5 Top-Down Analysis of Energy Demand

The practical use of bottom-up models for analyzing energy demand is faced with significant micro-data requirements. In order to keep such models manageable, the individual components of energy demand are usually linked to the same macroeconomic variables such as the Gross Domestic Product (GDP), per-capita income, and relative energy prices. This gives rise to the question, “Why not model energy demand directly as a function of these macro variables?”. This macro approach is presented in this chapter, exploring the role of population growth, economic growth, and in particular changes in relative prices. However, this approach raises issues of its own:

– How does one differentiate between short-term and long-term adjustments of demand?
– Do rising and declining prices have the same effect on energy demand?
– How can the effects of technological change be isolated from the effects of changes in energy prices?
– Is the relationship between energy and other production inputs, in particular capital, substitutive or complementary?

Issues not discussed in this chapter include the possible instability of estimated relationships over time and reverse causality, i.e. the fact that GDP may not be exogenous but is in turn influenced by the price of energy (as evidenced by the recessions caused by the two oil price shocks of 1973/1974 and 1979/1980).

The variables used in this chapter are:

C	Total cost
c	Unit cost
ΔE	Change in demand for energy, $E_{t+1} - E_t$
E_t	Energy demand in period t
GDP	Gross Domestic Product
$\eta_{E,GDP}$	Income elasticity of energy demand
$\eta_{E,p}$	Price elasticity of energy demand

P. Zweifel et al., *Energy Economics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-662-53022-1_5

K	Capital stock
L	Labor
M	Input of materials
p	Price index
p_E	(Inflation-adjusted) energy price index
PCI	Per-capita income
POP	Population
Q	Real output (quantity)
sh	Cost share
σ	Elasticity of substitution

5.1 Population Growth

A first approach to top-down modeling is the following tautological relationship between population POP_t and the aggregate energy demand E_t,

$$E_t = POP_t \frac{E_t}{POP_t}. \tag{5.1}$$

Because

$$\frac{d\log E}{dE} = 1/E, \quad d\log E = \frac{dE}{E} \quad \text{holds, and hence} \tag{5.2}$$

$$\frac{\Delta E_t}{E_t} \approx \frac{\Delta POP_t}{POP_t} + \frac{\Delta(E_t/POP_t)}{(E_t/POP_t)}. \tag{5.3}$$

Therefore, the logarithmic differentiation of Eq. (5.1) yields the percentage change in energy demand as the sum of the percentage change in population and the per-capita energy demand, at least to a first approximation. The development of the demand for energy is thus tautologically given by the sum of population growth and change of per-capita energy consumption.

Table 5.1 shows the corresponding values and their percentage changes between 2000 and 2011. In 2011, the world population of nearly 7 bn people consumed 12.7 bn toe of commercial energy. However, while the per-capita energy consumption in China grew by no less than 7.5% p.a., it declined in several developed countries. Evidently there must be other factors at work beyond population growth that explain the demand for energy and its trend. Population growth alone is a misleading indicator.

However, population is of importance in a different context. Table 5.1 reveals substantial differences in per-capita energy consumption between countries. In the United States it is roughly twice as high as in Germany or Japan. With 0.17 toe per capita, the figure for Indonesia (as of 2011) is at the other end of the spectrum. The

Table 5.1 Population and per-capita primary energy supply

	Population 2011 (mn)	Change 2000–2011 (% p.a.)	Primary energy per capita 2011 (toe)	Change 2000–2011 (% p.a.)
Brazil	196.9	1.1	1.37	2.2
China	1344.1	0.6	2.03	7.5
France	65.3	0.6	3.87	−0.6
Germany	81.8	0.0	3.81	−0.6
India	243.8	1.4	3.07	3.1
Indonesia	1221.2	1.5	0.17	1.3
Italy	59.4	0.4	2.82	−0.6
Japan	127.8	0.1	3.61	−1.1
Nigeria	164.2	2.7	0.72	−0.2
Pakistan	176.2	1.9	0.48	0.7
Russia	143.0	−0.2	5.11	1.8
Turkey	73.1	1.3	1.54	2.2
United Kingdom	63.3	0.7	2.97	−2.2
United States	311.6	0.9	7.03	−1.2

Data source: World Bank (2014)

international disparities in energy consumption can be visualized by the so-called Lorenz curve. On the horizontal axis of Fig. 5.1, countries are ranked according to their shares of the world's population, with e.g. China accounting for the first 20%. The vertical axis exhibits the countries' respective shares of global energy consumption. If per-capita energy consumption were completely equal among countries, the Lorenz curve would be a diagonal running from point (0; 0) to point (1; 1). With increasing inequality, the Lorenz curve moves away from this straight line. According to the solid line, 60% of the world population accounted for only 17% of energy consumption in 2002.[1]

Yet the dashed Lorenz curve of Fig. 5.1 shows that income inequality between countries is even more marked than energy inequality, with 60% of the world population disposing of about 8% of world income only. This difference is an expression of the fact that energy has to be regarded as an essential good. Poor people devote a bigger share of their income to it than the rich. This observation suggests that the demand for energy increases less than proportionally with increasing income. Therefore, the so-called income elasticity of energy demand is smaller than one in the long run (see Sect. 5.2).

The World Energy Council (1993) stipulated 1.5 toe per capita as the benchmark to ensure economic and social development. This implies that global energy consumption would have to be about 40% higher than at present. The extent to

[1]This figure is not based on individual energy consumption but country-wide per-capita consumption. The Lorenz curve would look even more convex if referring to individual consumption values.

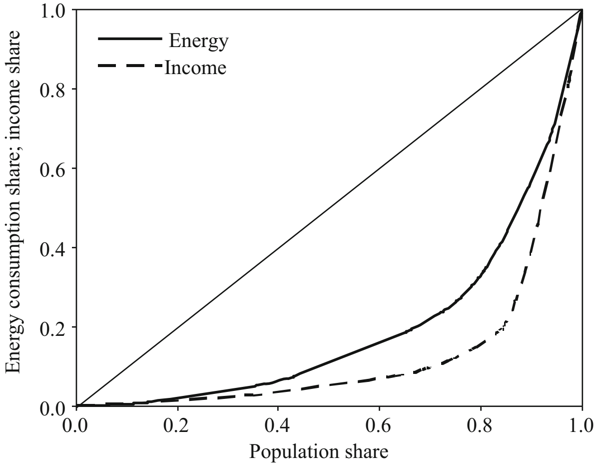


Fig. 5.1 Lorenz curves of the global energy consumption and income distribution. Data source: World Bank (2014)

which this postulate will be realized depends decisively on investment in energy technology, which in turn is driven by returns expected by investors and ability to pay of consumers.

5.2 Economic Growth

An increasing Gross Domestic Product (GDP) and the associated creation of value usually require an increased use of energy while at the same time leading to an improved capacity to pay for it. Similar to the tautology of Sect. 5.1, GDP can be inserted in the following way, with PCI_t denoting per-capita income,

$$E_t = POP_t \cdot \frac{GDP_t}{POP_t} \cdot \frac{E_t}{GDP_t} = POP_t \cdot PCI_t \cdot \frac{E_t}{GDP_t}. \tag{5.4}$$

After taking logarithms, the relative change in energy demand can be expressed as the sum of relative population growth, per-capita income growth, and the change in so-called energy intensity E_t/GDP_t,

$$\frac{\Delta E_t}{E_t} \approx \frac{\Delta POP_t}{POP_t} + \frac{\Delta PCI_t}{PCI_t} + \frac{\Delta(E_t/GDP_t)}{(E_t/GDP_t)}. \tag{5.5}$$

According to Table 5.2, energy intensity has decreased substantially between 2000 and 2011 in all countries sampled. Hence, consumption of energy has been increasing less than per-capita income, implying that the income elasticity of energy demand $\eta_{E,GDP}$ is smaller than one. It is defined as follows (see Eq. 5.2),

Table 5.2 Development of population, per-capita income, and energy intensity

Development 2000–2011	Population (% p.a.)	Per-capita income (% p.a.)	Energy intensity (% p.a.)	Primary energy consumption (% p.a.)
Brazil	1.1	2.4	−0.2	3.4
China	0.6	9.7	−2.1	8.1
France	0.6	0.6	−1.2	0.0
Germany	0.0	1.2	−1.8	−0.7
India	1.4	5.9	−2.6	4.6
Indonesia	1.5	3.9	−2.5	2.8
Italy	0.4	0.0	−0.6	−0.2
Japan	0.1	0.6	−1.7	−1.1
Nigeria	2.7	5.7	−5.6	2.5
Pakistan	1.9	2.2	−1.4	2.6
Russia	−0.2	5.0	−3.1	1.5
Turkey	1.3	2.9	−0.7	3.6
United Kingdom	0.7	1.1	−3.2	−1.5
United States	0.9	0.7	−1.9	−0.3

Data source: EIA (2014)

$$\eta_{E,GDP} = \frac{\text{percent change of } E}{\text{percent change of } GDP} = \frac{\partial \ln E}{\partial \ln GDP} = \frac{\partial E}{\partial GDP}\frac{GDP}{E}. \tag{5.6}$$

Here, the partial derivative indicates the *ceteris paribus* condition: the other determinants of energy demand (among them, the relative price of energy in particular) are held constant (see Sect. 5.3).

In normal circumstances, the income elasticity of energy demand is positive. The following distinctions can be made.

- $0 < \eta_{E,GDP} < 1$: In this case energy intensity E/GDP declines with growing income.
- $\eta_{E,GDP} > 1$: In this case the opposite holds. This is typical of developing countries, many of which are characterized by a backlog of demand at the going price.[2] This backlog is usually created by an artificially low price of energy imposed by the government.
- In the case of $\eta_{E,GDP} = 1$, energy intensity is independent of income.

[2]This statement serves as a reminder that actual energy consumption is interpreted as the outcome of supply and demand, both of which depend (among other things) on the relative price of energy. However, when the government fixes price below its equilibrium value, the quantity demanded exceeds the quantity supplied, creating a backlog in demand.

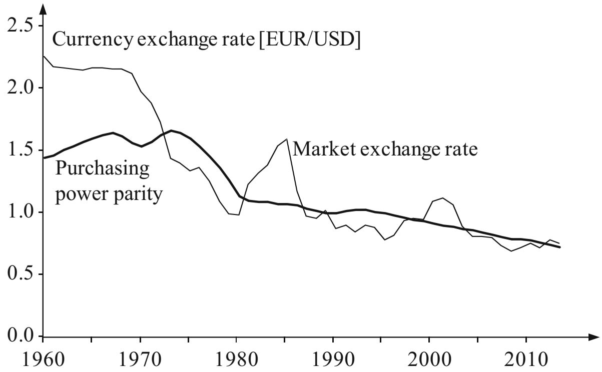


Fig. 5.2 Sample exchange rate and purchasing power parity. Data source: OECD

When GDP data of different years are to be compared, adjustment for inflation is necessary, as explained in Sect. 3.3. For international comparisons of energy intensity, one also has to convert GDP values expressed in national currency into a common currency (e.g. USD or EUR). In view of short-term exchange rate fluctuations, the average of a year (1995 in Table 5.2) or an average over several years (as is World Bank practice) may be the appropriate choice for depicting development over time.

Currency conversion can also be based on purchasing power parity (PPP), estimates of which are published by the OECD and the World Bank. This is a virtual exchange rate between two currencies, based on the notion that tradable goods have the same price everywhere. For example, if a hamburger costs 4 USD in the United States but 3 EUR in France (say), then 1 USD is apparently worth 0.75 EUR. When the hamburger is replaced by a basket of goods and services, one obtains the PPP. For most developing countries the application of PPP results in higher GDP values and hence lower estimates of energy intensity.[3] While exchange rates between industrialized countries tend to be closer to PPP values, there are deviations even in the EUR-USD exchange rate, as shown in Fig. 5.2.

5.3 The Price of Energy

The price of energy and its development over time (relative to the prices of other goods and services) are crucial determinants of the demand for energy. Considering the swath of energy prices, it is not easy to calculate a representative energy price index. Even for a given energy source, more than one price often exists. An

[3]The difficulty here is to establish the appropriate basket of goods and services. Goods and services have different weights between countries. In addition, price differences are justified by quality differentials, which must be filtered out. An alternative to taking a comprehensive basket of goods is to select one single good that is globally available. In the hamburger example, this results in the so-called Big Mac parity.

aggregated energy price index is constructed by weighting energy sources by their market shares. Since economic theory predicts that demand mainly depends on relative prices besides income or wealth, this index must be related to a macroeconomic price index (consumer price index, producer price index, or price index of the GDP). This ratio is often called the real price of energy. An increase over time signifies that energy prices grow faster than the average price of goods and services in general.

5.3.1 Short-Term and Long-Term Price Elasticities

Similar to the income elasticity, the energy own-price elasticity (often simply called price elasticity for short) is defined by using a change in the real price of energy p_E as the impulse,

$$\eta_{E,p_E} = \frac{\text{percent change of } E}{\text{percent change of } p_E} = \frac{\partial \ln E}{\partial \ln p_E} = \frac{\partial E}{\partial p_E}\frac{p_E}{E}. \tag{5.7}$$

Here as well, several cases need to be distinguished.

– $-1 < \eta_{E,pE} < 0$: The demand for energy is inelastic (to price). If the price of energy goes up (relative to the rate of inflation), the quantity of energy demanded declines less than proportionally (e.g. price goes up 10% but quantity sold only 4%, leaving a bottom line of 6% to sellers). This implies that inelastic demand results in a strong market position for suppliers (often called a seller's market).
– $\eta_{E,pE} < -1$: Demand for energy is elastic. If the (relative) price of energy increases, the quantity demanded declines more than proportionally (for instance a 10% price hike triggers a 12% fall in quantity sold, resulting in a bottom line of −2%). Elastic demand causes the market position of consumers to be strong (often called a buyer's market).

When energy markets become tight, the relative price of energy rises and the quantity traded decreases. This situation is shown in Fig. 5.3, where the supply curve shifts to the left. The original market equilibrium A is replaced by the one at point B, where the new supply function intersects with the short-term energy demand function. At first, a marked price hike combines with a limited decrease in quantity because immediate adjustment would be very costly for consumers. They often need to undertake an investment (e.g. by buying a car with higher fuel efficiency), a decision that is made only if the price change is viewed as permanent. Once undertaken, consumer adjustment gives rise to the long-term demand function (dotted line in Fig. 5.3), which is flatter (more price elastic) than its short-term counterpart and thus lies closer to the origin in the neighborhood of point B. This indicates reduced energy consumption at a given price, at least in the neighborhood of the initial equilibrium. Compared to B, the energy price drops slightly while the

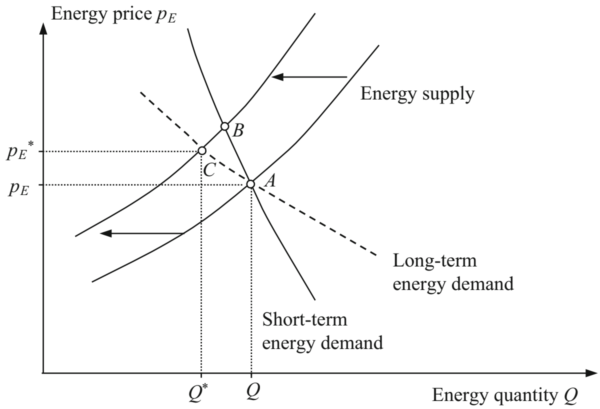


Fig. 5.3 Short-term and long-term effects of a reduction in energy supply

quantity traded continues to decrease until the new long-term equilibrium C is reached.

In normal circumstances, energy demand is a composite of demand for different energy sources E_i, $i = 1,\ldots, N$ with their relative prices p_i. This consideration gives rise to an extended definition of price elasticity,

$$\eta_{i,j} = \frac{\text{percent change of } E_i}{\text{percent change of } p_j} = \frac{\partial \ln E_i}{\partial \ln p_j} = \frac{\partial E_i}{\partial p_j}\frac{p_j}{E_i}. \tag{5.8}$$

For $i = j$ one obtains the own-price elasticity, for $i \neq j$ a so-called cross-price elasticity. While normally the own-price elasticity is negative, the cross-price elasticity can be of either sign. If the energy sources considered are substitutes, it is positive because a price increase $dp_j > 0$ leads to a response $dE_j < 0$, which triggers more demand for E_i, thus $dE_j > 0$. Yet E_i and E_j can also be complementary, in which case the cross-price elasticity is negative. A price increase $dp_j > 0$ again leads to a response $dE_j < 0$, which now causes a reduction $dE_i < 0$ in the complementary input E_i.

5.3.2 A Partial Energy Demand Model

A popular specification of a partial demand model reads[4]

$$E(t) = \alpha \cdot GDP(t)^{\beta} \cdot p_E(t)^{\gamma}. \tag{5.9}$$

[4]This approach can be understood as the result of the maximization of a so-called constant elasticity utility function, with energy and all other goods (at the price of 1) as its arguments, on the condition that the GDP is equivalent to income. See Varian (1992), Sect. 7.5.

Here, $E(t)$ symbolizes aggregate energy demand (in physical units), $GDP(t)$, the real (inflation-adjusted) Gross Domestic Product, and $p_E(t)$ the relative price of energy.

This formulation is partial rather than of the general-equilibrium type for three reasons:

– Consider a drop in the quantity of energy transacted. As the oil price shocks of 1973/1974 and 1979/1980 clearly demonstrated, this does not leave GDP unaffected. Therefore a reverse causality exists, running from $E(t)$ to $GDP(t)$.
– The observed consumption of energy is the outcome of an interaction between supply and demand. For example, if an increase in GDP causes the demand function to shift outward, the relative price of energy is predicted to go up. This time, causality runs from $GDP(t)$ to $p_E(t)$; superficially, it even seems to run from $E(t)$ to $p_E(t)$, indicating reverse causation.
– For all its importance, energy is just one factor of production both for households and firms. Therefore, a change in the relative price of energy has repercussions on the mix of factors of production. As a consequence, energy price developments may change inputs of other factors of production, which has an impact not only on the composition but also the size of GDP.

In keeping with the partial approach, economic theory indeed states that aggregate energy demand is determined by income and the relative price of energy. A mostly analogous formulation to Eq. (5.9) is

$$\frac{E(t)}{POP(t)} = \alpha \cdot \left(\frac{GDP(t)}{POP(t)}\right)^{\beta} \cdot p_E(t)^{\gamma} \tag{5.10}$$

with POP denoting the resident population. While the parameter α is a constant determining the general level of demand, β represents the income elasticity, and γ the price elasticity, respectively. This can be shown either by taking logarithms or by partial differentiation. In the latter case, one obtains

$$\frac{\partial E(t)}{\partial GDP(t)} = \alpha \cdot \beta \cdot GDP(t)^{\beta-1} \cdot p_E(t)^{\gamma} = \beta \cdot \frac{E(t)}{GDP(t)}. \tag{5.11}$$

Multiplication of both sides by $GDP(t)/E(t)$ yields β as the income elasticity $\eta_{E,GDP}$. Turning to the relative price of energy, one has

$$\frac{\partial E(t)}{\partial p_E(t)} = \alpha \cdot GDP(t)^{\beta} \cdot \gamma \cdot p_E(t)^{\gamma-1} = \gamma \cdot \frac{E(t)}{p_E(t)}. \tag{5.12}$$

Multiplication by $p_E(t)/E(t)$ shows that $\gamma = \eta_{E,Pe}$.

As it stands, the partial model does not permit to distinguish between short-term and long-term elasticities. This distinction is important because current energy consumption is the result of a reaction not only to current income and price, but also to past incomes and prices through the inherited stock of energy-using capital. This also implies that current energy consumption is to some extent determined by past energy consumption. The stronger this link, the longer it takes for a change in income or price to exert its full impact.

Two variants of this modified demand model are discussed in the literature. The stock adjustment hypothesis posits that consumers orient themselves to a desired (planned) energy consumption $E_p(t)$, which is a function of the desired stock of energy-using capital. In addition, the hypothesis assumes that this stock and hence planned energy consumption is determined by current income and the current relative price of energy, resulting in

$$E_p(t) = \alpha \cdot GDP(t)^{\beta} \cdot p_E(t)^{\gamma}. \tag{5.13}$$

During any given period t, however, there is a discrepancy between desired and actual (inherited) stock because adjustment is partial in view of its cost. If adjustment is completed up to a portion $(1-\rho)$ of the gap while ρ still is to be undertaken, one has

$$E(t) = E_p(t)^{1-\rho} \cdot E(t-1)^{\rho} \quad \text{with} \ \ 0 < \rho < 1. \tag{5.14}$$

The parameter ρ reflects the speed of adjustment. In the case of $\rho = 0$, adaptation to new market conditions happens without any delay, while in the case $\rho = 1$ no adjustment occurs at all. Note that ρ is to some degree an economic decision variable reflecting the benefits and costs of fast versus slow adjustment. This adjustment of the stock of energy-using capital is not explicitly modeled, in contrast with Eq. (4.1).

By substituting Eq. (5.14) into (5.13), one obtains according to the stock adjustment hypothesis,

$$E(t) = \alpha^{1-\rho} \cdot GDP(t)^{\beta\cdot(1-\rho)} \cdot p_E(t)^{\gamma\cdot(1-\rho)} \cdot E(t-1)^{\rho}. \tag{5.15}$$

The second approach is called habit persistence hypothesis. It states that the energy consumption $E(t)$ of period t is a function of expected future income $GDP^e(t)$ and expected relative energy price $p_E{}^e(t)$ rather than their current values,

$$E(t) = \alpha \ \ \cdot GDP^e(t)^{\beta} \cdot p_E^e(t)^{\gamma}. \tag{5.16}$$

Of course, an auxiliary hypothesis concerning the formation of expectations is needed. A popular alternative has been adaptive expectations, meaning that expectations are formed as an extrapolation from previous and current observation. If again a geometric mean is postulated, the pertinent functions read

$$\begin{aligned} GDP^e(t) &= GDP^e(t-1)^{\rho} \ \ \cdot GDP(t)^{1-\rho}, \\ p_E^e(t) &= p_E^e(t-1)^{\rho} \cdot p_E(t)^{1-\rho}. \end{aligned} \tag{5.17}$$

Here, $0 < \rho < 1$ denotes the parameter of adjustment as before, which now refers to expectations rather than the stock of energy-using capital. By substituting these expressions into Eq. (5.15) and taking into account

$$E(t-1)^{\rho} = \alpha^{\rho} \cdot GDP^{e}(t-1)^{\beta \cdot \rho} \cdot p_E^{e}(t-1)^{\gamma \cdot \rho} \tag{5.18}$$

by Eq. (5.16), Eq. (5.15) is obtained, with expected signs α, $\beta > 0$, $\gamma < 0$, and $0 < \rho < 1$.

In both approaches, energy demand $E(t)$ of period t (the dependent variable) is thus a function of income $GDP(t)$, relative energy price $p_E(t)$, and energy demand of the previous period $E(t–1)$, the lagged dependent variable. In order to estimate the parameters, appropriate data must be collected and econometric methods applied. In the case of model (5.18) a testable linear specification results when taking logs and adding an error term $\varepsilon(t)$,

$$\begin{aligned}\ln E(t) = (1-\rho)\ln\alpha + \beta(1-\rho)\ln GDP(t) + \\ \gamma(1-\rho)\ln p_E(t) + \rho E(t-1) + \varepsilon(t).\end{aligned} \tag{5.19}$$

The short-term income elasticity is $\beta(1–\rho) > 0$, the short-term price elasticity, $\gamma(1–\rho) < 0$. The long-term elasticities follow from considering the situation in which all impulses of a one-time income or price change have exerted their full effect, resulting in perfect adjustment (the unobserved energy-using stock of capital is constant). This means that energy demand is stationary,

$$E(t) = E(t-1). \tag{5.20}$$

In a stationary situation the time index may be omitted, resulting in

$$E = \alpha^{1-\rho} \cdot GDP^{\beta \cdot (1-\rho)} \cdot p_E{}^{\gamma \cdot (1-\rho)} \cdot E^{\rho} = \alpha \cdot GDP^{\beta} \cdot p_E^{\gamma}. \tag{5.21}$$

Therefore, β represents the long-term income elasticity and γ the long-term price elasticity. They can be obtained by dividing the estimated coefficients of Eq. (5.19) by the estimated ρ pertaining to the lagged dependent variable. The mean adjustment time (number of periods) following a one-time income or price change equals $1/(1–\rho)$. This follows from the fact that a discrepancy between desired and inherited energy consumption is reduced at the tune of $1–\rho$ per period. On average the discrepancy is thus eliminated in $1/(1–\rho)$ periods.

Yet this model is based on assumptions that prove to be restrictive:

– Rising and falling relative energy prices have a symmetric impact on energy demand, an assumption that is hardly plausible. So-called hysteresis is more likely, meaning that the consumption-reducing effect of a price hike continues even after price decreases again. After all, once equipment with higher energy efficiency is installed, it is not scrapped just because energy has become cheaper again. In order to model hysteresis, the price variable needs to be split in two,

$$\begin{aligned} p_E^+(t) &= p_E^+(t-1) + \max(0, p_E(t) - p_E(t-1)) \quad \text{(theup component)}, \\ p_E^-(t) &= p_E^-(t-1) + \min(0, p_E(t) - p_E(t-1)) \quad \text{(thedown component)}. \end{aligned} \tag{5.22}$$

Taking the price $p_E(0)$ of the base period 0 and using Eq. (5.22) again and again, the first variable p^+ contains the sum of all price increases beyond $p_E(0)$, while the second variable p^- contains the sum of all price decreases. Evidently

$$p_E(t) = p_E^+(t) + p_E^-(t) \tag{5.23}$$

holds. The modified Eq. (5.18) then reads,

$$E(t) = \alpha^{1-\rho} \cdot GDP(t)^{\beta\cdot(1-\rho)} \cdot p_E^+(t)^{\gamma\cdot(1-\rho)} \cdot p_E^-(t)^{\delta\cdot(1-\rho)} \cdot E(t-1)^\rho \tag{5.24}$$

where γ symbolizes the long-term price elasticity in case of price increases and δ for the long-term price elasticity in case of price decreases. Unfortunately, the explanatory variables often turn out to be highly correlated (giving rise to the so-called multicollinearity problem), rendering precise estimation of the parameters difficult.

– In the demand model presented, the mean adjustment time $1/(1–\rho)$ is independent of whether adjustment is triggered by change in income or energy prices. This assumption can be relaxed as well. For simplicity, consider the extreme case where the demand for energy reacts immediately to a change in income, while it reacts with a lag to a change in relative price, as before. In this case, Eq. (5.16) is modified as follows,

$$E(t) = \alpha \quad \cdot GDP(t)^\beta \cdot p_E^e(t)^\gamma \tag{5.25}$$

with $p_E{}^e(t)$ denoting the expected relative price of energy. In view of Eq. (5.17), this results in the specification

$$\frac{E(t)}{E(t-1)^\rho} = \alpha^{1-\rho} \frac{GDP(t)^\beta}{GDP(t-1)^{\beta\cdot\rho}} p_E(t)^{\gamma\cdot(1-\rho)}. \tag{5.26}$$

The parameter $\beta > 0$ represents the (short-term and long-term) elasticity, $\gamma(1–\rho) < 0$ the short-term and $\gamma < 0$, the long-term price elasticity. The mean adjustment time of energy demand to changes in price is once again $1/(1–\rho)$.

– A further assumption is that both short-term and long-term elasticities are constant. In particular, they do not depend on current values of income and price of energy. However, dropping this assumption calls for a much more complex modeling for demand (e.g. using the so-called translog specification, see Sect. 5.3.3).

Table 5.3 shows the results of a regression estimate of model (5.19) using annual data for the European Union and the United States covering the period from 1980 to 2013. The dependent variable is crude oil demand per capita. The explanatory variables are inflation-adjusted per-capita income and inflation-adjusted price of Brent crude. Demand for crude oil is fairly well explained: The coefficient of determination R^2 is between 0.88 and 0.91, indicating a high statistical fit. Though estimated income elasticities have the expected sign, they are statistically insignificant at the 1% level. However, the price elasticities and the lagged dependent variable are statistically significant at the 1% level.

Taking the results of Table 5.3 at face value, one is led to the following interpretation. In both the European Union and in the United States, the inflation-adjusted price of crude oil has a significant impact on demand. However, the long-term price elasticities are 0.220 or below in absolute value, which means that the demand for oil is inelastic. This view is confirmed by most econometric studies. Interestingly enough, consumers in the United States react with a shorter lag to oil price changes than Europeans.

Table 5.3 Income and price elasticities of crude oil demand

	EU-15	USA
Inflation-adjusted per-capita income (data source: World Bank)		
Short-term (β $(1-\rho)$)	0.022	0.008
Long-term (β)	0.146	0.037
Inflation-adjusted Brent price (data source: BP and World Bank)		
Short-term (γ $(1-\rho)$)	−0.034 (*)	−0.036 (*)
Long-term (γ)	−0.220 (*)	−0.163 (*)
Per-capita oil consumption (data source: BP and World Bank)		
Lagged dependent variable (ρ)	0.847 (*)	0.779 (*)
Adjustment lag (years)	6.5 (*)	4.5 (*)
Adjusted R^2	0.910	0.882
Standard error of estimate	0.0162	0.0147

Estimation period: 1980–2013; the significance of elasticities is denoted with * (1% level)

5.3.3 Substitution Between Energy and Capital

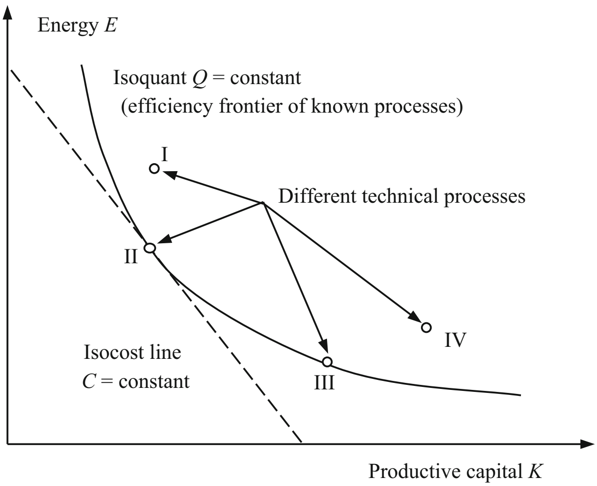
According to the process analysis discussed in Sect. 4.3, the relationship between energy demand and aggregate capital stock is substitutional rather than limitational (fixed proportions), meaning there is a choice between more and less energy-intensive modes of production and consumption. Substitution of energy is therefore possible through investment in capital goods. More generally, a given output quantity Q can be produced using more or less energy input E and commensurately modified quantities of other production factors. This is formally expressed by a production function,

$$Q = f(K, L, E, M) \tag{5.27}$$

which relates output Q to inputs of capital K, labor L, energy E, and materials M (non-energy raw materials). To be precise, Q denotes the maximum output achievable given the state of technology and input quantities, reflecting best practice.

Figure 5.4 depicts the production function by means of a so-called isoquant. An isoquant shows the quantities of production factors K and E (with inputs of labor L and materials M held constant in the present case) that are needed to produce a given quantity Q. The isoquant thus summarizes the efficient production frontier for a given quantity of output, depicting uses of an available technology ranging from energy-intensive to capital-intensive. Specifically, production process *II* is an energy-intensive variant that in turn uses little capital, whereas process *III* is capital-intensive but saves on energy. Production processes *I* and *IV* can be disregarded because of their excessive use of costly inputs; indeed, only technologies *II* and *III* are efficient.

Fig. 5.4 Efficient and inefficient technical processes



In order to decide which production alternative is minimum cost, the prices of production factors have to be considered. For K this is capital user cost p_K, an annuity which reflects interest and depreciation (net of tax exemptions and subsidies); for L this is the wage rate (including non-wage labor costs) p_L; for E this is the average energy purchase price p_E; for M this is average unit cost p_M. When L and M as well as their unit prices are held constant, all factor combinations that are compatible with constant total production cost lie on a straight line (the so-called isocost line; see Fig. 5.4). Per definition all points on the isocost line are characterized by constant total production cost,

$$C = p_K K + p_E E + p_L L + p_M M. \tag{5.28}$$

To calculate the slope of this line and neglecting L and M (thus $dL = dM = 0$), this equation can be differentiated to become

$$\begin{aligned} &dC = 0 = p_K dK + p_E dE \quad \text{and therefore} \\ &\frac{dE}{dK} = -\frac{p_K}{p_E}. \end{aligned} \tag{5.29}$$

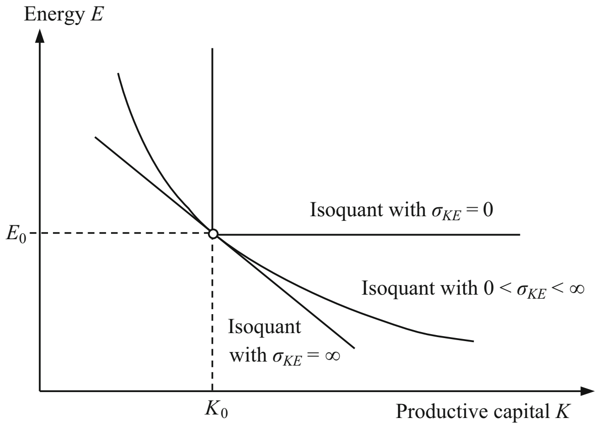
Thus the slope of the isocost line is $-p_K/p_E$. Consider a reduction in energy use, $dE < 0$. If the unit price of energy p_E is relatively high, the cost saving is substantial, permitting to use a lot more capital K while holding cost constant. Therefore, dE/dK takes on a low (absolute) value in this case. Conversely, if energy is cheap compared to the user cost of capital, $dE < 0$ generates a small cost saving which creates little room for an increased use of capital since this is relatively expensive. Accordingly, dE/dK takes on a high (absolute) value in this case.

Competitive pressure makes producers minimize cost. They therefore seek to attain the isocost line representing the lowest possible production cost. This is the one running closest to the origin in (K, E)-space, given the amount of output Q and hence the isoquant. Therefore, the isocost line needs to be tangent to the isoquant for cost minimization. This corresponds to the choice of technology *II* with its rather high energy intensity.

Note that the isocost line of Fig. 5.4 has a fairly steep slope, reflecting a situation where energy is cheap compared to the user cost of capital (as reflected by the annuity; see Sect. 3.1). If energy were to become more expensive relative to capital, the isocost line would exhibit a reduced slope, thus favoring a more capital-intensive mode of production. Therefore, a change in relative prices is predicted to affect the choice of production process within the technology available. This constitutes producers' short-run response, while the choice of technology (to be discussed in Sect. 5.4) amounts to their long-run adjustment.

For a given technology, the curvature of the isoquant representing it evidently is of great importance. The more pronounced the curvature, the smaller is the adjustment in the factor mix in response to a given change in relative factor prices. In the extreme case of a limitational technology, isoquants have an angular shape, which means that there cannot be any adjustment to a change in relative prices. Producers

Fig. 5.5 Isoquants with different elasticities of substitution



are stuck at the corner as it were (see Fig. 5.5). With reference to a pair of inputs, one defines the elasticity of substitution σ_{KE} as the parameter reflecting the degree of substitutability in production. It answers the question, "By how much (in percent) does the mix of inputs change when their relative price changes by 1%?". In terms of Fig. 5.4, one has

$$\sigma_{KE} = \frac{d(K/E)}{dR} \frac{R}{K/E} \quad \text{with the slope of the tangent given by} \\ R = -\frac{dK}{dE} = -\frac{\partial Q}{\partial E} \Big/ \frac{\partial Q}{\partial K}. \tag{5.30}$$

Obviously, the slope of the (tangent to an) isoquant reflects relative marginal productivities. It is known as the marginal rate of (factor) substitution. Given choice of a minimum-cost production process, the marginal rate of substitution R is just equal to the (negative) relative price of the factors (the slope of the isocost line),

$$R = \frac{p_K}{p_E}. \tag{5.31}$$

Therefore, the elasticity of substitution can also be defined in terms of a change in relative factor prices,

$$\sigma_{KE} = \frac{d(E/K)}{d(p_K/p_E)} \frac{p_K/p_E}{E/K} = \frac{d\ln(E/K)}{d\ln(p_K/p_E)} = \sigma_{EK}. \tag{5.32}$$

The symmetry follows from $d\ln(K/E) = -d\ln(E/K)$ and $d\ln(p_K/p_E) = -d\ln(p_E/p_K)$. If capital and energy are substitutive factors of production, the elasticity of substitution must lie in the interval $0 < \sigma_{KE} < \infty$. A high value of σ_{KE} indicates that substitution between these factors is easy. With an elasticity of substitution $\sigma_{KE} = 3$ e.g., a 10% increase in the relative price of energy results in a 30% reduction of the E/K ratio. If $\sigma_{KE} = 1.2$, the E/K ratio falls by 12% only. In the extreme case of $\sigma_{KE} = 0$, there is no substitution possibility between K and E but a fixed input relationship between them (this amounts to a fixed-proportions (limitational) production function, also called Leontief production function; see Sect. 2.5).

The situation becomes more complex when more than two production factors are considered. The partial elasticity of substitution defined above has to be replaced by the so-called Allen elasticity (see Allen 1938, Sect. 19.4). Any two factors of production may now be complementary rather than substitutive. For instance, labor has historically been substituted by capital and energy, making capital and energy complements in production. Since the elasticity of substitution is defined to be positive in the case of substitutability, the Allen elasticity is negative in the case of complementarities.

This raises the question of whether energy and capital are complements or substitutes in the context of the four-factor production function $Q = f(K, L, E, M)$. This is an empirical question which can only be answered by applying econometric methods. In doing so, one usually prefers not to focus on the isoquants but rather on (minimum) cost, which is a scalar measure. As shown by Fig. 5.4, the isocost line contains the same information as the isoquant in the neighborhood of the minimum cost combination of inputs. Indeed the problem, "Minimize production cost for a given output level" leads to the same solution as the so-called dual formulation, "Maximize output for a given cost budget". Thus, the dual to maximizing output given a cost constraint reads

$$C = C(Q, p_K, p_L, p_E, p_M) = Q \cdot c(p_K, p_L, p_E, p_M). \tag{5.33}$$

It states that minimum total cost C depends on the amount of output Q to be achieved and the (relative) prices of inputs. Since unit cost c is given by C/Q, one can also analyze unit cost c. Strictly speaking, this is possible only if scaling up by Q does not matter, i.e. if the cost function $C(Q, p_K, p_L, p_E, p_M)$ is homogenous of the first degree in Q. This is the case when a change of all production factors (e.g. doubling all of them) leads to an analogous change (doubling) of output, amounting to constant returns to scale.

This leaves the choice of functional form. Preferably, the functional form should not impose *a priori* restrictions on crucial parameters such as the elasticity of substitution. A popular solution is the so-called translog function (see Christensen et al. 1973; Berndt and Wood 1975). It results from a second-degree Taylor approximation to an arbitrary function, with the arguments and the dependent variable in logarithms. In the case of the average cost function, the translog form becomes

$$\begin{aligned}
&\ln c = \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + \frac{1}{2}\sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} \ln p_i \ln p_j, \\
&\sum_{i=1}^{N} \alpha_i = 1, \ \sum_{i,j=1}^{N} \beta_{ij} = \sum_{i,j=1}^{N} \beta_{ji}, \ \beta_{ij} = \beta_{ji}.
\end{aligned} \tag{5.34}$$

Here $\alpha_0 > 0$ is a constant and N is the number of production factors ($N = 4$ in the present context). The α_i are the first-order derivatives of the unit cost function with reference to the inputs. They sum up to one because of the assumed homogeneity of the first degree. The β_{ij} are the second-order derivatives. By Young's theorem, the order of differentiation does not matter for continuously differentiable functions, implying $\beta_{ij} = \beta_{ji}$.

Differentiating Eq. (5.34), one obtains

$$\frac{\partial \ln c}{\partial \ln p_i} = \frac{\partial c}{\partial p_i} \cdot \frac{p_i}{c} = \alpha_i + \frac{1}{2} \sum_{j=1}^{N} \beta_{ij} \cdot \ln p_j \tag{5.35}$$

Shephard's lemma states that the derivative of the minimum cost function with respect to factor price p_i yields the optimal input quantity x_i (see e.g. Varian 1992, Chap. 5). Therefore one obtains

$$\frac{x_i \cdot p_i}{c} =: sh_i = \alpha_i + \sum_{j=1}^{N} \beta_{ij} \ln p_j \tag{5.36}$$

with sh_i denoting the cost share of the i-th factor of production (see Diewert 1974). Thus, the shares of K, L, E, and M can be linearly related to the logarithm of their prices p_K, p_L, p_E, and p_M, making estimation of the β_{ij} by ordinary least-squares (OLS) possible.

Also, the Allen partial elasticities of substitution between capital and energy can be recovered from the cost shares and price elasticities as follows (see Allen 1938, Sects. 19.5 and 19.6),

$$\sigma_{KE} = \frac{\eta_{K,p_E} + sh_E \cdot \eta_{E,p_E}}{sh_E}. \tag{5.37}$$

This is intuitive: Energy and capital are substitutes if their cross-price elasticity $\eta_{K,pE}$ is positive, resulting in a positive value of σ_{KE} (the own-price elasticity $\eta_{K,pE}$ is always negative). Conversely, they are complements if their cross-price elasticity is strongly positive and the own-price elasticity of energy as well as its cost share sh_E are small in absolute value, resulting in a negative value of σ_{KE}.

Econometric estimation of substitution elasticities between energy and capital was motivated by the first oil price shock of 1973. Policy-makers wanted to know whether it was easy or difficult to substitute energy by other production factors, in particular capital. The first evidence exhibited in Table 5.4 was disappointing:

Table 5.4 Elasticities of substitution between capital, labor, and energy

Elasticity of substitution	Berndt and Wood	Griffin and Gregory	Hunt (1984)	Hunt (1986)
	United States		Great Britain	
	1975	1976	(neutral technological change)	(non-neutral techn. change)
σ_{KL}	1.01	0.06	1.58	0.37
σ_{KE}	−3.22	1.07	−1.64	2.68
σ_{LE}	0.64	0.87	0.84	0.08

Berndt and Wood (1975) found a complementary relation between energy and capital. Yet another estimate by Griffin and Gregory (1976) points to substitutability ($\sigma_{KE} = 1.07$). This triggered a lively discussion among economists (see e.g. Solow 1987). Later studies using more recent data and including technological change also show ambiguous results. However, the estimate presented in the last column of Table 5.4 confirms substitutability between energy and capital once it is assumed that producers have a choice of technology. This leads to the conclusion that companies have not exhausted the substitutional potential suggested by bottom-up process analysis to the same extent as the potential for automation, which amounts to replacing labor by both capital and energy.

In conclusion, the relationship between energy and capital cannot be determined with sufficient precision even to this day. Likely reasons are the limited validity of aggregate data, difficulty in distinguishing between the short term and long term (K and E may be complementary in the short run but substitutes in the long run), and the challenges posed by isolating the effects of technological change.

5.4 Technological Change

In economics, technological change is defined in the following way. Technological change enables a larger output Q to be produced with the same input quantities of capital K, labor L, energy E, and materials M. An equivalent way of expressing the same idea is to say that a given output quantity Q can be produced using smaller quantities of production factors. An improvement in quality is a possible outcome, too.

In Fig. 5.6, technological change is depicted by a shift of the isoquant towards (and not away from) the origin of (K, E)-space. In the figure on the left, technological change does not affect the input mix as long as relative factor prices do not change, thus indicating neutral technological change with respect to energy and capital. In the figure on the right, technological change is energy-saving because the transition exhibits a lower E/K ratio at a given factor price ratio. Clearly, changing relative factor prices can also influence the choice of production technology, in addition to technological change.

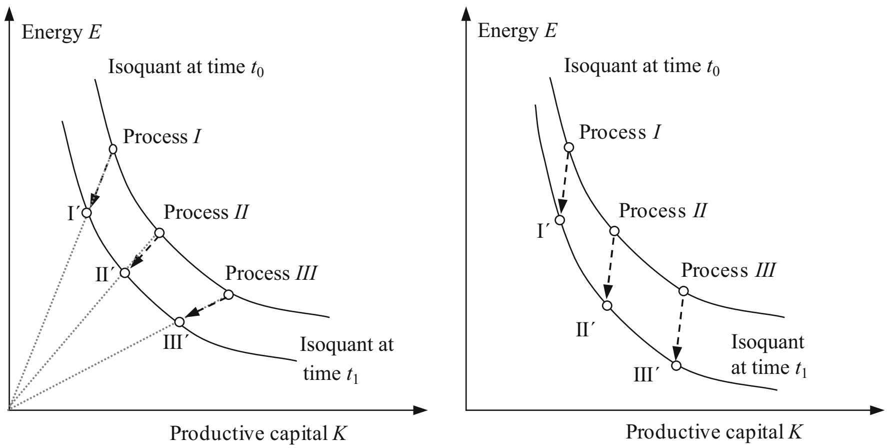


Fig. 5.6 Isoquants reflecting technological change

A mathematical formulation of factor-augmenting technological change is the following (see Stoneman 1983),

$$Q = f(K, L, E, M) = g[a_K(t)K, \quad a_L(t)L, \quad a_E(t)E, a_M(t)M] \tag{5.38}$$

where t denotes time and $a_K(t)$, $a_L(t)$, $a_E(t)$, and $a_M(t)$ are functions indicating factor-augmenting changes resulting in savings of capital, labor, energy, and materials. These functions may depend on investment in research and development, education and training of the labor force, improved management, institutional reforms, and much more. But if technological change is to be advantageous, these functions must obey

$$\frac{da_K}{dt} \geq 0; \quad \frac{da_L}{dt} \geq 0; \quad \frac{da_L}{dt} \geq 0, \quad \frac{da_M}{dt} \geq 0. \tag{5.39}$$

Using these functions, the direction of technological change can be defined. For example, $da_L(t)/dt > 0$ while $da_K(t)/dt = da_E(t)/dt = da_M(t)/dt = 0$ indicates labor-saving technological change because only labor input is scaled up as it were (which implies less of it is actually used at unchanged relative factor prices). But if $da_L(t)/dt = da_K(t)/dt = da_E(t)/dt = da_M(t)/dt$, then all factors of production benefit from technological change to the same degree, a case which is often referred to as Hicks-neutral technological change. In the past, however, technological change has not been neutral but first and foremost labor-saving. Using the expression for sh_i derived from the translog unit cost function (5.36) and complementing it with $\gamma_i \cdot t$ to reflect technological change, one obtains (see Binswanger 1974; Hunt 1986)[5]

[5]In this formula, technological change is understood as autonomous. In fact, it may be linked to investment in capital and additions to the workforce. Thus, technological change is incorporated in

$$sh_i(t) = \alpha_i + \sum_j \beta_{ij} \ln p_j(t) + \gamma_i t. \tag{5.40}$$

As before, $sh_i(t)$ stands for the cost share of production factor i and $p_j(t)$ for the respective price. As the shares of the factor costs have to add up to 1, one has

$$\sum_j \gamma_j = 0. \tag{5.41}$$

Hicks-neutral technological change is equivalent to $\gamma_K = \gamma_L = \gamma_E = \gamma_M = 0$. However, using British industry data covering the years from 1960 to 1980, Hunt (1986) finds evidence suggesting $\gamma_L < \gamma_E < \gamma_K$. Therefore, in Great Britain at least, technological innovation has been above all labor-saving. It has also been energy-saving, though hardly capital-saving.

This finding gives rise to the question of which factors bring about such a bias to technological change. According to the induced bias hypothesis (see Stoneman 1983, Chap. 4), the direction of the technological change is conditioned by the market, meaning it depends on the development of relative prices. This was first formulated by Hicks (1932, 124f),

> A change in the relative prices of the factors of production is itself a spur to invention, and to innovation of a particular kind—directed to economizing the use of a factor which has become relatively expensive.

If the price of the production factor labor increases compared to the cost of other production factors, then labor-saving technological change will come about in due time. In addition to the movement along the isoquant as in Fig. 5.6, a change in the isoquant itself also takes place, which leads to a further substitution of labor even when relative factor prices no longer change.

It can be argued that the oil crises of the 1970s with their twin price shocks have guided innovation efforts towards improved energy efficiency. Interestingly, these efforts continued into the 1990s when relative oil prices were lower again, possibly because of the (expected) scarcity of energy resources and governments aiming at reducing greenhouse gas emissions and supporting energy-efficient investments. Apparently, the price hikes triggered an enduring technological change which has decoupled the demand for energy from economic growth. Quite possibly, this decoupling will be enhanced by the renewed increase in the relative price of oil between 1999 and 2014. Historic case studies show that problem awareness in the energy industry has influenced the direction technological change (see Weizsäcker 1988) as entrepreneurs hope to make a profit by developing energy-saving and environment-friendly technologies and products. These hoped-for innovation gains thus may play an important role in the future demand for energy.

---

the factors of production. This feature can be taken into account by the capital vintage model presented in Sect. 4.1.

References

Allen, R. G. D. (1938). *Mathematical analysis for economists*. London: MacMillan.

Berndt, E. R., & Wood, D. O. (1975). Technology, prices and the derived demand for energy. *Review of Economics and Statistics, 57*, 259–268.

Binswanger, H. P. (1974). A microeconomic approach to induced innovation. *Economic Journal, 84*, 940–958.

Christensen, L. R., Jorgenson, D. W., et al. (1973). Transcendental logarithmic production frontiers. *Review of Economics and Statistics, 55*, 28–45.

Diewert, W. E. (1974). An application of the Shephard duality theorem: A generalized Leontief production function. *Journal of Political Economy, 79*, 481–507.

EIA. (2014). *Miscellaneous data files*. Washington: Energy Information Administration. Retrieved from www.eia.gov/

Griffin, J. M., & Gregory, P. R. (1976). An intercountry translog model of energy substitution responses. *American Economic Review, 66*, 845–857.

Hicks, J. R. (1932). *The theory of wages*. London: MacMillan.

Hunt, L. C. (1984). Energy and capital: Substitutes or complements? Some results for the UK industrial sector. *Applied Economics, 16*, 783–789.

Hunt, L. C. (1986). Energy and capital: Substitutes or complements? *Applied Economics, 18*, 729–735.

Solow, J. L. (1987). The capital-energy complementary debate revisited. *American Economic Review, 77*, 605–614.

Stoneman, P. (1983). *The economic analysis of technological change*. Oxford: Oxford University Press.

Varian, H. L. (1992). *Microeconomic analysis* (3rd ed.). New York: Norton.

Weizsäcker, C. C. von (1988). Innovationen in der Energiewirtschaft (Innovations in the energy economy). *Zeitschrift für Energiewirtschaft, 3*, 141–146.

World Bank. (2014). *World development indicators*. Washington. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators