Note on the Economics of Energy Demand[1]

This note discusses the following topics in the economics of energy demand: energy demand as a derived demand, methods of estimating energy demand elasticities, results of demand elasticity estimation, and some findings that suggest that the basic economic model of energy demand is incomplete.

Derived Demand

It is hard to believe that anybody really wants to consume electricity, natural gas, gasoline, or coal directly; doing so would be unpleasant at best and fatal at worst. The demand for energy is, rather, *derived* from the demand for services it is used to produce.[2] Gasoline, vehicles, and labor are the main inputs into the production of transportation services, for instance. Some residential demand for electricity is derived from the demand for food preparation. Commercial demand for electricity is largely derived from demands for heat, light, cooling, and, increasingly, computation; and these are ultimately derived from demand for firms' outputs. As Smil (2000) documents, global energy demand increased dramatically during the 20th century, but the efficiency with which energy services were produced increased even more rapidly.

A simple transportation example illustrates some properties of derived demand. In the *short run*, with John's car given and neglecting the influence of income, weather, and other variables, it is reasonable to suppose that the miles of auto transportation he demands, T, is lower the higher is the cost per mile of driving, C_T. Let us further suppose that the relation between these quantities can be approximated by

$$T = a(C_T)^{-e}, \tag{1}$$

where a and e are positive constants. Note that e is the elasticity of demand of T with respect to its cost, C_T:

$$-(dT/dC_T)(C_T/T) = e. \tag{2}$$

In the short run, ignoring the cost (positive or negative) of the time John spends driving, the production function for this energy service, miles of auto transportation, is approximately given by

$$T = mG, \tag{3}$$

where m is a positive constant that gives John's car's mileage in miles per gallon, and G is the consumption of gasoline used to produce T miles of transportation.

[1] This note was prepared for classroom use only. Please do not reproduce, cite, or quote from it.

[2] The theory of derived demand was initially developed to show how an industry's demand for labor depended on the demand for its output, as well as the production function used to produce that output.

Denote the per-gallon price of gasoline by P_G. In order to derive the demand equation for gasoline, we need to solve for T and C_T in the energy service demand equation (1) as functions of G and P_G. In this case the simple production function (3) makes this easy:

(4a) $T[\text{miles}] = m[\text{miles/gallon}]G[\text{gallons}]$, and

(4b) $C_T[\$/\text{mile}] = P_G[\$/\text{gallon}]/m[\text{miles/gallon}]$.

Units are shown in brackets – they are always worth checking. Substituting from (4) into (1), we obtain

(5a) $mG = a(P_G/m)^{-e}$, which simplifies to

(5b) $G = a(m)^{e-1}(P_G)^{-e}$.

A couple of things are worth noting about the derived demand equation (5). First, the price elasticity of demand for gasoline is exactly equal to the price elasticity of demand for transportation services. This is not generally the case, however; the derived demand for an input may be more or less elastic than the output demand function from which it is derived.[3] Second, unless $e = 0$, doubling John's car's mileage will not cut his demand for gasoline in half. In fact, if $e > 1$, increasing his car's mileage will increase his demand for gasoline. This happens because raising m lowers the per-mile cost of driving, thus tending to increase John's demand for transportation. (This point is made by critics of CAFE standards, but almost nobody argues that e in fact exceeds one for cars in the real world.)

What about the *long run*, in which all inputs (including, in particular, John's car) are variable?[4] In the long run an increase in the price of energy (gasoline in the example above) will reduce energy demand through both an *output effect* and a *substitution effect*.

The *output effect* is the whole story in the example above: John can only respond to an increase in the price of gasoline by reducing his demand for the transportation services it produces. It follows that the less elastic is the demand for the final service (the smaller is e in the example above), the less elastic will be the demand for the input that produces it, because the output effect will be smaller.

The *substitution* effect arises when it is possible to respond to an energy price increase by producing the same quantity of services using less energy, by substituting other inputs. In our example, in the long run John might well respond to an increase in the price of gasoline by buying a higher-mileage car. A general principle, first stated clearly by Paul Samuelson, implies that increasing the number of inputs a firm or household can vary in response to a change in the price of any one input gives more scope for the substitution effect and thus makes demand more price-elastic. So one generally expects long-run price elasticities of demand to exceed short-run

[3] Suppose, for instance, that John's car could run just as well on diesel fuel and that diesel fuel and gasoline were selling at the same price. Even a tiny increase in the price of gasoline would then be expected to reduce the demand for gasoline to zero.

[4] In the economic long run, technology is assumed fixed. To avoid various sorts of confusion it is important to distinguish between changes made over time in response to, e.g., price changes with a given set of technological possibilities, on the one hand, and changes caused by improvements in technology over time, on the other hand.

price elasticities, and, as we discuss below, the available evidence is consistent with this expectation.

It is also interesting to know how energy demand varies with household income or, in the case of businesses, with the volume of production. Samuelson's general principle that more scope for reaction implies a larger reaction means that long-run income (or production) elasticities of demand for energy will tend to exceed short-run elasticities, and that, too, is what is generally observed.

Estimating Demand Elasticities

Many technical articles have been written on problems associated with the estimation of demand elasticities. Here we will briefly discuss two that arise regularly in demand estimation: ensuring identification and allowing for dynamic behavior.

It is important to recognize that in practice, it is usually at least as important to scrutinize the available data carefully as to use appropriate statistical methods. Schmalensee and Stoker (1999), for instance, report a statistical study of gasoline demand using survey data on individual households and state-of-the-art estimation methods. They report plausible price elasticity estimates but note that subsequent discussions with the government agency that provided the data revealed that the household-level price data did not in fact measure prices that households faced, so the estimated elasticities were meaningless!

Identification

Among the first applications of statistical methods in economics was the estimation of demand functions for wheat, corn, and other agricultural commodities. Data were abundant, and the response of prices to changes in harvests was of great interest to farmers and others. Unfortunately, a number of early studies concluded that increases in price would increase the quantity demanded! Because the scatters of observed price/quantity observations tended upward, the estimated demand curves had positive slopes.

Prices and quantities are determined by both supply and demand, of course. Suppose the demand curve is perfectly stable and the supply curve is shifted by changes in such things as the weather and farmers' costs. In this case, which the analysts were implicitly assuming, shifts in the supply curve will serve to trace out the demand curve. Now suppose the supply curve is stable and changes in income, tastes, and other factors shift the demand curve. In this case, the shifts in the demand curve would trace out the supply curve.

In reality, neither the demand curve nor the supply curve in any market is likely to be perfectly stable, and the scatter of points generated by their intersections (that is, by the market) could have a positive, negative, or zero slope, depending on the shapes of the curves and the relative sizes of their shifts. In this situation one can only observe the solution to a pair of *simultaneous equations,* and the problem of figuring out how to estimate the parameters of one or more of the underlying equations is termed the *identification* problem.

Because prices in energy markets are generally affected by simultaneous (or near-simultaneous) decisions by suppliers and demanders, the identification problem also arises in attempts to estimate demand functions for particular forms of energy. Thus Kamerschen and Porter (2004) obtain estimates of upward-sloping electricity demand curves. Their solution – and solutions to the identification problem in other settings – involves finding variables that shift one of the curves but not the other. As the two extreme cases discussed above suggest, to estimate the demand curve one needs variables that shift supply but not demand, while to estimate the supply curve one needs variables that shift demand but not supply.

To see how this works in a very simple case, suppose the supply function in the market for wheat is assumed to be

(6a) $Q_S = \alpha + \beta P - \gamma P_G + \varepsilon$

where Q_S is the quantity of wheat supplied, P is the price of wheat, P_G is the price of gasoline, ε is an error or disturbance terms (i.e., a random variable summarizing the effects of omitted factors), and α, β, and γ are unknown positive constants. Similarly, suppose the demand function in this market is assumed to be

(6b) $Q_D = a - bP + cI + e$,

where Q_D is the quantity of wheat demanded, P is as before, I is consumer income, e is an error or disturbance term, and a, b, and c are unknown positive constants. The observations available are where supply equals demand:

(6c) $Q = Q_D = Q_S$.

In order actually to estimate the coefficients in (6), one would have to make some assumptions about the random (or error) terms ε and e and use an appropriate statistical method. Before getting that far, however, one has to ask whether it is possible *in principle* to estimate those coefficients, even if the error terms were zero. To address that question, set the error terms in (6) equal to zero and solve for Q and P as functions of I and P_G:

(7a) $P = [(a-\alpha)/(b+\beta)] + [c/(b+\beta)]I + [\gamma/(b+\beta)]P_G$, and

(7b) $Q = [(a\beta + \alpha b)/(b+\beta)] + [c\beta/(b+\beta)]I - [b\gamma/(b+\beta)]P_G$

Given data on all the quantities involved, it should be possible to estimate the coefficients in brackets in equations (7). Let us assume that we have "good" estimates of these (reduced form) coefficients. Is it possible to use those estimates to compute estimates of b and/or β?

Four cases naturally arise. First, if $c = \gamma = 0$, we are back where we started: all we can estimate are the intercepts in (7), and there is no way to go from them to any of the (structural) coefficients in (6). Now suppose that $\gamma = 0$ and $c > 0$, so we only have a variable that shifts the demand curve. In this case the ratio of the estimated coefficients of I gives us an estimate of β, but there is no way to get an estimate of b. In this case the supply curve shifts, and the data trace out the demand curve. Similarly, if $c = 0$ and $\gamma > 0$, we can estimate b but not β. Finally, if $\gamma > 0$ and $c > 0$, we can go back from estimates of the coefficients in (7) to all the coefficients in (6). In this case, we say that both equations are identified.

Dynamic Behavior

We noted above that the long-run response to change in prices or incomes is likely to differ from and, generally, to be larger than the short-run response. In the case of energy, the difference between the short run and the long run mainly has to do with changes in fixed assets like automobiles, appliances, industrial equipment, and the like. However, it also reflects job and housing choices and the associated commuting and shopping patterns, as well as changes in transportation for a given commuting pattern e.g. car-pools, public transportation, bicycles, and walking. If one had household or firm data that included information on fixed assets and other durable choices, one might be able to model the process of changing those assets as part of the response to changes in price and other variables. But such data are rarely if ever available.

Instead, modelers commonly make assumptions about the form of response over time and use data to estimate the speed of that response. Perhaps the most common assumption is the so-called partial adjustment model:

(8a) $$Q(t) - Q(t\text{-}1) = \lambda[Q^*(X_1, X_2, \ldots, X_N) - Q(t\text{-}1)],$$

where Q(t) is the quantity demanded in period t, Q* is the long-run demand function that describes behavior after all adjustments (including acquiring new assets) are complete, the X's are the variables that determine long-run demand, and λ is a positive constant to be estimated. The assumption here is that the adjustment process eliminates a constant fraction, λ, of the difference between long-run (or equilibrium) demand and actual demand in each period. Re-arranging (8a) yields

(8b) $$Q(t) = \lambda Q^*(X_1, X_2, \ldots, X_N) + (1-\lambda)\, Q(t\text{-}1),$$

Making some assumption about the form of Q* (as in (6b), above) yields an equation in which Q(t) is a function of Q(t-1) and the X's, with the coefficient of Q(t-1) providing an estimate of λ. Long-run elasticities are computed from the Q* function; short-run elasticities are smaller by the factor λ.

There are many other approaches to modeling dynamic adjustment, of course. Huntington (2010), for instance, hypothesizes that the choice of vehicles or other fixed assets is affected by the world oil price only when it rises above its previous maximum, P_{MAX}. He decomposes the actual world oil price, P, into two components: P_{MAX} and $[P - P_{MAX}]$. He argues that the world oil price is not affected by US demand, so that he can simply estimate an equation like (6b) directly, without worrying further about identification. (One could question that assumption, of course.) He finds that the first of these components has a stronger effect on demand than the second, yielding long-run elasticity estimates above the short-run estimates.

Demand Elasticity Estimates

There are literally hundreds of published studies that estimate price and income elasticities for various forms of energy in various places over various periods using data of various sorts and employing various statistical methods and modeling strategies.[5] Not surprisingly, the results of these studies vary substantially – probably a great deal more than the underlying behavior.

[5] If I is income and Q is quantity demanded, the income elasticity of demand is simply $(\partial Q/\partial I)(I/Q)$, where the partial derivative is used when demand has more than one determinant.

Price and income elasticities of demand are not physical constants, of course. Estimates of these quantities based on historical data are simply summaries of the behavior of particular human beings in particular times and places. Patterns of human behavior do not generally change suddenly and substantially, so well-constructed elasticity estimates can help predict responses to future price and income changes. But behavior does change.

Two general patterns emerge clearly from this large literature. First, long-run elasticities are larger than short-run elasticities, as the theory predicts, and often much larger. Given that long-run adjustment of energy demand generally involves changes in long-lived assets like cars and buildings, this is not a surprise. Second, short-run price and income elasticities tend to be small, generally well under 0.5. This, too, is not a great surprise: think about the major energy services your parents consume and think how easily it would be for them to decrease their consumption substantially in a short time in response to an energy price increase. Would they live in the dark? Turn the heat way down in the winter? Give up air conditioning? Cook less? Walk to work and to shop?

The demands for gasoline and residential electricity seem to have been the most studied: see Goodwin et al (2004) and Espey and Espey (2004) for comprehensive and fairly recent reviews. The table below summarizes one person's view of the ranges of "reasonable" estimates of price and income elasticities for these commodities in rich countries:

		Gasoline	Electricity
Price Elasticity	Short-Run	.15 – .25	.20 – .40
	Long-Run	.50 – .70	.50 – .80
Income Elasticity	Short-Run	.30 – .50	.15 – .30
	Long-Run	.60 – 1.10	.80 – 1.10

An important reason for the high long-run income elasticity for gasoline seems to be a tendency for people to buy larger and, generally, less fuel-efficient vehicles as they become wealthier. Home appliances presumably play a similar role in shaping long-run residential electricity demand.

Estimates of various other energy demand elasticities are scattered throughout the literature. In the US National Energy Modeling System (built and operated by the Energy Information Agency), for instance, the assumed short-run and long-run price elasticities of demand for natural gas are .14 and .40, respectively.

As noted above, behavior patterns do change – most often change gradually but sometimes abruptly – so that estimated demand elasticities must be used with caution. Hughes et al (2008) and others have argued that the price elasticity of demand for gasoline fell substantially after 1980, for instance. This may be a reaction to the fact that with lower oil prices and more efficient cars, the inflation-adjusted per-mile cost of driving and the share of gasoline in household budgets were both much lower in the 1980s and 1990s than in the 1960s and 1970s. Similarly, as we will discuss in class, the demand function for electricity appears to have shifted substantially in recent decades.

Challenges to the Economic Model

In recent years, the management consulting firm McKinsey & Company has joined a number of advocacy organizations and contended that firms and households in the U.S. (and in other nations) have failed to make many economic investments in energy efficiency. They find (McKinsey 2009) that investments of about \$520 billion would yield a net payoff of \$680 billion. McKinsey measures the net payoff as the net present value of energy savings, computed using a 7% discount rate, minus the initial investment cost.

The first reaction of many economists when assertions of this sort were first made in the 1970s was that the engineering calculations involved were too optimistic, particularly about the costs of making old buildings more energy-efficient. These criticisms had some validity, but most observers now seem convinced that there are in fact real examples of investments in energy efficiency that would be profitable at reasonable discount rates but are not made, though the magnitude of these opportunities is still controversial.

Hausman and Joskow (1982) offer a simple explanation for this failure to harvest attractive, low-hanging fruit: they point to a number of studies that showed that in making decisions regarding energy-efficient versus energy-inefficient appliances, consumers acted as if they had very high discount rates. Suppose, for instance, that spending an extra \$100 on a central air conditioner would reduce electricity cost by \$10.50 per year. Assuming, for simplicity, that the air conditioner lasts forever and that interest is compounded continuously,[i] the net present value of this investment using a 7% discount rate, as McKinsey did, would be

$$NPV_7 = 10.5/.07 - 100 = 150 - 100 = 50,$$

and this would appear to be an attractive investment. But if the household were using a 20% discount rate instead the net present value would be negative,

$$NPV_{20} = 10.5/.20 - 100 = 52.5 - 100 = -47.5,$$

indicating a very unattractive investment.

There are two basic problems with this explanation, however. First, while consumers who can't get loans might use a discount rate well above published loan rates, it is by no means obvious why most consumers would do so. It is perhaps even less obvious why businesses would behave as if they used very high discount rates, particularly large business that can access the capital markets and that have formal processes to evaluate possible investments.

Second, it is clear that at least some consumers do not treat all investments in energy savings in this way. About two million Toyota Prius cars have been sold world-wide, more than 800,000 in the US. The main difference between a Prius and a Toyota Corolla is that the Prius gets better mileage. As we will discuss in class, buying a Prius to save on gasoline makes sense only if one drives much more than average and applies a very low discount rate. But it is not rational for consumers to use two different discount rates to evaluate two different sorts of investment in energy efficiency. Either consumers are being irrational or the choice between a Prius and a Corolla involves more than comparing up-front costs and fuel costs – or both.

McKinsey (2009) offers some alternative explanations for society's apparent persistent under-investment in energy efficiency, and we will discuss them and other factors in later class sessions.

References

Espey, J.A. and M. Espey (2004), "Turning on the Lights: A Meta-Analysis of Residential Electricity Demand Elasticities," *Journal of Agricultural and Applied Economics*, 36 (April): 65-81.

Goodwin, P., J. Dargay, and M. Hanly (2004), "Elasticities of Road Traffic and Fuel Consumption with Respect to Price and Income: A Review," *Transport Reviews*, 24 (May): 275-292.

Hughes, J.E., C.R. Knittel, and D. Sperling (2008), "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand," *Energy Journal*, 29: 93-114.

Huntington, H. (2010), "Short- and Long-Run Adjustments in U.S. Petroleum Consumption," *Energy Economics*, 32: 63-72.

Kamerschen, D.R., and D.V. Porter (2004), "The Demand for Residential, Industrial, and Total Electricity," *Energy Economics*, 26: 87-100.

McKinsey & Company (2009), Unlocking Energy Efficiency in the U.S. Economy: Executive Summary, July. Available at http://www.mckinsey.com/clientservice/electricpowernaturalgas/us_energy_efficiency/

Schmalensee, R., and T.M. Stoker (1999), "Household Gasoline Demand in the United States," *Econometrica*, 67 (May): 645-62.

Smil, V. (2000), "Energy in the Twentieth Century: Resources, Conversions, Costs, Uses, and Consequences," *Annual Review of Energy and the Environment*, 25: 21-51.

---

[i] Suppose interest is compounded n times per year and the rate of interest is r per year. Then the value in t years of one dollar invested today is

$$[1 + (r/n)]^{nt}$$

Taking the limit as $n \to \infty$, this becomes exp(rt), and the present value of a dollar to be received t years from now is just exp(-rt). Suppose now that savings are S per year and that they occur uniformly during the year. The present value of savings from time zero to time T is given by

$$PV = \int_0^T Se^{-rt}dt = \frac{S}{r}[1 - e^{-rT}].$$

As $T \to \infty$, PV goes to S/r, which is used the analysis of air conditioner choise.

In the analysis of the decision to buy a Prius, if the cost difference between a Prius and a Corolla is C, the length of time the car must be driven in order for the gasoline savings to make up for the higher purchase price, the breakeven time, is found by setting PV = C in the equation above and solving for T. If r = 0, PV=ST, and the solution is just T=C/S. If rC > S, the equation has no finite solution: no matter how long the car is driven, the savings in gasoline cost will not make up for the higher purchase price of the Prius.