimate the demand function.
  - The estimated sign of b must be negative to reflect a demand function.
- Inverse Demand Function: $p = g + hQ + e$
  - Price is a function of quantity; price is on the left-hand side and the quantity on the right-hand side; p is the dependent variable and Q is the explanatory variable
  - e is the random error
  - It is based on the previous demand function, so $g = –a/b > 0$ and $h = 1/b < 0$ and has a specific linear form
  - If a manager surveys how much customers were willing to pay for various units of a product or service, he would estimate the inverse demand equation.
  - The estimated sign of h must be negative to reflect an inverse demand function.

Example Inverse Demand Function, Portland Fish:

$p = g + hQ + e$, so estimation is $\hat{p} = \hat{g} + \hat{h}Q$

- *g* and *h* are the true coefficients
- The OLS regression provides estimates of these coefficients, $\hat{g}$ and $\hat{h}$, which we can use to predict the expected price, $\hat{p}$, for a given quantity. It is assumed $e = 0$.
- Use Microsoft Excel Trendline option for scatterplots to estimate $\hat{g}$ and $\hat{h}$ using OLS and get the respective graph and function.
- The estimated coefficients are $\hat{g}$ = 1.96 and $\hat{h}$ = 0.15. So the inverse demand curve is $\hat{p}$= 1.96 – 0.15*Q*
- The estimated change in price needed to induce buyers to purchase one more unit (1000 lbs) is $\hat{h}$= –$0.15 = –15¢.

Multivariate Regression: $p = g + hQ + iY + e$

- Multivariate Regression is a regression with two or more explanatory variables, for instance $p = g + hQ + iY + e$
- This is an inverse demand function that incorporates both quantity and income as explanatory variables.
- *g*, *h*, and *i* are coefficients to be estimated, and *e* is a random error

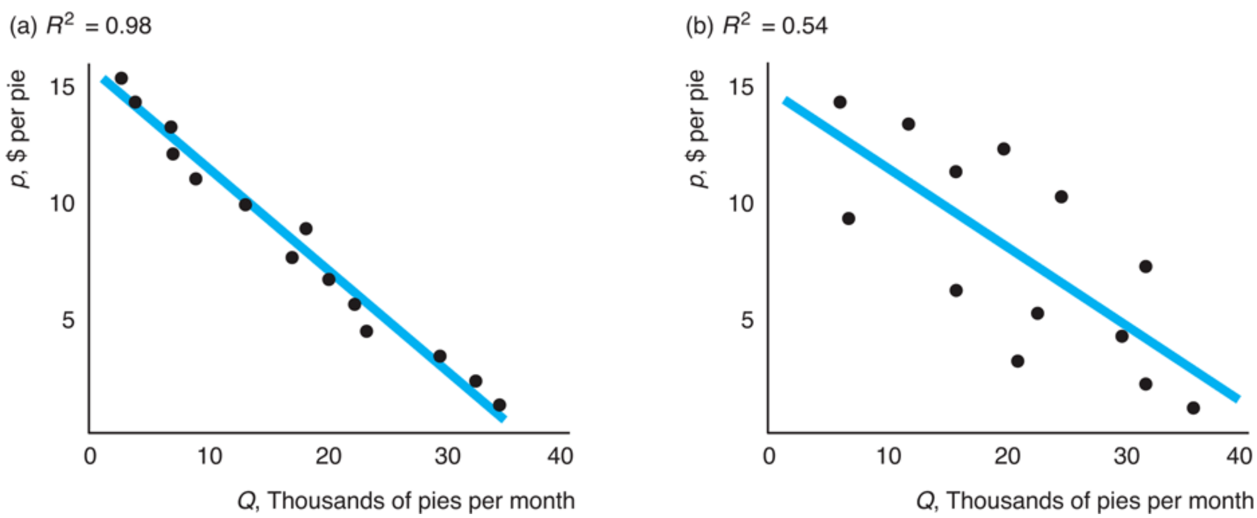
Corresponding Estimated Regression: $\hat{p} = \hat{g} + \hat{h}Q + \hat{i}Y$

- $\hat{g}$, h, and $\hat{i}$ are the estimated coefficients and $\hat{p}$ is the *predicted value* of *p* for any given levels of *Q* and *Y*.
- The objective of an OLS multivariate regression is to fit the data so that the sum of squared residuals is as small as possible.
- A multivariate regression is able to isolate the effects of each explanatory variable holding the other explanatory variables constant.

- Goodness of Fit and the R^2 Statistic
  - The R^2 (*R*-squared) statistic is a measure of the goodness of fit of the regression line to the data.
  - The R^2 statistic is the share of the dependent variable’s variation that is explained by the regression.
  - The R^2 statistic must lie between 0 and 1.
  - 1 indicates that 100% of the variation in the dependent variable is explained by the regression.

- Figure 3.5 shows a regression with $R^2 = 0.98$ in panel a and another with $R^2 = 0.54$ in panel b. Data points in panel a are close to the linear estimated demand, while they are more widely scattered in panel b.

- **Figure 3.5 Two Estimated Apple Pie Demand Curves with Different R^2 Statistics**



3.3 Properties & Significance of Coefficients

- Key Questions When Estimating Coefficients
  - How close are the estimated coefficients of the demand equation to the true values, for instance â respect to the true value *a*?
  - How are the estimates based on a sample reflecting the true values of the entire population?
  - Are the sample estimates on target?
- Repeated Samples
  - We trust the regression results if the estimated coefficients were the same or very close for regressions performed with repeated samples (different samples).
  - However, it is costly, difficult, or impossible to gather repeated samples or sub-samples to assess the reliability of regression estimates.
  - So, we focus on the properties of both <u>estimating methods</u> and <u>estimated coefficients</u>.
- OLS Desirable Properties
  - Ordinary Least Squares (OLS) is an unbiased estimation method under mild conditions. It produces an estimated coefficient, $\hat{a}$, that equals the true coefficient, a, on average.

- OLS estimation method produces estimates that vary less than other relevant unbiased estimation methods under a wide range of conditions.

- Estimated Coefficients and Standard Error
  - Each estimated coefficient has an standard error.
  - The smaller the standard error of an estimated coefficient, the smaller the expected variation in the estimates obtained from different samples.
  - So, we use the standard error to evaluate the significance of estimated coefficients.
- A Focus Group Example
- Estimate a linear D curve $Q = a + bp + e$
  - We use OLS using Microsoft Excel
  - Excel LINEST (for *line est*imate) function gives the parameter estimates and errors.
-
  - The estimate of b is –1.438 (cell A12), and its estimated standard error is 0.090 (cell A13).
  - The estimate for a is 53.857 (cell B12), and its estimated standard error is 2.260 (cell B13).
  - R^2 is 0.977, which is close to the maximum possible value (cell A14). This high R^2 indicates that the regression line explains almost all the variation in the observed quantity.
- Use the estimated D to estimate the q demanded at any p
  - If the price is 27 ($27,000), we expect consumers to buy $\hat{Q}$ = 53.857 – (1.438*27) = 15.031.
  - We round this estimate to 15 cars, given that cars are sold in discrete units.
  - If this focus group represented a large group, perhaps a thousand times larger, the quantity demanded estimate would be 15,031 cars.
- Confidence Intervals
  - A confidence interval provides a range of likely values for the true value of a coefficient, centered on the estimated coefficient.
  - A 95% confidence interval is a range of coefficient values such that there is a 95% probability that the true value of the coefficient lies in the specified interval.

- Simple Rule for Confidence Intervals
  - Rule: If the sample has more than 30 degrees of freedom, the 95% confidence interval is approximately the estimated coefficient minus/plus twice its estimated standard error.
  - A more precise confidence interval depends on the estimated *standard error* of the coefficient and the number of degrees of freedom. The relevant number can be found in a *t*-statistic distribution table.
- Hypothesis Testing
  - Suppose a firm's manager runs a regression where the demand for the firm's product is a function of the product's price and the prices charged by several possible rivals.
  - If the true coefficient on a rival's price is 0, the manager can ignore that firm when making decisions.
  - Thus, the manager wants to formally test the null hypothesis that the rival's coefficient is 0.
- Testing Approach Using the t-statistic
  - One approach is to determine whether the 95% confidence interval for that coefficient includes zero.
  - Equivalently, the manager can test the null hypothesis that the coefficient is zero using a *t*-statistic. The *t*-statistic equals the estimated coefficient divided by its estimated standard error. That is, the *t*-statistic measures whether the estimated coefficient is large relative to the standard error.
- Statistically Significantly Different from Zero
  - In a large sample, if the *t*-statistic is greater than about two, we reject the null hypothesis that the proposed explanatory variable has no effect at the 5% significance level or 95% confidence level.
  - Most analysts would just say the explanatory variable is statistically significant.

3.4 Regression Specification

- Selecting Explanatory Variables
  - A regression analysis is valid only if the regression equation is correctly specified.
- Criteria for Regression Equation Specification
  - It should include all the observable variables that are likely to have a meaningful effect on the dependent variable.

- It must closely approximate the true functional form.
- The underlying assumptions about the error term should be correct.
- We use our understanding of causal relationships, including those that derive from economic theory, to select explanatory variables.

- Selecting Variables, Mini Case: $Y = a + bA + cL + dS + fX + e$
  - The dependent variable, *Y*, is CEO compensation in 000 of dollars.
  - The explanatory variables are assets *A*, number of workers *L*, average return on stocks *S* and CEO's experience *X*.
  - OLS regression: $\hat{Y} = -377 + 3.86A + 2.27L + 4.51S + 36.1X$
  - *t*-statistics for the coefficients for *A*, *L*, *S* and *X* : 5.52, 4.48, 3.17 and 4.25.
  - Based on these *t*-statistics, all 4 variables are 'statistically significant.'

- Statistically Significant vs. Economically Significant
  - Although all these variables are statistically significantly different than zero, not all of them are economically significant.
  - For instance, *S* is statistically significant but its effect on CEO's compensation is very small: an increase in shareholder return of one percentage point would add less than $5 thousand per year to the CEO's wage.
  - So, *S* is statistically significant but economically not very important.

- Correlation and Causation
  - Two variables are correlated if they move together. The *q* demanded and *p* are negatively correlated: *p* goes up, *q* goes down. This correlation is causal, changes in *p* directly affect *q*.
  - However, correlation does not necessarily imply causation. For example, sales of gasoline and the incidence of sunburn are positively correlated, but one doesn't cause the other.
  - Thus, it is critical that we do not include explanatory variables that have only a spurious relationship to the dependent variable in a regression equation. In estimating gasoline demand we would include price, income, sunshine hours, but never sunburn incidence.

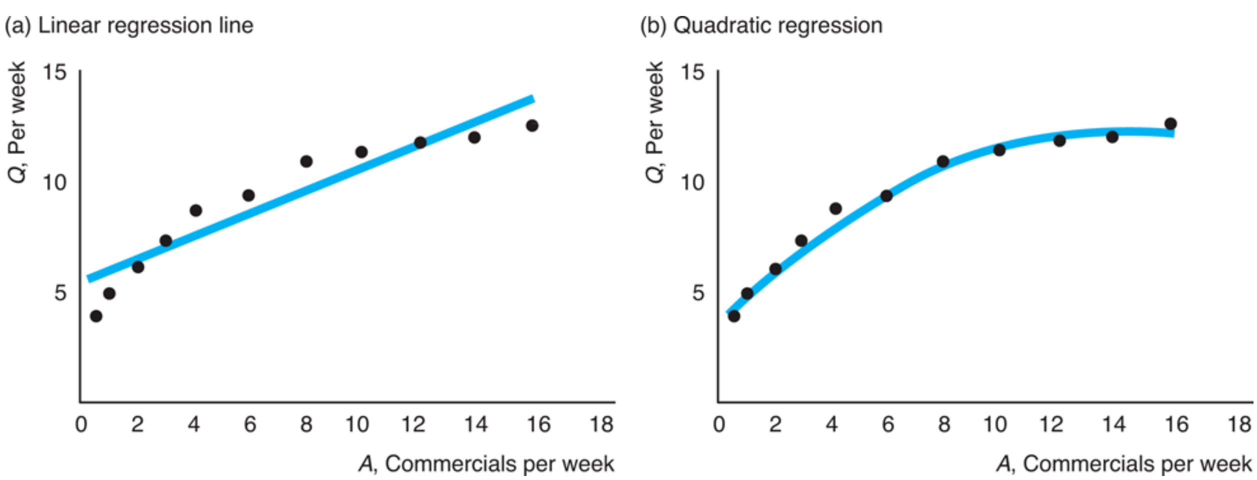
- Omitted Variables
  - These are the variables that are not included in the regression specification because of lack of information. So, there is not too much a manager can do.

- However, if one or more key explanatory variables are missing, then the resulting coefficient estimates and hypothesis tests may be unreliable.
- A low R^2 may signal the presence of omitted variables, but it is theory and logic that will determine what key variables are missing in the regression specification.

- Functional Form
  - We cannot assume that demand curves or other economic relationships are always linear.
  - Choosing the correct functional form may be difficult.
  - One useful step, especially if there is only one explanatory variable, is to plot the data and the estimated regression line for each functional form under consideration.

- Graphical Presentation
  - In Figure 3.6, the quadratic regression ($Q = a + bA + cA^2 + e$) in panel b fits better than the linear regression ($Q = a + bA + e$) in panel a.

- **Figure 3.6 The Effect of Advertising on Demand**



- Extrapolation
  - Extrapolation seeks to forecast a variable as a function of time.
  - Extrapolation starts with a series of observations called time series.
  - The time series is <u>smoothed</u> in some way to reveal the underlying pattern, and this pattern is then extrapolated into the future.
  - Two linear smoothing techniques are <u>trend line</u> and <u>seasonal variation</u>.
  - Not all time trends are linear

- Trend Line: $R = a + bt + e$, *where t is time*
  - If this is the trend for Heinz Revenue, *a* and *b* are the coefficients to be estimated.

- The estimated trend line is $R = 2{,}089 + 27.66t$, with statistically significant coefficients.
- Heinz could forecast its sales in the first quarter of 2014, which is quarter 37, as 2089 + (27.66 × 37) = $3.112 billion.

- Seasonal Variation: $R = a + bt + c_1D_1 + c_2D_2 + c_3D_3 + e$
  - Heinz revenue data shows a quarterly trend that is captured with seasonal dummy variables, D_1, D_2, and D_3.
  - The new estimated trend is $R = 2{,}094 + 27.97t + 93.8D_1 - 125.3D_2 - 8.60D_3$, with all coefficients statistically significant.
  - The forecast value for the first quarter of 2014 is 2094 + (27.97 × 37) + (93.8 × 1) – (125.3 × 0) – (8.60 × 0) = $3.223 billion.

3.5 Forecasting

- Theory-Based Econometric Forecasting
  - We estimated Heinz revenue with time trend and dummy seasonal variables. However, revenue is determined in large part by the consumers' demand curve, and the demand is affected by variables such as income, population, and advertising. Extrapolation (pure time series analysis) ignored these structural (causal) variables.
  - *Theory-based econometric forecasting* methods incorporate both extrapolation and estimation of causal or explanatory economic relationships.
  - We use these estimates to make *conditional forecast*s, where our forecast is based on specified values for the explanatory variables.

Managerial Solution

- Managerial Problem
  - How can managers use the data to estimate the demand curve facing iTunes? How can managers determine if a price increase is likely to raise revenue, even though the quantity demanded will fall?
- Solution
  - To generate data, authors asked a focus group of 20 Canadian college students in 2008 how many songs they downloaded from iTunes when p was 99¢ and how many they would have downloaded at various other prices.
  - The estimated linear demand curve is $Q = 1024 - 413p$.
  - The t-statistic is –12.8, so this coefficient for price is significantly different from zero. The R^2 statistic is 0.96, so the regression line fits the data closely.
  - Apple's manager could use such an estimated demand curve to determine how revenue ($R = p \times Q$), varies with price. At p = 99¢, 615 songs were downloaded, so R_1 = $609. When p = $1.24, the number of songs drop to 512, R_2 = $635. Revenue increased by $26.
  - If the general population has similar tastes to the focus group, then Apple's revenue would increase if it raised its price to $1.24 per song.