Estimating Short and Long-Run Demand Elasticities: A Primer with Energy-Sector Applications

Cuddington, John and Dagher, Leila

2013

Estimating Short and Long-Run Demand Elasticities:
A Primer with Energy-Sector Applications

John T. Cuddington*
and Leila Dagher**

January 28, 2014

ABSTRACT: Many empirical exercises estimating demand functions, whether in energy economics or other fields, are concerned with estimating dynamic effects of price and income changes over time. This paper first reviews a number of commonly used dynamic demand specifications to highlight the implausible a priori restrictions that they place on short and long-run elasticities. Such problems are easily avoided by adopting a general-to-specific modeling methodology. Second, it discusses functional forms and estimation issues for getting point estimates and associated standard errors for both short and long-run elasticities – key information that is missing from many published studies. Third, our proposed approach is illustrated using a dataset on Minnesota residential electricity demand.
Keywords: SR elasticity, LR elasticity, demand function, ADL
JEL Classifications: C13 (econometric and statistical methods, estimation), C51 (econometric modeling, model construction and estimation), Q41 (energy demand and supply)

Acknowledgements: We thank the referees and the editor for very helpful comments and suggestions that have greatly improved the paper. Part of this work was done while the second author was a visiting scholar at the Joint Institute for Strategic Energy Analysis at the National Renewable Energy Laboratory.

---

* William J. Coulter Professor of Mineral Economics, Colorado School of Mines, jcudding@mines.edu
** Assistant Professor, American University of Beirut, ld08@aub.edu.lb

Many empirical exercises estimating demand and supply functions are concerned with estimating dynamic effects of price and income changes over time.[1] Researchers are typically interested in estimating both short-run (SR) and long-run (LR) elasticities, along with their standard errors. Energy demand analysis offers many applications; see Dahl (1993) for a comprehensive survey of energy elasticity estimates. For example, consider a public utility requesting a rate increase from the public service commission. The utility and regulators want to know how a proposed price hike will impact demand in the SR and the LR. Searching the literature on energy demand elasticity estimates, one finds that authors often fail to provide standard errors for either their short or long-run elasticity estimates, or both.[2,3] Thus, it is hard to know whether the LR elasticities are statistically different from their SR counterparts. Moreover, it is difficult to determine whether elasticity estimates across studies are really statistically different from each other.

Section I of this paper first reviews a number of commonly used dynamic demand specifications and highlights the implausible a priori restrictions that they place on short and long-run elasticities. We emphasize that these restrictions are easily avoided, and are indeed testable, when a general-to-specific modeling approach is employed. Section II discusses estimation issues, including a simple way to get standard errors as well as point estimates for both short and long-run elasticities. Section III provides an empirical application – estimating residential demand for electricity in Minnesota. Section IV concludes.

I. Alternative Dynamic Demand Specifications and Implications for SR and LR Elasticities

Modern time-series econometricians emphasize the merits of beginning with a hypothesized data generating process (DGP) for all variables in the data sample being analyzed. In the case of sectoral supply and demand analysis, the DGP will typically involve a system of potentially simultaneous equations. As Andrew Harvey (1990, p. 2) has stressed: "Econometric models typically consist of sets of equations which incorporate feedback

[1] The analysis in this paper is equally relevant for estimating dynamic demand or dynamic supplies elasticities. For specificity, this paper focuses on demand estimation.

[2] Examples include Baughman and Joskow (1976, p.315, Table2), Chern and Just (1980, p. 40, Table 1), Chang and Hsing (1991, p.1255, Table 2), Arsenault et al. (1995, p.167, Table 3), Joutz and Trost (2007, p.5, Table ES2), and Huntington (2007, p. 755, Table 5).

[3] A referee comments that "Quite often these need to be dropped for publication constraints. Authors will mention in the text that they are available upon request, posted on a web-site as a working paper, or respond to requests when readers who ask. In fact, I would argue the opposite point. More often, firms and government agencies that use the published numbers and or obtain the additional information do two things. First they do not acknowledge the work of the researchers and or use the results incorrectly. This occurs frequently in public documents submitted before regulatory bodies and even in journals."

effects from one variable to another. Treating the estimation of a single equation from such a system as an exercise in multiple regression will, in general, lead to estimators with poor statistical properties."

In many cases, however, a system of equations can be reduced to a single equation. The assumptions needed to reduce the empirical analysis to a single-equation exercise (or so-called partial system) with no loss of information regarding the parameters of interest are often testable within a systems framework. Even when there is some loss of information, a limited information approach may have merits. According to Juselius (2006, p. 198): "Note, however, that in order to know whether we can estimate from a partial system we need first to estimate the full system and test in that system. But if we need to estimate the full system, why would we bother to discuss estimation in a partial system? Two reasons come to mind: (1) by conditioning on weakly exogenous variables, one can often achieve a partial system which has more stable parameters than the full system and (2) it is sometimes very likely a priori that weak exogeneity holds. In particular when the number of potentially relevant variables to include in the VAR model is large it can be useful to impose weak exogeneity restrictions from the outset."

Studies of energy demand elasticities have often used a single-equation DGP by assuming "that the particular market conditions of electrical and natural gas energy favor single equation analyses free from any endogeneity problem" (Balestra 1967; Uri 1975; Bohi 1981). The most common justification given is that the supply of electricity and natural gas may be considered perfectly elastic because supply is rarely, if ever, interrupted, and construction of pipeline and transmission and distribution lines are made with the purpose of satisfying not only immediate but also future consumption. As a result, most of the time there is excess capacity (Balestra 1967). Most studies implicitly assume that all regressors are (weakly) exogenous, so that these estimation approaches yield asymptotically valid statistical inference.

Suppose the empirical task at hand is to estimate a demand function for residential electricity demand (*q*) using time series data. For expositional simplicity, demand is assumed to depend only on own real price (*p*), the real price of substitutes (*ps*), and real income (*y*).[4,5]

$$q_t = \beta_0 + \beta_p p_t + \beta_s ps_t + \beta_y y_t + \varepsilon_t \tag{1.1}$$

[4] Other demand drivers include population or number of households and weather variables. The empirical example in Section III includes them.
[5] Juselius (2006, Chapter 19: "Specific-to-General and General-to-Specific") advocates the use of general-to-specific for model selection, but specific-to-general for variable inclusion: "[C]ontrary to the 'general to specific' approach of the statistical modeling process (i.e. of imposing more and more restrictions on the unrestricted VAR), it appeared more advantageous to follow the principle of 'specific to general' in the choice of information set." (p.11)

where all variables are in natural logs. Typically, this equation will have serially correlated errors, which is taken as prima facie evidence that dynamic considerations are important when modeling demand for many commodities. Here, four popular approaches for modeling the dynamics in order to estimate SR and LR price and income elasticities are considered. The first three are, in fact, nested as special cases of the general autoregressive distributed lag (ADL) model. After discussing these approaches, we emphasize the merits of a general-to-specific methodology.

Approach 1: Estimate the LR demand function with an AR(1) error process

$$\begin{aligned} &q_t = \beta_0 + \beta_p p_t + \beta_s ps_t + \beta_y y_t + \varepsilon_t \\ &where \\ &\varepsilon_t = \rho \varepsilon_{t-1} + u_t \end{aligned} \tag{1.2}$$

In this specification, the error term ε_t can be interpreted as the deviation of quantity demanded from the LR demand equation. The speed of adjustment toward the LR equilibrium is given by $1-\rho$.

To estimate the AR(1) model in (1.2), a generalized least squares (GLS) estimator is typically used. Alternatively, the long-run relationship is quasi-differenced to yield the following regression:

$$q_t = \beta_0(1-\rho) + \beta_p p_t + \beta_s ps_t + \beta_y y_t - \rho\beta_p p_{t-1} - \rho\beta_s ps_{t-1} - \rho\beta_y y_{t-1} + \rho q_{t-1} + u_t \tag{1.3}$$

This equation is then estimated using non-linear least squares (NLS) regression to obtain estimates of the LR price, cross-price, and income elasticities $(\beta_p, \beta_s, \beta_y)$ and other structural parameters (β_0, ρ) .[6]

Approach 2: Estimate a partial adjustment model (PAM)[7]

In the partial adjustment model (PAM), the long-run level of demand q^* is:

$$q_t^* = \beta_0 + \beta_p p_t + \beta_s ps_t + \beta_y y_t + \varepsilon_t \tag{1.4}$$

A partial adjustment mechanism describes how actual quantity q_t adjusts gradually towards q^* with speed of adjustment λ where $0 < \lambda < 1$:

$$q_t = q_{t-1} + \lambda(q_t^* - q_{t-1}) + u_t \tag{1.5}$$

[6] The iterative Cochrane-Orcutt method was used to estimate such equations before the advent of econometric software that can easily carry out NLS estimation.
[7] The PAM has also been called the stock adjustment model, the Koyck model, the lagged endogenous variable model, and the flow adjustment model.

Substituting (1.4) into (1.5) produces an equation that is nonlinear in the five structural parameters $(\beta_0, \beta_p, \beta_s, \beta_y, \lambda)$:

$$q_t = \lambda\beta_0 + \lambda\beta_p p_t + \lambda\beta_s ps_t + \lambda\beta_y y_t + (1-\lambda)q_{t-1} + \nu_t \tag{1.6}$$

Again, this specification is easily estimated using NLS regression, yielding both parameter estimates and their associated standard errors. With the PAM, however, authors typically just apply OLS regression involving the regressors in the (just-identified) equation above, and then reverse engineer the long-run elasticities. Getting their associated standard errors is tricky, however, so they are often not calculated or not reported.

Note that both the AR(1) and PAM specifications include the contemporaneous price p_t on the right-hand side, suggesting a need to use instrumental variables (IV) estimation to avoid the possibility of endogeneity bias.

Approach 3: Estimate an error correction model (ECM)

A single-equation error-correction model (as opposed to a *vector* error-correction system) is similar to the PAM except that the long-run demand q^* enters with a one-period lag:

$$\begin{aligned} \Delta q_t &= \lambda(q^*_{t-1} - q_{t-1}) + \varepsilon_t \\ &= \lambda(\beta_0 + \beta_p p_{t-1} + \beta_s ps_{t-1} + \beta_y y_{t-1} - q_{t-1}) + \varepsilon_t \end{aligned} \tag{1.7}$$

Rewriting (1.7) with the log-level of q rather than the log-difference as the dependent variable for comparability to the previous specifications yields:

$$q_t = \lambda\beta_0 + \lambda\beta_p p_{t-1} + \lambda\beta_s ps_{t-1} + \lambda\beta_y y_{t-1} + (1-\lambda)q_{t-1} + \varepsilon_t \tag{1.8}$$

The PAM and the ECM are quite similar: p_t, ps_t, and y_t enter the PAM specification, whereas the one-period lags, p_{t-1}, ps_{t-1}, and y_{t-1}, are included in the ECM. λ provides information about the speed of adjustment in both models.

The absence of the contemporaneous price among the regressors in the ECM is a restrictive a priori assumption when estimating demand or supply equations for most commodities. Even though we may presume that the very-short-run price elasticity is low, forcing it to be zero seems questionable with quarterly or annual frequency data, at least. Omitting contemporaneous price from the demand equation might seem to legitimize the use of OLS or non-linear LS rather than IV estimation, but it may merely replace the criticism of simultaneity bias with that of omitted variable bias. Moreover, consistency of the OLS estimator in ADLs typically requires the assumption of weak exogeneity of the regressors.[8]

The "simple" ECM in (1.7) assumes that the error term is serially uncorrelated. After discussing the ADL model below, more general ECMs with lagged differences of the regressors will be considered.

Approach 4: Estimate an autoregressive distributed lag (ADL) model

The autoregressive distributed lag or ADL*(L,R,V,S)* model regresses quantity demanded on *L* lags of itself, *R* lags of prices, *V* lags of cross prices, and *S* lags of income:

$$q_t = \gamma_0 + \sum_{l=1}^{L} \gamma_{ql} q_{t-l} + \sum_{r=0}^{R} \gamma_{pr} p_{t-r} + \sum_{v=0}^{V} \gamma_{sv} ps_{t-v} + \sum_{s=0}^{S} \gamma_{ys} y_{t-s} + u_t \tag{1.9}$$

Sometimes, ADL models include contemporaneous values of the additional regressors, as shown above. Other times, only lagged values are included. We'll allow contemporaneous (not just lagged) values of the explanatory variables to enter the demand equation for reasons just discussed above. To make the lag intervals explicit, we might label an ADL model as ADL*(1-L,0-R,0-V,0-S)*, this implies that lags 1 to *L*, 0 to *R*, 0 to *V*, and 0 to *S* of *q*, *p*, *ps*, and *y*, respectively, enter the equation.

Note that the ADL*(1-L,0-R,0-V,0-S)* model is the most general of the four specifications above. Indeed, even the ADL*(1,0-1,0-1,0-1)* model nests the other three specifications as special cases in the sense that all involve some subset of the following regressors: an intercept plus $(q_{t-1}, p_t, p_{t-1}, ps_t, ps_{t-1}, y_t, y_{t-1})$. Table 1 shows the (unrestricted) coefficients in the *ADL(1,0-1,0-1,0-1)* specification in column 1. The cells in the subsequent columns show the zero and nonlinear coefficient restrictions that the other models – AR(1), PAM, and ECM – impose on the ADL*(1,0-1,0-1,0-1)* specification. These restrictions may be expressed either in terms of the ADL coefficients or in terms of the underlying structural parameters in the respective models. Table 1 shows both.

[8] Strict exogeneity is sufficient for consistent estimation. Since our goal is to conduct valid conditional inference, however, we want to rely on the less restrictive condition that our variables are weakly exogenous (Engle et al., 1983).

Table 1: Dynamic Demand Specifications:
Restrictions on Estimated Parameters

	ADL(1,0-1,0-1,0-1)	*AR1 error*	*PAM*	*Simple ECM*
Intercept	γ_0	$\gamma_0 = (1-\rho)\beta_0$	$\gamma_0 = \lambda\beta_0$	$\gamma_0 = \lambda\beta_0$
p_t	γ_{p0}	$\gamma_{p0} = \beta_p$	$\gamma_{p0} = \lambda\beta_p$	$\gamma_{p0} = 0$
p_{t-1}	γ_{p1}	$\gamma_{p1} = -\rho\beta_p$ $\gamma_{p1} = -\gamma_{q1}\gamma_{p0}$	$\gamma_{p1} = 0$	$\gamma_{p1} = \lambda\beta_p$
ps_t	γ_{s0}	$\gamma_{s0} = \beta_s$	$\gamma_{s0} = \lambda\beta_s$	$\gamma_{s0} = 0$
ps_{t-1}	γ_{s1}	$\gamma_{s1} = -\rho\beta_s$ $\gamma_{s1} = -\gamma_{q1}\gamma_{s0}$	$\gamma_{s1} = 0$	$\gamma_{s1} = \lambda\beta_s$
y_t	γ_{y0}	$\gamma_{y0} = \beta_y$	$\gamma_{y0} = \lambda\beta_y$	$\gamma_{y0} = 0$
y_{t-1}	γ_{y1}	$\gamma_{y1} = -\rho\beta_y$ $\gamma_{y1} = -\gamma_{q1}\gamma_{y0}$	$\gamma_{y1} = 0$	$\gamma_{y1} = \lambda\beta_y$
q_{t-1}	γ_{q1}	$\gamma_{q1} = \rho$	$\gamma_{q1} = 1-\lambda$	$\gamma_{q1} = 1-\lambda$

Implausible A Priori Restrictions on Elasticities

One might think that the restrictions that the AR(1), PAM and ECM impose on the ADL*(1,0-1,0-1,0-1)* model are innocuous enough, and have the advantage of parsimony, especially when the basic models are extended to allow for additional regressors such as weather and seasonal dummies, etc. We show here, however, that all three specifications impose very implausible a priori restrictions on the relationships between short and long-run elasticities. These restrictions are easily avoided by estimating the general ADL specification using a general lag selection criterion such as the Akaike or Schwarz criterion. Given that the first three approaches are nested as special cases within the ADL model, we first calculate the elasticities in that model.

ADL*(1-L,0-R,0-V,0-S)* Model Elasticity Calculations

Let $\eta(q, p, k)$ equal the *cumulative* percentage response of *q* to a permanent percentage point change in *p after k periods*. These are so-called 'dynamic multipliers', or 'dynamic elasticities' when working with log-log specifications. The SR and LR price elasticities implied by the ADL model are easily calculated. The SR elasticity is simply the coefficient on the first price term:

$$\eta(q,p,0) \equiv \frac{dq_t}{dp_t} = \gamma_{p0} \tag{1.10}$$

Assuming stability, the LR elasticity can be found by first setting *q* in all time periods equal to $\bar{q}$ and all price terms equal to $\bar{p}$. Next calculate the total derivative:

$$\eta(q,p,\infty) \equiv \frac{d\bar{q}}{d\bar{p}} = \frac{\sum_{r=0}^{R} \gamma_{pr}}{1 - \sum_{l=1}^{L} \gamma_{ql}} \tag{1.11}$$

Note that stability of the demand function requires the denominator is positive:

$$1 - \sum_{l=1}^{L} \gamma_{ql} > 0. \tag{1.12}$$

The corresponding SR and LR *income* elasticities are:

$$\eta(q,y,0) \equiv \frac{dq_t}{dy_t} = \gamma_{y0} \tag{1.13}$$

$$\eta(q,y,\infty) \equiv \frac{d\bar{q}}{d\bar{y}} = \frac{\sum_{s=0}^{S} \gamma_{ys}}{1 - \sum_{l=1}^{L} \gamma_{ql}} \tag{1.14}$$

In general, the SR price elasticity may be bigger than, smaller than, or equal to its LR counterpart. Consider a special case, however, where only a single price term enters the ADL. In this case the short-run price elasticity must be less than the LR price elasticity (or equal to it, if there are no lags of the dependent variable):

$$\eta(q,p,0) = \gamma_{p0} \leq \eta(q,p,\infty) = \frac{\gamma_{p0}}{1 - \sum_{l=1}^{L} \gamma_{ql}} \tag{1.15}$$

Suppose, in addition, that there is just a single income term (and one or more lags of the dependent variable) in the ADL. Then the SR income elasticity must be less than the LR income elasticity:

$$\eta(q, y, 0) = \gamma_{y0} \leq \eta(q, y, \infty) = \frac{\gamma_{y0}}{1 - \sum_{l=1}^{L} \gamma_{ql}} \tag{1.16}$$

In general, it is unwise to impose a priori restrictions on the relative magnitudes of short-run vs. long-run price, income, or cross-price elasticities. Pindyck and Rubinfeld (2005, Ch. 2.5) have a useful discussion of relative magnitudes of short-run and long-run price and income elasticities: "For many goods [e.g. gasoline], demand is much more price elastic in the long run than in the short run…On the other hand, for some goods [e.g. automobiles and other durable goods] just the opposite is true – demand is more elastic in the short run than in the long run… Income elasticities also differ from the short run to the long run. For most goods and services, the income elasticity of demand is larger in the long run than in the short run. … For a durable good, the opposite is true." (pp. 39-40) For many mineral and energy products, the short-run income elasticity in the face of business cycle fluctuations, say, is presumed to be high, while the long-run income elasticity (reflecting trend growth in income) is generally considered to be near unity.

Note that if the ADL contains one price term and one income term, the *ratio* of the LR to SR elasticities is identical for both demand determinants:

$$\frac{\eta(q, p, \infty)}{\eta(q, p, 0)} = \frac{\eta(q, y, \infty)}{\eta(q, y, 0)} = \frac{1}{1 - \sum_{l=1}^{L} \gamma_{ql}} \tag{1.17}$$

The same 'ratio restriction' applies to the cross-price elasticity! This ratio restriction is an extremely implausible one to impose a priori when estimating demand functions. How can it be avoided? Here's our recommendation: *Make sure that each determinant of demand (e.g. price, cross-price, and income) enters with at least two time subscripts and that no non-linear restrictions are placed on their coefficients.* Typically lags 0 and 1 or lags 1 and 2 will be used. (The number of lags of the dependent variable does not matter.) Note that the ADL*(1,0-1,0-1,0-1)* specification is general enough to accomplish this, whereas the three special cases above are not. The same general rule holds when estimating dynamic *supply* functions. Each supply determinant (except the lagged dependent variable) should enter twice with different time subscripts in the supply equation.

We now calculate the SR and LR price elasticities implied by the other three specifications: the AR(1) model, PAM, and ECM.

AR(1) Model Elasticity Calculations

From (1.3), it is easy to compute the short-run and long-run own-price elasticities for the AR(1) error model:

$$\eta(q,p,0) \equiv \frac{dq_t}{dp_t} = \beta_p < 0$$

$$\eta(q,p,\infty) \equiv \frac{d\bar{q}}{d\bar{p}} = \frac{\beta_p(1-\rho)}{(1-\rho)} = \beta_p < 0 \tag{1.18}$$

Surprisingly, the AR(1) model imposes the a priori restriction that the SR and LR price elasticity must be equal to each other![9] The same feature holds for income (and cross-price elasticities) as well:

$$\eta(q,y,0) \equiv \frac{dq_t}{dy_t} = \beta_y$$

$$\eta(q,y,\infty) \equiv \frac{d\bar{q}}{d\bar{y}} = \frac{\beta_y(1-\rho)}{(1-\rho)} = \beta_y \tag{1.19}$$

These are hardly restrictions that one would want to impose a priori when estimating demand functions to compare the SR and LR effects of price or income changes. Indeed, Mizon (1995) argues that AR corrections for the *error process* are almost never appropriate, in part because of the COMFAC (common factor) restrictions that they impose on dynamic regression equations.

This result—the forced equality between SR and LR elasticities—also holds when the error term in (1.2) is a higher-order AR(p) error process. The resulting price elasticities in this case equal:

[9] Intermediate-term elasticities are also equal to the SR=LR elasticities, as one can confirm by simulating dynamic elasticities for this model.

$$\eta(q,p,0) \equiv \frac{dq_t}{dp_t} = \beta_p < 0$$

$$\eta(q,p,\infty) \equiv \frac{d\overline{q}}{d\overline{p}} = \frac{\beta_p(1-\rho_1-\rho_2-\rho_3-...)}{(1-\rho_1-\rho_2-\rho_3-...)} = \beta_p < 0 \tag{1.20}$$

The PAM Elasticity Calculations

As Table 1 shows, the partial adjustment model contains only a single price term and a single income term. Therefore, it follows immediately from the discussion of the ADL *(1,0-1,0-1,0-1)* model above that this specification also contains the implausible LR/SR ratio restriction:

$$\eta(q,p,0) \equiv \frac{dq_t}{dp_t} = \lambda\beta_p$$
$$\eta(q,p,\infty) \equiv \frac{d\overline{q}}{d\overline{p}} = \beta_p$$

$$\eta(q,y,0) \equiv \frac{dq_t}{dy_t} = \lambda\beta_y$$
$$\eta(q,y,\infty) \equiv \frac{d\overline{q}}{d\overline{y}} = \beta_y \tag{1.21}$$

Note that the partial adjustment coefficient should be between zero and one: $0 < \lambda < 1$. This being the case, *all* short-run elasticities (price, cross-price, and income) are necessarily less than their LR counterparts. As argued above, this is not a reasonable a priori restriction to impose on price, cross-price or income elasticities. These shortcomings of the PAM were highlighted by Fisher, Cootner, and Baily (1972). See also, Pei and Tilton (1999) and Chan and Lee (1997).

The ECM Elasticity Calculations

Comparing the ECM to PAM yields an interesting conclusion. Because (1.8) contains only lagged (not contemporaneous) regressors, the SR elasticities take effect with a one-period lag. Stated differently, the *very* SR elasticities are forced to equal zero by omission of the contemporaneous terms. The instantaneous, SR (i.e., one-period ahead), and LR price elasticities are:

$$\eta(q,p,0) \equiv \frac{dq_t}{dp_t} = 0$$
$$\eta(q,p,1) \equiv \frac{dq_t}{dp_{t-1}} = \gamma_{p1} = \lambda\beta_p \tag{1.22}$$
$$\eta(q,p,\infty) \equiv \frac{d\overline{q}}{d\overline{p}} = \frac{\gamma_{p1}}{1-\gamma_{q1}} = \beta_p$$

The instantaneous, one-period, and LR income elasticities are:

$$\eta(q,y,0) \equiv \frac{dq_t}{dy_t} = 0$$
$$\eta(q,y,1) \equiv \frac{dq_t}{dy_{t-1}} = \gamma_{y1} = \lambda\beta_y \tag{1.23}$$
$$\eta(q,y,\infty) \equiv \frac{d\overline{q}}{d\overline{y}} = \frac{\gamma_{y1}}{1-\gamma_{q1}} = \beta_y$$

As there is only a single lag term for both price and income, the undesirable features of the PAM also appear in the ECM. Namely, SR elasticities are forced to be less that their LR counterparts for *all* (own-price, cross-price, and income) elasticities. Moreover, the *ratio* of LR/SR elasticities is equal for all elasticities (e.g. price, cross-price, income, etc.).

How can we estimate SR and LR price and income elasticities of demand in a way that imposes no a priori restrictions on their relative magnitudes? The practical implication of the foregoing calculations bears repeating: *Estimate an ADL model using a standard lag selection criterion (e.g., the Schwarz or Akaike information criterion) but be sure to allow for at least two terms involving p, ps, and y; the number of lags of the dependent variable (q) does not matter.* More generally, in specifying dynamic regression equations, it is essential to adopt a general-to-specific modeling methodology -- beginning with sufficient lags of each variable to insure that serial correlation has been expunged from the error process. Avoid the use of AR (or ARMA) error processes in multivariate regression models, per Mizon's (1995) recommendation.

More General ECMs with lagged differences

Oftentimes the simple ECM specification above is generalized to include a number of lagged differences of the various regressors in order to 'mop up' serial correlation in the error process. Do the implausible restrictions on the ECM elasticities still hold if the ECM contains lagged differences as recommended by general-to-specific methodology? The answer is "no." Adding lagged differences to the ECM is a simple way to eliminate the implausible a priori restrictions! Indeed, every ADL – whether it includes contemporaneous regressors or not – can be *always* rewritten as an equivalent ECM.[10] The two are isomorphic, but the ECM parameterization is especially useful when cointegrated *I(1)* variables are involved. Moreover, the ECM variant of the ADL is particularly useful, because it allows us to identify the SR and LR elasticities directly. When the SR and LR elasticities in this equation are estimated with a consistent estimation approach, the associated standard errors for both are obtained automatically – that is to say, without auxiliary 'hand' calculations.

Panel Data Estimation

Most panel data sets used in energy demand studies are short panel data sets. Hence to be able to estimate SR versus LR elasticities most of these studies simply add a lagged dependent variable to the regression. The resulting model is thus the PAM, which suffers from the same weaknesses noted above. Alberini et al. (2011) provide a thorough exposition of the PAM in a panel framework. See also Bernard et al. (2011) and Garcia-Cerrutti (2000).

Recently, new developments in dynamic panel techniques have allowed some researchers to apply panel ADL models. See Baltagi (2008) for a good discussion of the panel ADL. However, to our knowledge no energy demand studies have used these techniques.

A General Functional Form for Estimating SR and LR Demand Elasticities

The estimated SR price elasticity and its associated standard error can be obtained directly from the point estimates on the first price term in the ADL*(1-L,0-R)* model in (1.24), $\hat{\gamma}_{p0}$, and its standard error:

$$q_t = \gamma_0 + \sum_{l=1}^{L} \gamma_{ql} q_{t-l} + \sum_{r=0}^{R} \gamma_{pr} p_{t-r} + u_t \tag{1.24}$$

To simplify the math in what follows (without impacting the generality of the results), we ignore the distributed lags on income and the price of substitutes.

Is there a simple way to estimate *both* the SR and LR elasticities and their standard errors directly? The answer is yes and involves the use of the isomorphic ECM. First, consider the ADL*(1-L,0-R)* where there is a

[10] David Hendry elaborates on this in a number of papers. See, e.g., Hendry (2008, Section 2.3): ADLs as Equilibrium-Correction Models.

contemporaneous (unlagged) price term. As Appendix I shows, the ADL can always be written in canonical form as an ECM:

$$\Delta q_t = \gamma_0 + \lambda(q_{t-1} - \beta_p p_{t-1}) + \gamma_{p0}\Delta p_t - \sum_{r=1}^{R-1} \gamma_{p,r+1}\Delta p_{t-r} - \sum_{l=1}^{L-1} \gamma_{q,l+1}\Delta q_{t-l} + u_t \tag{1.25}$$

where

$$\lambda \equiv \gamma_{q1} + \gamma_{q2} + \ldots + \gamma_{qL} - 1 < 0 \tag{1.26}$$

and

$$\beta_p \equiv \frac{\gamma_{p0} + \gamma_{p1} + \ldots + \gamma_{pR}}{1 - \gamma_{q1} - \gamma_{q2} - \ldots - \gamma_{qL}} \tag{1.27}$$

Note that the total number of lags of the dependent variable drops by one when the ADL in levels is rewritten in canonical form. This re-parameterization is particularly useful because the coefficient on Δp_t is the SR price elasticity and the coefficient on p_{t-1} in the error-correction term is the LR elasticity shown in (1.27). The parameters in the ECM, along with their associated standard errors, can be estimated directly using NLS.

For an ADL*(1-L,1-R)* that does *not* contain a contemporaneous price term:

$$q_t = \gamma_0 + \sum_{l=1}^{L} \gamma_{ql} q_{t-l} + \sum_{r=1}^{R} \gamma_{pr} p_{t-r} + u_t, \tag{1.28}$$

the canonical form needed to identify the SR and LR price elasticities is slightly different:

$$\Delta q_t = \gamma_0 + \lambda(q_{t-2} - \beta_p p_{t-2}) + \gamma_{p1}\Delta p_{t-1} - \sum_{r=2}^{R-1} \gamma_{p,r+1}\Delta p_{t-r} + (\gamma_{q1} - 1)\Delta q_{t-1} - \sum_{l=2}^{L-1} \gamma_{q,l+1}\Delta q_{t-l} + u_t. \tag{1.29}$$

The coefficient on Δp_{t-1} is the SR price elasticity and the coefficient on p_{t-2} in the error-correction term is the LR elasticity analogous to (1.27), but with $\gamma_{p0} = 0$.

Here's a summary of the procedure.

1. Estimate an ADL*(1-L,0-R,0-V,0-S)* model using a standard lag selection method such as AIC, SC, or sequential likelihood ratio tests. Selection of lag lengths *(L,R,V,S)* can be based on a grid search using the Schwarz or Akaike information criterion. Note that there are many, many cases to consider *(*i.e. the product *L*R*V*S)*. Thus many algorithms only consider symmetric cases where *L=R=V=S*, as in the standard VAR approach. This drastically reduces the number of regressions to be considered. We adopt this simplification below. Be wary about lag choices that involve only a single term of a given demand determinant, as this will impose the magnitude restriction on SR and LR elasticities: SR < LR.

2. Re-estimate the chosen ADL*(1-L,0-R,0-V,0-S)* model as an equivalent ECM with *L-1, R-1, V-1* and *S-1* lagged differences of *(q, p, ps, y)*. The standard error of regression for the ADL and ECM should be *identical*, as the specifications are isomorphic.
3. The LR elasticities and their standard errors are read directly from the coefficients in the error correction term of the ECM, while the SR elasticities and their standard errors are read from the coefficients of the first difference terms in the ECM.

II. Estimating SR and LR Elasticities and their Standard Errors

There are several important factors that determine the appropriate estimation technique for an ADL: (i) the presence or absence of contemporaneous terms where simultaneity bias may be an issue, (ii) the stationarity of variables, and (iii) the presence or absence of a common stochastic trend (cointegration) among non-stationary variables. We briefly examine the different cases below.

Case 1: Only *lags* of the independent variables (not contemporaneous values) appear in the ADL and all variables are stationary.

This case is not commonly encountered in energy demand analysis, because non-stationary regressors are typically present.[11] Nonetheless, we briefly discuss how to proceed when all variables appear to be *I(0)*. The omission of contemporaneous prices is quite widespread. It is often justified in analyses of residential demand for electricity or natural gas, for example, by arguing that most end-users only learn of a price change when they receive their monthly bill, i.e. typically a month after the price change occurs. See, e.g., Munley et al. (1990), Joutz and Trost (2007), and Dagher (2011). This becomes less plausible when working with lower frequency data. Moreover, one can test this hypothesis, rather than making an a priori assumption that it holds.

If the ADL contains only stationary and lagged regressors (and the model is correctly specified in the sense that there are no omitted variables and the functional form is correct), then it is well known that OLS estimates of the dynamic demand coefficients are consistent. If the White robust coefficient covariance matrix is used when carrying out hypothesis tests, *F* and *t* tests will have their standard asymptotic distributions. See Stock and Watson (2007, Ch. 14) for details.

[11] Indeed, Hendry and Juselius (2000 part I, p.2) make a more general claim: "It seems clear that stationarity assumptions must be jettisoned for most observable economic time series."

Case 2: The ADL contains contemporaneous as well as lagged regressors; all variables are stationary.

As mentioned above, one rarely encounters an energy demand regression in which all variables are stationary. If the ADL contains stationary, contemporaneous as well as lagged regressors, the regressors must be weakly exogenous for OLS estimates to be consistent and conditional inference to be valid.[12]

While it may be reasonable to assume that this condition holds for the contemporaneous income and cross-price terms in a dynamic demand specification, it is typically not reasonable for own-price effects. To overcome simultaneity bias, one needs to use an instrumental variables estimator such as two stage least squares or generalized methods of moments.[13]

Case 3: Some variables are nonstationary, but are not cointegrated[14]

This section discusses the validity of OLS estimation of the ADL*(L,R,V,S)* model and the equivalent ECM*(L-1, R-1, V-1, S-1)* when some but not necessarily all of the variables are *I(1)*, but not cointegrated. Note that, if there is no long-run equilibrium relationship among the variables, then any empirical quest to estimate LR elasticities is misguided! The researcher might, nonetheless, be interested in SR elasticities. An appropriate approach would be to estimate an ADL using first-differenced (i.e., stationary) variables simply by differencing the non-stationary variables and then estimating the resulting regression equation.

We have assumed that all independent variables are weakly exogenous. As usual, endogenous regressors require that instrumental variables estimation be used.

Case 4: Variables are non-stationary and cointegrated.

We now turn to examining the ADL model where the *I(1)* variables are cointegrated.[15] Here the validity of OLS is less sensitive to the presence of contemporary (unlagged) first-differences of potentially endogenous variables, notably the own-price effects. In case of cointegration, the cointegrating vector can be estimated

[12] It is well-known that OLS estimation of the coefficient on an endogenous variable (i.e. one that is correlated with the error term) results in biased and inconsistent parameter estimates.
[13] A recent study that uses only stationary variables but allows for contemporaneous effects is Huntington's (2007) paper on industrial natural gas consumption in the USA. He finds all variables (i.e. the price of natural gas, the price of distillate fuel oil, heating degree days, structural output, and capacity utilization) are stationary. OLS regressions that allow for contemporaneous effects are used in order to estimate SR and LR demand elasticities. Yet there are no tests of the weak exogeneity condition needed to validate parameter estimation and inference using OLS.
[14] If only the regressand and one of the regressors are *I(1)* while all other regressors are *I(0)*, then all of the estimated coefficients will still have their standard asymptotic distributions (West, 1988; Hamilton, 1994, p.555).
[15] Chan and Lee (1997) consider such a case. Using annual data for the years 1953 to 1990, they find that their regression variables are all *I(1)* and cointegrated. They use both an ECM and an ADL model to estimate short-run and long-run demand elasticities for coal in China, but they restrict all variables to have the same number of lags and they go up to a maximum of three lags only. Moreover, they do not provide standard errors for their long-run estimates. The methodology proposed here provides a simple way to obtain SR and LR elasticity estimates with their standard errors in similar cases without imposing the above-mentioned restrictions.

consistently[16] by running an OLS regression in levels – the first-stage of the Engle-Granger two-step procedure -- without the need for IV estimation, even if some or all of the variables are not strictly exogenous (Stock and Watson, 1988; Diebold and Nerlove, 1990; Hamilton, 1994 p.588). This property should not be affected with the addition of lagged variables in an ADL, since if X_t and Y_t are cointegrated, then X_t and Y_{t-i} will also be cointegrated because Y_t and Y_{t-i} are cointegrated for all i (Cuthbertson et al., 1992 p.133). Hence, when estimating an ADL that includes both the contemporaneous terms of the cointegrated variables and their lags, OLS should give consistent estimates. Pesaran and Shin (1999) show that OLS estimation of ADL model coefficients yields consistent short-run parameter estimates and super-consistent long-run parameter estimates, provided there exists a *unique* cointegrating relationship among these variables in the ADL. Moreover, valid inferences can be made using the standard distributions if the variables are weakly exogenous. One might wonder: how can standard asymptotical distributions apply for the estimated long-run relationships? Doesn't superconsistency imply nonstandard sampling distributions, even asymptotically? Pesaran and Shin (1999, pp. 381-389) show that when written in ADL form, the coefficients have a 'mixture normal distribution asymptotically and standard inferences are therefore asymptotically valid.' See also Hamilton (1994, p. 602), Sims, Stock and Watson (1990), or Watson (1994). If, however, there is more than one long-run relationship among the *I(1)* variables in the ADL model, then the estimated coefficients might be a linear combination of the true underlying parameters. Hence, the desirable properties of the OLS estimator described no longer hold.

Turning now to the equivalent ECM equation, if *q, p, ps,* and *y* are *I(1)* and cointegrated, then all terms in the ECM are *I(0).* That is, the lagged differences are *I(0)* and the error correction term $q_{t-1} - \beta_0 - \beta_p p_{t-1} - \beta_s ps_{t-1} - \beta_y y_{t-1}$ is *I(0)* as well. Hence, if *q, p, ps, y* are *I(1)* but cointegrated, the error term will be stationary and standard OLS is valid, as was the case when all three variables are *I(0).* If the variables are not cointegrated, however, the coefficients of the error correction term will have nonstandard distributions and hence the usual methods of statistical inference will be invalid. In an ECM it does not matter if some of the variables are endogenous, because no contemporaneous terms appear in the equation.

[16] The OLS estimates in the case of cointegration are said to be 'super consistent' because their order of convergence is T and not $T^{0.5}$ (Stock, 1987; Engle and Granger, 1987).

Rather than beginning with a single-equation error-correction model, we estimate a vector-error correction model, and then test whether this system can be reduced to a single-equation conditional ECM with no loss of information. This null hypothesis is not rejected in our application.

III. Empirical Application

This section illustrates the approach to estimating short-run and long-run price and income elasticities discussed in the previous sections. We estimate residential electricity demand over the period 1998:1 to 2006:12 in Xcel Energy's Minnesota service area. Electricity demand, Q_e, is posited to be a function of the real price of electricity P_e, the real price of natural gas P_g, real income Y, the number of customers N, cooling degree days *CDD*, heating degree days *HDD*, and monthly dummy variables. All variables, except *CDD* and *HDD* and monthly dummies, are in logs.

The order of integration of each series is considered first. The ADF test results presented in Table 2 suggest that, at the 5% level of significance, the null hypothesis of (one or more) unit roots is not rejected for any of our series Q_e, P_e, P_g, Y, N, *CDD*, and *HDD.* After first differencing, however, all become stationary. We conclude that all of the stochastic series are *I(1).* The dummies, on the other hand, are *I(0).*

Table 2: ADF Tests on (Log) Levels and First Differences

Series	ADF t-Stat	P-Value	AIC Lag Choice
Q_e	-1.03	0.74	12
P_e	-0.75	0.83	11
P_g	-2.05	0.27	12
Y	-0.94	0.77	2
N	-0.66	0.85	1
CDD	-1.95	0.31	12
HDD	-2.09	0.25	12
$D(Q_e)$	-6.29	0.00	11
$D(P_e)$	-3.64	0.01	12
$D(P_g)$	-5.90	0.00	8
D(Y)	-6.51	0.00	1
D(N)	-8.59	0.00	0

D(CDD)	-13.59	0.00	10
D(HDD)	-5.94	0.00	11

Are the variables above cointegrated? Using the Johansen trace and maximum eigenvalue tests, the null hypothesis of no cointegration is easily rejected.[17] Both the trace and maximum eigenvalue tests suggest one cointegrating relationship at the 1% level. However, at the 5% level, the trace test implies that there are two cointegrating relations, while the maximum eigenvalue test still suggests one cointegrating relationship. See Table 3 for the test results. Based on economic theory, it seems reasonable to expect a single long-run relationship linking consumption, prices, income, and number of customers. Note that the monthly dummies and the weather variables have been treated as exogenous variables in the test.[18]

Table 3: Johansen Cointegration Test Results
Sample: 1998M01 2006M12
Included observations: 106
Series: Q_e, P_e, P_g, N, Y
Exogenous series: S1 S2 S3 S5 S6 S7 S8 S9 S10 S11 S12 CDD HDD
Lags interval: 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.01 Critical Value	Prob.**
None *	0.77	211.37	77.82	0.00
At most 1	0.23	53.50	54.68	0.01
At most 2	0.17	26.22	35.46	0.12

Trace test indicates 1 cointegrating eqn(s) at the 0.01 level

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.01 Critical Value	Prob.**
None *	0.77	157.87	39.37	0.00
At most 1	0.23	27.29	32.72	0.05
At most 2	0.17	19.91	25.86	0.07

[17] Note that throughout this Section, we limit ourselves to cases where all variables have the same number of terms. This dramatically reduces the number of cases to consider and matches the standard approach in the literature.
[18] Some variables are very likely to be weakly exogenous to a system, such as the monthly dummies and weather variables in our case. Juselius (2006 p.198) suggests that testing might not be necessary in those cases and a partial system can thus be used from the outset.

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.01 level
* denotes rejection of the hypothesis at the 0.01 level
**MacKinnon-Haug-Michelis (1999) p-values

In the unrestricted VAR setting, the AIC selected a specification with 2 lags, implying a VECM of order 1. The estimates are reported in Appendix II. (The arbitrary normalization sets Q_e equal to unity). We also carried out weak exogeneity tests with respect to the LR coefficients by testing the statistical significance of the loading factors or speed-of-adjustment coefficients on deviations from the long-run equilibrium conditions in the VECM. The results suggest that P_e, P_g, Y, and N are all weakly exogenous for the long-run elasticities. Thus it is possible to estimate a *fully-efficient* conditional demand function that includes contemporaneous values of the changes in all weakly exogenous variables using OLS estimation. If P_e had turned out to be weakly endogenous, any attempt to estimate the conditional model with OLS would be plagued by endogeneity bias. The full model would have to be estimated in that case.

Given that all variables except $D(Q_e)$ appear to be weakly exogenous in the VECM, we proceed in the text by estimating a demand equation that includes contemporaneous variables when estimating the short-run elasticities. Both an ECM and its equivalent ADL model were estimated. As discussed in Section II Case 4 (where there is a unique cointegrating relationship), the OLS estimates will be consistent and hypothesis testing can be based on standard distributions. This is the specification that minimizes the Akaike criterion when the maximum lag length is set at six and the minimum lag is zero. Both the SR elasticities and the LR elasticities, along with their respective standard errors, can be read directly from the ECM estimation output.

Table 4: OLS Estimates of the Equivalent ECM

Cointegrating Eq.	Coefficient	Std. Error	Prob.
$Q_e(-1)$	1.000		
$P_e(-1)$	-0.61	1.04	0.56
$P_g(-1)$	0.26	0.29	0.37
Y(-1)	-0.35	3.16	0.91
N(-1)	0.41	5.73	0.94
Error Correction	**Coefficient**	**Std. Error**	**Prob.**
Cointegrating Eq.	0.13	0.09	0.19
C	0.84	6.83	0.90

$D(Q_e(-1))$	-0.13	0.08	0.13
$D(Q_e(-2))$	0.01	0.08	0.89
$D(Q_e(-3))$	0.11	0.06	0.05
$D(P_e)$	-0.16	0.08	0.05
$D(P_e(-1))$	-0.04	0.11	0.70
$D(P_e(-2))$	-0.06	0.09	0.52
$D(P_e(-3))$	-0.02	0.08	0.85
$D(P_g)$	0.01	0.03	0.76
$D(P_g(-1))$	0.02	0.04	0.57
$D(P_g(-2))$	-0.07	0.04	0.06
$D(P_g(-3))$	-0.07	0.04	0.04
D(Y)	-1.17	1.12	0.29
D(Y(-1))	4.50	1.21	0.00
D(Y(-2))	-1.63	1.25	0.19
D(Y(-3))	-0.52	1.13	0.19
D(N)	1.12	1.97	0.57
D(N(-1))	-3.56	2.06	0.09
D(N(-2))	-1.23	1.86	0.51
D(N(-3))	2.60	1.84	0.16
Adjusted R-squared	0.98		
S.E. of regression	0.03		

The results for the *isomorphic* ADL model are omitted here to save space.

Table 5 reports the results for the simple AR1 and PAM specifications so that they can be compared to our chosen conditional ECM model. The ECM has a higher adjusted R-squared and a lower standard error than the AR(1) and PAM. It is also the specification that minimizes the Akaike value.

Table 5: OLS Estimates of the AR(1) and PAM Specifications

Eq Name: *Dep. Var:*	AR(1) Q_e	PAM Q_e
C	-19.33	-2.67
	(12.39)	(5.23)
P_e	-0.07	-0.04
	(0.11)	(0.07)
P_g	0.06	0.00
	(0.04)	(0.02)
Y	-0.47	0.04
	(0.75)	(0.32)
N	2.57	0.37

	(1.26)	(0.56)
D(CDD)	0.00 (0.00)	0.00 (0.000)
D(HDD)	0.00 (0.00)	0.00 (0.000)
$Q_e(-1)$		0.77 (0.07)
AR(1)	0.53 (0.09)	
Observations	106	107
Adj. R-squared	0.95	0.97
S.E. regression	0.04	0.03

Table 6 summarizes the estimated SR and LR elasticities from the AR1, PAM, and ADL/ECM specifications. As noted in Section 1, the AR1 specification imposes the restriction that each SR elasticity be equal to its LR counterpart, while the PAM specification imposes the implausible 'ratio restriction.' We can see from the Table that the ratio of the LR elasticity to the SR elasticity for the three determinants of demand is equal to (4.5) in all cases. The specification with the minimum Akaike criterion is neither an AR1 nor a PAM, but a more general ADL*(4,0-4,0-4,0-4,0-4)* model.

Table 6: Summary of Short-Run and Long-Run Elasticity Estimates from the Various Dynamic Demand Specifications

	$\varepsilon_{pe,sr}$ SR own-price elasticity	$\varepsilon_{pe,lr}$ LR own-price elasticity	$\varepsilon_{cp,sr}$ SR cross-price elasticity	$\varepsilon_{cp,lr}$ LR cross-price elasticity	$\varepsilon_{y,sr}$ SR income elasticity	$\varepsilon_{y,lr}$ LR income elasticity	Akaike Criterion
AR1	-0.07 (0.11)	-0.07 (0.11)	0.06 (0.04)	0.06 (0.04)	-0.47 (0.75)	-0.47 (0.75)	-3.27
PAM	-0.04 (0.06)	-0.19 (0.29)	0.00 (0.02)	0.01 (0.09)	0.04 (0.32)	0.18 (1.37)	-3.89
ECM/ ADL	-0.16 (0.08)	-0.60 (1.04)	0.01 (0.03)	0.26 (0.29)	-1.18 (1.12)	-0.39 (3.14)	-3.97

Standard errors in parentheses

IV. Conclusions

A review of the literature on energy demand reveals two common deficiencies: (i) the omission of standard errors when reporting short-run and especially long-run elasticities and (ii) the use of restricted models without testing the relevant restrictions. This paper first reviews a number of commonly used dynamic demand specifications and highlights the implausible a priori restrictions that they place on short and long-run elasticities. It then shows which specifications do not impose any restrictions. The discussion suggests that the ADL or corresponding ECM should be employed in practice, rather than using the AR or PAM specifications. Second, we propose a simple way to get standard errors as well as point estimates for *both* short and long-run elasticities. Our approach is illustrated using data on Minnesota residential electricity demand.

Although our focus is on demand estimation in the energy sector, the issues raised are also relevant when estimating dynamic supply equations. Moreover, they apply in a wide variety of contexts beyond the energy sector where estimating short and long-run elasticities is a recurring topic of interest.

References

Alberini, A., Gans, W., and Velez-Lopez, D. (2011). "Residential Consumption of Gas and Electricity in the U.S.: The Role of Prices and Income." *Energy Economics* 33: 870-881.

Arsenault, E., J.T. Bernard, C.W. Carr, and E. Genest-Laplante (1995). "A Total Energy Demand Model of Québec." *Energy Economics* 17(2): 163-171.

Baltagi, B. H. (2008). *Econometrics*, 4th edition. Heidelberg: Springer-Verlag.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*, 4th edition. Chichester: John Wiley and Sons.

Baughman, M. L. and P.L. Joskow (1976). "Energy Consumption and Fuel Choice by Residential and Commercial Consumers in the United States." *Energy Systems and Policy* 1(4): 305-323.

Balestra, P. (1967). *The Demand for Natural Gas in the United States*. Amsterdam: North Holland Publishing Co.

Bernard, J. T., Bolduc, D, and Yameogo, N. D. (2011). "A Pseudo-panel Data Model of Household Electricity Demand." *Resource and Energy Economics* 33: 315-325.

Bohi, D. R. (1981). *Analyzing Demand Behavior*. Baltimore, MD: Johns Hopkins University. Published for Resources for the Future.

Chan, H. L. and S. K. Lee (1997). "Modeling and Forecasting the Demand for Coal in China." *Energy Economics* 19: 271-287.

Chang, H. S. and Y. Hsing (1991). "The Demand for Residential Electricity: New Evidence on Time-Varying Elasticities." *Applied Economics* 23: 1251-1256.

Charemza, W. and D. Deadman (1997). *New Directions in Econometric Practice*, 2nd Edition. UK: Edward Elgar Publishing.

Chern, W. S. and R.E. Just (1980). "Regional Analysis of Electricity Demand Growth." *Energy* 5: 35-46.

Cuthbertson, K., S. Hall, and M. Taylor (1992). *Applied Econometric Techniques.* Ann Arbor: The University of Michigan Press.

Dagher, L. (2011). " Natural Gas Demand at the Utility Level: An Application of Dynamic Elasticities." *Energy Economics* 34: 961-969.

Dahl, C. A. (1993). "A Survey of Energy Demand Elasticities in Support of the Development of the NEMS." Prepared for the US Department of Energy under contract De-Ap01-93EI23499. Colorado School of Mines.

Diebold, F.X. and M. Nerlove (1990). "Unit Roots in Economic Time Series: A Selective Survey." In G.F. Rhodes and T.B. Fomby, eds., *Advances in Econometrics*, Volume 8: 3-69. Greenwich, CT: JAI Press.

Engle, R.F., and C.W.J. Granger (1987). "Co-integration and Error Corrections Representation, Estimation, and Testing." *Econometrica* 55: 251-76.

Engle, R.F., D.F. Hendry, and J-F Richard (1983). "Exogeneity." *Econometrica* 51 (2): 277-304.

Fisher, F.M., P. H. Cootner, and M.N. Baily (1972). "An Econometric Model of the World Copper Industry." *Bell Journal of Economics and Management Science* 3 (2): 568–609.

Garcia-Cerrutti, L. M. (2000). "Estimating Elasticities of Residential Energy Demand from Panel County Data using Dynamic Random Variable Models with Heteroskedastic and Correlated Error Terms." *Resource and Energy Economics* 22: 355-366.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Harvey, A. (1990). *Econometric Analysis of Time Series*, 2nd edition. Cambridge: MIT Press.

Hendry, D. F. (2008). “Equilibrium-Correction Models.” In Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics,* 2nd Edition in Eight Volumes. Hampshire: Palgrave MacMillan.
Hendry, D., K. Juselius (2000). “Explaining Cointegration Analysis: Part I.” *Energy Journal* 21: 1-42.
Huntington, H. G. (2007). “Industrial Natural Gas Consumption in the United States: An Empirical Model for Evaluating Future Trends.” *Energy Economics* 29: 743-759.
Joutz, F., and R.P. Trost (2007). “An Economic Analysis of Consumer Response to Natural Gas Prices.” American Gas Association.
Juselius, K. (2006). *The Cointegrated VAR Model: Methodology and Applications*. Oxford, UK: Oxford University Press.
Mizon, G. E. (1995). "A Simple Message for Autocorrelation Correctors: Don't." *Journal of Econometrics* 69: 267-288.
Munley, V. G., L. W Taylor, and J. P. Formby (1990). “Electricity Demand in Multi-family, Renter-occupied Residences.” *Southern Economic Journal* 57(1): 178-194.
Pei F., and J.E. Tilton (1999). “Consumer Preferences, Technological Change, and the Short-run Income Elasticity of Metal Demand.” *Resources Policy* 25: 87-109.
Pesaran, M. H., and Y. Shin (1999). “An Autoregressive Distributed Lag Modeling Approach to Cointegration Analysis.” In S. Strom, ed., *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*. Cambridge: Cambridge University Press.
Pindyck, R. S. and D. L. Rubinfeld (2005). *Micrcoeconomics*, 6th edition. Essex: Pearson Education.
Sims, C., J. Stock, and M. Watson (1990). “Inference in Linear Time Series Models with Some Unit Roots.” *Econometrica* 58: 113-144.
Stock, J. H. (1987). “Asymptotic Properties of Least Squares Estimators of Co-integrating Vectors.” *Econometrica* 55: 1035-56.
Stock, J. H., and M. W. Watson (1988). “Variable Trends in Economic Time Series.” *Journal of Economic Perspectives* 2 (3): 147-174.
Stock, J. H., and M. W. Watson (2007). *Introduction to Econometrics*, 2nd edition. Boston, MA: Pearson Education.
Urbain, J. P. (1992). “On weak exogeneity in error correction models.” *Oxford Bulletin of Economics and Statistics* 54: 187-207.
Uri, N. D. (1975). *Towards an Efficient Allocation of Electrical Energy*. (pp. 11-26). Lexington, MA: Heath.
Watson, M. W. (1994). “Vector Autoregressions and Cointegration.” In R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics*, Volume IV: 2843-2915. Amsterdam: Elsevier Science.
West, K. D. (1988). “Asymptotic Normality When Regressors Have a Unit Root.” *Econometrica* 56: 1397-1417.

Appendix I: ADL and Equivalent ECM

Here we prove that for every ADL*(L,R)* specification[19]

$$q_t = \gamma_0 + \sum_{l=1}^{L} \gamma_{ql} q_{t-l} + \sum_{r=0}^{R} \gamma_{pr} p_{t-r} + u_t$$

there is an equivalent ECM with lagged differences:

$$\Delta q_t = \delta_0 + \lambda(q_{t-1} - \beta_p p_{t-1}) + \sum_{l=1}^{L-1} \delta_{ql} \Delta q_{t-l} + \sum_{r=0}^{R-1} \delta_{pr} \Delta p_{t-r} + u_t$$

The coefficients in the error correction term are equal to the long-run elasticities:

$$\beta_p = \frac{\sum_{r=0}^{R} \gamma_{pr}}{1 - \sum_{l=1}^{L} \gamma_{ql}}$$

and the coefficient on the first difference term, Δp_t, is equal to the SR elasticity.

[19] We start with a general specification that includes contemporaneous and lagged regressors and later present the analogous derivation for a specification that includes only lagged regressors.

$$\delta_{p0} = \gamma_{p0}$$

It is interesting to note that an ADL model can always be rewritten as a single-equation ECM that includes lagged differences to capture serial correlation that may exist in the no-lag version. This is discussed below.

1. ADL*(1,1)* case:[20]

$$q_t = \gamma_0 + \gamma_{q1} q_{t-1} + \gamma_{p0} p_t + \gamma_{p1} p_{t-1} + u_t$$

Subtract q_{t-1} from both sides of the equation and add and subtract $\gamma_{p0} p_{t-1}$ from the right-hand side to get:

$$\Delta q_t = \gamma_0 + (\gamma_{q1} - 1) q_{t-1} + \gamma_{p0} \Delta p_t + (\gamma_{p0} + \gamma_{p1}) p_{t-1} + u_t$$

which can be rearranged as:

$$\Delta q_t = \gamma_0 + \gamma_{p0} \Delta P_t + (\gamma_{q1} - 1)(q_{t-1} - \frac{\gamma_{p0} + \gamma_{p1}}{1 - \gamma_{q1}} p_{t-1}) + u_t$$

2. ADL*(2,2)* case:

$$q_t = \gamma_0 + \gamma_{q1} q_{t-1} + \gamma_{q2} q_{t-2} + \gamma_{p0} p_t + \gamma_{p1} p_{t-1} + \gamma_{p2} p_{t-2} + u_t$$

Subtract q_{t-1} from both sides of the equation, then add and subtract $\gamma_{p0} p_{t-1}$, $\gamma_{p2} p_{t-1}$, and $\gamma_{q2} q_{t-1}$ from the right-hand side. This yields:

$$\Delta q_t = \gamma_0 + (\gamma_{q1} + \gamma_{q2} - 1) q_{t-1} - \gamma_{q2} \Delta q_{t-1} + \gamma_{p0} \Delta p_t + (\gamma_{p0} + \gamma_{p1} + \gamma_{p2}) p_{t-1} - \gamma_{p2} \Delta p_{t-1} + u_t$$

which can be rearranged as:

$$\Delta q_t = \gamma_0 + \gamma_{p0} \Delta p_t - \gamma_{p2} \Delta p_{t-1} - \gamma_{q2} \Delta q_{t-1} + (\gamma_{q1} + \gamma_{q2} - 1)(q_{t-1} - \frac{\gamma_{p0} + \gamma_{p1} + \gamma_{p2}}{1 - \gamma_{q1} - \gamma_{q2}} p_{t-1}) + u_t$$

3. ADL*(L,R)* case:

$$q_t = \gamma_0 + \gamma_{q1} q_{t-1} + \gamma_{q2} q_{t-2} + ... + \gamma_{qL} q_{t-L} + \gamma_{p0} p_t + \gamma_{p1} p_{t-1} + ... + \gamma_{pR} p_{t-R} + u_t$$

Subtract q_{t-1} from both sides of the equation and add and subtract $\gamma_{p0} p_{t-1}$, $\gamma_{p2} p_{t-1}$, … $\gamma_{pR} p_{t-1}$ and, $\gamma_{q2} q_{t-1}$, $\gamma_{q3} q_{t-1}$, … $\gamma_{qL} q_{t-1}$ from the right-hand side, and after rearranging we get the ECM:

[20] This case is adapted from Baltagi (2008 p. 141) and Charemza and Deadman (1997).

$$\Delta q_t = \gamma_0 + \gamma_{p0}\Delta p_t - \gamma_{p2}\Delta p_{t-1} - ... - \gamma_{pR}\Delta p_{t-(R-1)} - \gamma_{q2}\Delta q_{t-1} - ... - \gamma_{qL}\Delta q_{t-(L-1)}$$
$$+(\gamma_{q1} + \gamma_{q2} + ... + \gamma_{qL} - 1)(q_{t-1} - \frac{\gamma_{p0} + \gamma_{p1} + ... + \gamma_{pR}}{1 - \gamma_{q1} - \gamma_{q2} - ... - \gamma_{qL}} p_{t-1}) + u_t$$

which can be re-written as:

$$\Delta q_t = \gamma_0 + \gamma_{p0}\Delta p_t - \sum_{r=1}^{R-1} \gamma_{p,R+1}\Delta p_{t-r} - \sum_{l=1}^{L-1} \gamma_{q,L+1}\Delta q_{t-l} + (\gamma_{q1} + \gamma_{q2} + ... + \gamma_{qL} - 1)(q_{t-1}$$
$$- \frac{\gamma_{p0} + \gamma_{p1} + ... + \gamma_{pR}}{1 - \gamma_{q1} - \gamma_{q2} - ... - \gamma_{qL}} p_{t-1}) + u_t$$

Note that due to the presence of a contemporaneous p term on the right-hand side of the ADL there is an extra term ($\gamma_{p0}\Delta p_t$) of first differences of the independent variable in the ECM compared to the dependent variable.

A somewhat different transformation of the ADL *(3,3)* specification can be used when the ADL contains no contemporaneous regressors in order to get an ECM that contains both SR and LR elasticity estimates. Suppose the estimated ADL is ADL*(1-3, 1-3)*:

$$q_t = \gamma_0 + \gamma_{q1}q_{t-1} + \gamma_{q2}q_{t-2} + \gamma_{q3}q_{t-3} + \gamma_{p1}p_{t-1} + \gamma_{p2}p_{t-2} + \gamma_{p3}p_{t-3} + u_t$$

Subtract q_{t-1} from both sides of the equation to get:

$$\Delta q_t = \gamma_0 + (\gamma_{q1} - 1)q_{t-1} + \gamma_{q2}q_{t-2} + \gamma_{q3}q_{t-3} + \gamma_{p1}p_{t-1} + \gamma_{p2}p_{t-2} + \gamma_{p3}p_{t-3} + u_t$$

Next, add and subtract the following terms from the right-hand side of the equation: $(\gamma_{q1} - 1)q_{t-2}$, $\gamma_{q3}q_{t-2}$, $\gamma_{p1}p_{t-2}$, and $\gamma_{p3}p_{t-2}$ to get:

$$\Delta q_t = \gamma_0 + (\gamma_{q1} - 1)(q_{t-1} - q_{t-2}) + (\gamma_{q1} + \gamma_{q2} + \gamma_{q3} - 1)q_{t-2} - \gamma_{q3}(q_{t-2} - q_{t-3}) + \gamma_{p1}(p_{t-1} - p_{t-2}) +$$
$$(\gamma_{p1} + \gamma_{p2} + \gamma_{p3})p_{t-2} - \gamma_{p3}(p_{t-2} - p_{t-3}) + u_t$$

$$\Delta q_t = \gamma_0 + (\gamma_{q1} - 1)\Delta q_{t-1} - \gamma_{q3}\Delta q_{t-2} + \gamma_{p1}\Delta p_{t-1} - \gamma_{p3}\Delta p_{t-2} - (1 - \gamma_{q1} - \gamma_{q2} - \gamma_{q3})$$
$$\left[q_{t-2} - \frac{(\gamma_{p1} + \gamma_{p2} + \gamma_{p3})}{(1 - \gamma_{q1} - \gamma_{q2} - \gamma_{q3})} p_{t-2} \right] + u_t$$

where the coefficient in the error correction term is equal to the long-run elasticity and the coefficient on the first difference term, Δp_{t-1}, is equal to the SR elasticity.

For a general ADL*(L,R)*, subtract q_{t-1} from both sides of the equation and add and subtract $\gamma_{p1}p_{t-2}$, $\gamma_{p3}p_{t-2}$, … $\gamma_{pR}p_{t-2}$ and, $(\gamma_{q1} - 1)q_{t-2}$, $\gamma_{q3}q_{t-2}$, … $\gamma_{qL}q_{t-2}$ from the right-hand side, to get the equivalent ECM.

$$\Delta q_t = \gamma_0 + \gamma_{p1}\Delta p_{t-1} - \sum_{r=2}^{R-1}\gamma_{p,r+1}\Delta p_{t-r} + (\gamma_{q1}-1)\Delta q_{t-1} - \sum_{l=2}^{L-1}\gamma_{q,l+1}\Delta q_{t-l} - (1-\gamma_{q1}-\gamma_{q2}-...-\gamma_{qL})(q_{t-2}$$

$$-\frac{\gamma_{p1}+...+\gamma_{pR}}{1-\gamma_{q1}-\gamma_{q2}-...-\gamma_{qL}}p_{t-2}) + u_t$$

Appendix II

Unrestricted and restricted Vector Error Correction Models (VECMs) are reported below for interested readers. The conditional ECM reported in the text is fully efficient in our application. All estimated equations also include the monthly dummies, but for brevity their estimated coefficients are not reported. By testing whether the cointegrating relationships are present in the various equations or not, we can decide whether the variables P_e, P_g, Y, and N are weakly exogenous for the long-run elasticities or not. Note that even if the parameters of interest are both the short-run and long-run elasticities, this procedure is sufficient to reject weak exogeneity (Urbain, 1992, p.202).

Appendix Table 1: Estimates of the VECM with 1 cointegrating relationship

Vector Error Correction Estimates
Date: 11/09/13 Time: 15:41
Sample (adjusted): 1998M03 2006M12
Included observations: 106 after adjustments
Standard errors in ( ) & t-statistics in [ ]

Cointegration Restrictions:
B(1,1)=1
Convergence achieved after 1 iteration
Restrictions identify all cointegrating vectors
Restrictions are not binding (LR test not available)

Cointegrating Eq:	CointEq1
Q_e(-1)	1.00
P_e(-1)	0.14 (0.06) [2.30]
P_g(-1)	-0.03 (0.02) [-1.74]
Y(-1)	-0.30 (0.25) [-1.17]
N(-1)	-1.35 (0.43) [-3.13]

C	8.26				

Error Correction:	D(Qe)	D(Qe)	D(Pg)	D(Y)	D(N)
CointEq1	-0.97	**0.06**	**-0.14**	**-0.01**	**0.00**
	(0.06)	**(0.11)**	**(0.28)**	**(0.01)**	**(0.00)**
	[-15.74]	**[0.52]**	**[-0.51]**	**[-0.69]**	**[0.57]**
$D(Q_e(-1))$	-0.05	0.04	0.43	-0.00	0.00
	(0.05)	(0.08)	(0.21)	(0.01)	(0.00)
	[-0.99]	[0.45]	[2.05]	[-0.50]	[0.15]
$D(P_e(-1))$	0.05	-0.29	-0.05	0.01	0.00
	(0.06)	(0.11)	(0.27)	(0.01)	(0.00)
	[0.90]	[-2.75]	[-0.18]	[1.58]	[0.05]
$D(P_g(-1))$	-0.01	0.09	-0.17	0.00	0.00
	(0.02)	(0.04)	(0.10)	(0.00)	(0.00)
	[-0.42]	[2.18]	[-1.66]	[0.69]	[0.48]
D(Y(-1))	1.24	-0.26	-2.02	0.40	0.05
	(0.77)	(1.40)	(3.51)	(0.10)	(0.06)
	[1.60]	[-0.18]	[-0.57]	[3.88]	[0.88]
D(N(-1))	-2.97	-0.12	0.17	0.16	0.18
	(1.47)	(2.67)	(6.70)	(0.19)	(0.11)
	[-2.02]	[-0.04]	[0.02]	[0.82]	[1.69]
C	-0.14	0.02	-0.19	0.01	-0.00
	(0.03)	(0.06)	(0.14)	(0.00)	(0.00)
	[-4.48]	[0.28]	[-1.34]	[1.29]	[-0.08]
CDD	0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	[20.09]	[0.13]	[0.54]	[-0.98]	[-0.44]
HDD	0.00	-0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	[6.45]	[-0.24]	[1.73]	[-0.99]	[0.70]
R-squared	0.98	0.73	0.34	0.47	0.28
Adj. R-squared	0.98	0.67	0.19	0.35	0.12
Sum sq. resids	0.06	0.18	1.16	0.00	0.00
S.E. equation	0.03	0.05	0.12	0.00	0.00
F-statistic	277.24	12.09	2.29	3.97	1.74
Log likelihood	249.42	186.44	88.72	464.20	526.41
Akaike AIC	-4.33	-3.14	-1.30	-8.38	-9.55
Schwarz SC	-3.83	-2.64	-0.79	-7.88	-9.05
Mean dependent	0.00	0.00	0.00	0.00	0.00
S.D. dependent	0.18	0.08	0.13	0.00	0.00

Appendix Table 2: VECM estimation with weak exogeneity restrictions

Vector Error Correction Estimates
Date: 11/09/13 Time: 16:14

Included observations: 106 after adjustments
Standard errors in ( ) & t-statistics in [ ]

Cointegration Restrictions:
B(1,1)=1, A(2,1)=0, A(3,1)=0, A(4,1)=0, A(5,1)=0
Convergence achieved after 6 iterations.
Restrictions identify all cointegrating vectors
LR test for binding restrictions (rank = 1):

Chi-square(4)	2.54
Probability	0.64

Cointegrating Eq:	CointEq1
$Q_e(-1)$	1.00
$P_e(-1)$	0.16
	(0.06)
	[2.49]
$P_g(-1)$	-0.03
	(0.02)
	[-1.92]
Y(-1)	-0.33
	(0.26)
	[-1.30]
N(-1)	-1.26
	(0.44)
	[-2.89]
C	7.38

Error Correction:	D(Qe)	D(Pe)	D(Pg)	D(Y)	D(N)
CointEq1	-0.97	0.00	0.00	0.00	0.00
	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)
	[-17.12]	[NA]	[NA]	[NA]	[NA]
$D(Q_e(-1))$	-0.05	0.04	0.43	-0.00	0.00
	(0.05)	(0.08)	(0.21)	(0.01)	(0.00)
	[-1.06]	[0.49]	[2.05]	[-0.52]	[0.15]

D(P_e(-1))	0.06	-0.29	-0.05	0.01	0.00
	(0.06)	(0.11)	(0.27)	(0.01)	(0.00)
	[0.97]	[-2.74]	[-0.18]	[1.58]	[0.05]
D(P_g(-1))	-0.01	0.09	-0.17	0.00	0.00
	(0.02)	(0.04)	(0.10)	(0.00)	(0.00)
	[-0.52]	[2.18]	[-1.67]	[0.69]	[0.48]
D(Y(-1))	1.30	-0.28	-2.00	0.40	0.05
	(0.77)	(1.40)	(3.51)	(0.10)	(0.06)
	[1.68]	[-0.20]	[-0.57]	[3.89]	[0.88]
D(N(-1))	-2.92	-0.11	0.17	0.16	0.18
	(1.47)	(2.67)	(6.70)	(0.19)	(0.11)
	[-1.98]	[-0.04]	[0.02]	[0.82]	[1.69]
C	-0.14	0.02	-0.19	0.01	-0.00
	(0.03)	(0.06)	(0.14)	(0.00)	(0.00)
	[-4.44]	[0.27]	[-1.33]	[1.30]	[-0.09]
CDD	0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	[20.08]	[0.14]	[0.54]	[-0.98]	[-0.44]
HDD	0.00	-0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	[6.38]	[-0.23]	[1.72]	[-1.00]	[0.71]
R-squared	0.98	0.73	0.34	0.47	0.28
Adj. R-squared	0.98	0.67	0.19	0.35	0.12
Sum sq. resids	0.06	0.18	1.16	0.00	0.00
S.E. equation	0.03	0.05	0.12	0.00	0.00
F-statistic	276.78	12.08	2.29	3.96	1.74
Log likelihood	249.34	186.41	88.71	464.18	526.41
Akaike AIC	-4.33	-3.14	-1.30	-8.38	-9.55
Schwarz SC	-3.82	-2.64	-0.79	-7.88	-9.05
Mean dependent	0.00	0.00	0.00	0.00	0.00
S.D. dependent	0.18	0.08	0.13	0.00	0.00

The weak exogeneity restrictions are not rejected here, so the conditional single-equation approach in the text is fully efficient.