Chapter 4

Elasticities of Substitution

R. Robert Russell
University of California, Riverside

Abstract

This chapter lays out the theoretical foundation of the measurement of the degree of substitutability among inputs utilized in a production process. It proceeds from the well-settled (Hicksian) notion of this measure for two inputs (typically labor and capital) to the more challenging conceptualization for technologies with more than two inputs (most notably, Allen-Uzawa and Morishima elasticities). Dual elasticities of substitution (also called elasticities of complementarity) and gross elasticities of substitution (measuring substitutability for non-homothetic technologies taking account of output changes) are also covered. Also analyzed are functional representations of two-input technologies with constant elasticity of substitution (CES) and of n-input technologies with constant and identical elasticities for all pairs of inputs. Finally, the chapter explores the relationship between elasticity values and the comparative statics of factor income shares and the relationships between certain elasticity identities and separability conditions rationalizing consistent aggregation of subsets of inputs.

4.1 Introduction.

In his classic book on the *Theory of Wages*, the Oxford University economist John R. Hicks [1932] introduced two concepts that persist to this day as important components of both microeconomic and macroeconomic analysis: (1) elasticity of substitution[1] and (2) input neutrality (alternatively, input bias) of technological change. Each of these constructs is fundamental to the analysis of changing factor income shares as an economy (or other production unit) expands (or, for that matter, contracts). This chapter focuses on the first of these concepts; Chapter 5 covers the second.

As noted by Blackorby and Russell [1989, p. 882] in their discussion of the elasticity of substitution, "Hicks' key insight was to note that [in a two-factor economy] the effect of changes in the capital/labor ratio (or the factor price ratio) on the distribution of income (for a given output) can be completely characterized by a scalar measure of curvature of the isoquant."[2] This measure, the (two-input) *elasticity of substitution*, is a logarithmic derivative of the input-quantity ratio with respect to the technical rate of substitution between the two inputs, holding output constant. The elasticity of substitution and its relationship to the comparative statics of income shares is exposited in Section 4.2, where the famous SMAC (Arrow, Chenery, Minhaus, and Solow [1961]) theorem on the functional characterization of constancy of the elasticity of substitution (CES) is also discussed.

Generalization of the elasticity of substitution to allow for more than two inputs began with suggestions by Hicks and Allen [1934]. One suggestion was to employ the constructions defining the original Hicksian notion for any two inputs, holding the other input quantities fixed. This idea was further explored by McFadden [1963], but since then it has faded from the picture, largely because its failure to allow for optimal adjustment of other inputs means that it generally fails to provide information about the comparative

[1] The concept was independently formulated by Cambridge University economist Joan Robinson [1933] in her comparably classic book on *The Economics of Imperfect Competition*. Abba Lerner [1933] and A. C. Pigou [1934] also contributed to the understanding of the concept at its genesis.

[2] While the use of the word "curvature" in this quote conveys the appropriate intuition, it is nevertheless technically incorrect, in part because curvature, formally defined, is a unit-dependent mathematical concept. See de la Grandville [1997] for a clear exposition of this point.

statics of relative income shares of any two inputs.[3]

The other generalization formulated by Hicks and Allen [1934]—and further analyzed by Hicks [1938], Allen [1938], and Uzawa [1962]—is now known as the Allen elasticity of substitution (AES) or the Allen-Uzawa elasticity of substitution (AUES). This elasticity is a share-weighted (constant-output) cross elasticity of demand. An alternative generalization, first formulated by Morishima [1967] (in Japanese and unfortunately never translated into English) and independently discovered by Blackorby and Russell [1975], is a constant-output cross elasticity of demand minus a constant-output own price elasticity of demand.[4] Blackorby and Russell [1981] named this concept the "Morishima elasticity of substitution" (MES) and argued that it, unlike the AES, preserves the salient properties of the original Hicksian notion when the number of inputs is expanded to more than two. As these (and other) elasticity concepts are most evocatively described using dual representations of the technology, Section 4.3 presents some useful duality constructs. Features of the AES and the MES are then explored in Section 4.4, particularly their relationships to the comparative statics of factor income shares and functional representations of the technology when the elasticities are invariant with respect to changes in input quantities.

The curvature of a two-input isoquant can be equivalently (and again informally[5]) represented by the inverse of the Hicksian elasticity, the logarithmic derivative of the technical rate of substitution—*i.e.*, the shadow price ratio—with respect to the quantity ratio. These concepts are dual to one another. Of course, in the case of only two variables, these two elasticities are simple inverses of one another (and have inverse implications for the curvature of the isoquant). With more than two inputs, however, the analogous dual concepts—duals to the Allen and Morishima elasticities—are not simple inverses of one another. This dual structure, developed by Blackorby and

---

[3]As pointed out by Blackorby and Russell [1981, p. 882], "[O]nly if the two variables were separable from all other variables would [this elasticity] provide information about shares; if we were to require all pairs to have this property, the production function would be additive. When combined with homotheticity (an assumption maintained in all these studies . . .) this implies that the production function is CES, in which case [the elasticities] are constant for all pairs of inputs."

[4]The Morishima elasticity is a generalization of Robinson's [1933] characterization of the two-input elasticity and for this reason is called the "Robinson elasticity of substitution by Kuga and Murota [1972] and Kuga [1979].

[5]See footnote 2 above

Russell [1975, 1981], is examined in Section 4.5.2.

Stern [2010] points out that the dual Morishima elasticities do not reflect differential movements along an isoquant (essentially because only one input is varied in the calculation). He proposes an elasticity, named the symmetric elasticity of complementarity, that constrains differential changes in quantities to be contained in an isoquant. This elasticity, unlike the dual Morishima elasticity, is symmetric.[6] It is presented in Section 4.5.3.

The Allen and Morishima elasticities of substitution are calculated for differential movements along a constant-output surface. If the technology is homothetic, this is not a restriction, but in general the comparative-static calculations on, say, income shares hold only for cases where outputs are exogenous. The extension of the Allen elasticity to incorporate output effects was broached by Mundlak [1968] and formulated in the dual (using the profit function) by Lau [1978]. Extensions of these results to Morishima elasticities can be found in Davis and Shumway [1996] and Blackorby, Primont, and Russell [2007]. In line with the latter, I refer to these concepts as gross elasticities and describe and analyze them in Section 4.6.

In a widely cited paper, Berndt and Christensen [1973a] were the first to notice that identities among certain pairs of Allen elasticities are equivalent to corresponding separability restrictions on the production function. Essentially, a set of inputs is separable from a distinct input if the technical rates of substitution among inputs in the set are independent of the quantity of the excluded input. Separability is a powerful concept in part because it has implications for the possibility of consistent aggregation of inputs (i.e., aggregation across different types of labor inputs to form an aggregate input in the functional representation of the technology).[7] If a subset of inputs is separable from all inputs excluded from the subset, there exists an aggregator over the inputs in the subset, which then is an aggregate input into the production function. Berndt and Christensen discovered that a subset of inputs is separable from a distinct input if and only if the Allen elasticities of substitution between the excluded and each of the inputs in the separable

[6] It is dual, not to the Morishima elasticity, but to McFadden's [1963] shadow elasticity of substitution.

[7] Separability is also a necessary condition for decentralization of an optimization problem (as in, *e.g.*, two-stage budgeting). See Blackorby, Primont, and Russell [1978, Ch. 5] for a thorough exposition of the connection between separability and decentralized decision making.

subset are identical. Russell [1975] and Blackorby and Russell [1976, 1981] later generalized these results and extended them to comparable identity restrictions for Morishima elasticities of substitution. These relationships between certain elasticity identities and functional structure are covered in Section 4.7.

Chapter 4.8 concludes with a brief discussion of the extensive theoretical and empirical literature in which the elasticity of substitution plays a salient role.

4.2 Two-Input Elasticity of Substitution: Early Formulations and Characterizations.

4.2.1 Definition.

Hicks was particularly interested in the substitutability between labor and capital, and more particularly in the relative income shares of these two inputs. Let us therefore denote the input quantity vector, in an obvious notation, by $\langle x_\ell, x_k \rangle \in \mathbf{R}^2_+$ and the (scalar) output quantity by $y \in \mathbf{R}_+$. The production function, $F : \mathbf{R}^2_+ \to \mathbf{R}_+$, is assumed to be increasing, strictly quasi-concave, and homothetic.[8] For convenience, we restrict our analysis to the interior of quantity space, $\mathbf{R}^2_{++}$, and assume that F is continuously twice differentiable on this space.[9]

Homotheticity of F implies that the technical rate of substitution between labor and capita,[10]

$$trs_{\ell,k} = F_\ell(x_\ell, x_k)/F_k(x_\ell, x_k) =: TRS_{\ell,k}(x_\ell, x_k),$$

[8] As we shall see in Section 3, the homotheticity assumption can be dropped when the elasticity concept is formulated in the dual.

[9] We could extend our analysis to all of $\mathbf{R}^2_+$ by employing directional derivatives at the boundary but instead leave this technical detail to the interested reader.

[10] Subscripts on functions indicate differentiation with respect to the specified variable. The relation, $=$, should be interpreted as an identity throughout this chapter (*i.e.*, as holding for all allowable values of the variables). Also, $A := B$ means the relation defines A, and $A =: B$ means the relation defines B.

is homogeneous of degree zero, so that

$$trs_{\ell,k} = TRS_{\ell,k}(1, x_k/x_\ell) =: t(x_k/x_\ell)$$

and

$$\ln trs_{\ell,k} = \ln t(x_k/x_\ell) =: \theta(\ln(x_k/x_\ell)).$$

Owing to strict quasi-concavity of F, t is strictly monotonic, hence invertible, and we can write

$$\frac{x_k}{x_\ell} = t^{-1}(trs_{\ell,k})$$

or

$$\ln \frac{x_k}{x_\ell} =: \phi(\ln trs_{\ell,k}). \tag{4.1}$$

The (two-input) *elasticity of substitution* is defined as the log derivative of the input-quantity ratio with respect to the technical rate of substitution,

$$\sigma = \phi'(\ln trs_{\ell,k}),$$

or, equivalently, as the inverse of the log derivative of the technical rate of substitution with respect to the input-quantity ratio,

$$\sigma = \frac{1}{\theta'(\ln(x_k/x_\ell))}.$$

As the production function is strictly quasi-concave, the elasticity of substitution lies in the open interval, $(0, \infty)$. Relatively large values of σ indicate that the rate at which one input can be substituted for the other is relatively insensitive to changes in the input ratio: in the vernacular, substitution is "relatively easy" (isoquants are "relatively flat"). Conversely, lower values of σ reflect "relatively difficult" substitution (and "strong curvature" of isoquants). As $\sigma \to \infty$, the isoquants converge to (parallel) linear line segments (perfect substitution), and as $\sigma \to 0$, the isoquants converge to Leontief (fixed proportions) isoquants.

4.2.2 Comparative Statics of Income Shares.

As shown by Hicks [1932], the value of the elasticity of substitution has unambiguous implications for the effects of changes in relative factor prices or in relative factor quantities on relative factor shares in a competitive (price taking) economy. In an obvious notation for factor prices, the share of capital relative to labor is $s = p_k x_k / p_\ell x_\ell$. In a competitive economy, where $trs_{\ell,k} = p_\ell / p_k$,

$$\begin{aligned} \ln s &= \ln(x_k/x_\ell) - \ln(p_\ell/p_k) \\ &= \phi(\ln(p_\ell/p_k)) - \ln(p_\ell/p_k) =: \hat{S}(\ln(p_\ell/p_k)), \\ &= \ln(x_k/x_\ell) - \theta(\ln(x_k/x_\ell)) =: \tilde{S}(\ln(x_k/x_\ell)), \end{aligned}$$

so that the proportional effect on relative shares of a change in the price ratio is

$$\hat{S}'(\ln(p_\ell/p_k)) = \phi'(\ln(p_\ell/p_k)) - 1 = \sigma - 1$$

and the proportional effect on relative shares of a change in the quantity ratio is

$$\tilde{S}'(\ln(x_k/x_\ell)) = 1 - \theta'(\ln(x_k/x_\ell)) = 1 - \frac{1}{\sigma}.$$

Thus, the comparative statics of functional income shares when σ is constant is encapsulated in

$$\frac{d\ \hat{S}(\ln(p_\ell/p_k))}{d\ \ln(p_\ell/p_k)} \begin{pmatrix} < \\ = \\ > \end{pmatrix} 1 \quad \Longleftrightarrow \quad \sigma \begin{pmatrix} < \\ = \\ > \end{pmatrix} 1$$

or

$$\frac{d\ \tilde{\mathcal{S}}(\ln(x_k/x_\ell))}{d\ \ln(x_k/x_\ell))} \begin{pmatrix} < \\ = \\ > \end{pmatrix} 1 \quad \Longleftrightarrow \quad \sigma \begin{pmatrix} > \\ = \\ < \end{pmatrix} 1.$$

4.2.3 Constant Elasticity of Substitution.

Implicit in Hicks's analysis of substitutability is the assumption that the elasticity is constant for all values of the input vector. This begs the question of the restrictiveness of this assumption, a question that later was answered by Arrow, Chenery, Minhas, and Solow [1961] (affectionately known as SMAC in much of the literature). Assuming that the technology is convex and homogeneous of degree one, they showed that the elasticity of substitution is constant (independent of x_k/x_ℓ) if and only if

$$F(x_\ell, x_k) = \left(\alpha_\ell x_\ell^\rho + \alpha_k x_k^\rho\right)^{1/\rho}, \quad 1 \geq \rho \neq 0, \qquad (CES_2)$$

or

$$F(x_\ell, x_k) = \gamma x_\ell^\alpha x_k^{1-\alpha}, \quad \gamma > 0,\ 1 > \alpha > 0. \tag{CD_2}$$

Thus, constancy of the elasticity of substitution implies that the production function must be belong to the *CES family* of technologies, (CES_2) or (CD_2).[11] Moreover, the fixed-proportions technology (Leontief [1953]) is generated as a limiting case of (CES_2):

$$\lim_{\rho \to -\infty} \left(\alpha_\ell x_\ell^\rho + \alpha_k x_k^\rho\right)^{1/\rho} = \min\{\alpha_\ell x_\ell, \alpha_k x_k\}. \tag{L_2}$$

Some simple calculations establish the following equivalences: (CES_2) $\iff$ $\sigma = 1/(1-\rho)$; (CD_2) $\iff$ $\sigma = 1$; and (L_2) $\iff$ $\sigma \to 0$.[12]

I need not bring to the attention of the reader the extent to which the CES family of production functions has been, over the years, a perennial workhorse in both theoretical and empirical research employing production functions.

4.3 Digression: Dual Representations of Multiple-Input, Multiple-Output Technologies.

The Allen and Morishima elasticities, originally formulated by Hicks and Allen [1934] and Morishima [1967] in the (primal) context of a single-output production function, are more generally and evocatively exposited in the dual

[11] The Cobb-Douglas production function was well known at the time of the SMAC derivation, having been proposed much earlier (Cobb and Douglas [1928]). The (CES_2) production function made its first appearance in Solow's [1956] classic economic growth paper, but the functional form had appeared much earlier in the context of utility theory: Bergson (Burk) [1936] proved that additivity of the utility function and linear Engel curves (expenditures on individual goods proportional to income for given prices) implies that the utility function belongs to the CES family. As the SMAC authors point out, the function (CES_2) itself was long known in the functional-equation literature (see Hardy, Littlewood, and Pólya [1934, p. 13]) as the "mean value of order ρ".

[12] The SMAC theorem is easily generalized to homothetic technologies, in which case the production function is a monotonic transformation of (CES_2) or (CD_2); in the limiting case as $\sigma \to 0$ it is a monotonic transformation of (L_2).

(using the cost function) as first shown, respectively, by Uzawa [1962][13] and by Blackorby and Russell [1975, 1981], Kuga [1979], and Kuga and Murota [1972]. The dual approach also allows us to move seamlessly from technologies with a single output to those with multiple outputs. Finally, duality theory is needed for the development of the dual elasticities of substitution in Section 4.5. This section lays out the requisite duality theory.[14]

Denote the ordered set of inputs by $N = \langle 1, \ldots n \rangle$ and the ordered set of outputs by $M = \langle 1, \ldots m \rangle$. Input and output quantity vectors are denoted $x \in \mathbf{R}^n_+$ and $y \in \mathbf{R}^m_+$, respectively. The technology set is the set of all feasible $\langle$input, output$\rangle$ combinations:

$$T := \{\langle x, y \rangle \in \mathbf{R}^{n+m}_+ \mid x \text{ can produce } y\}.$$

While the nomenclature suggests that feasibility is a purely technological notion, a more expansive interpretation is possible: feasibility could incorporate notions of institutional and political constraints, especially when we consider entire economies as the basic production unit.

An input requirement set for a fixed output vector y is

$$L(y) := \{x \in \mathbf{R}^n_+ \mid \langle x, y \rangle \in T\}.$$

We assume throughout that, for all $y \in \mathbf{R}^m_+$, $L(y)$ is closed and strictly convex (relative to $\mathbf{R}^n_+$)[15] and satisfies strong input disposability

$$L(y) = L(y) + \mathbf{R}^n_+ \ \forall \ y \in \mathbf{R}^m_+,$$

output monotonicity,[16]

$$\bar{y} > y \Rightarrow L(\bar{y}) \subset L(y),$$

---

[13]Because the Allen elasticities are now typically exposited in the dual, they are often called the "Allen-Uzawa elasticities," and I resort to that nomenclature on occasion as well. Uzawa's approach was later extended by Blackorby and Russell [1975, 1981].

[14]Thorough expositions of duality theory can be found in, *e.g.*, Blackorby, Primont, and Russell [1978], Chambers [1988], Cornes [1992], Diewert [1974, 1982], Färe and Primont [1995], Fuss and McFadden [1978], and Russell [1997].

[15]These assumptions are stronger than needed for much of the conceptual development that follows, but in the interest of simplicity I maintain them throughout.

[16]Vector notation: $\bar{y} \geq y$ if $\bar{y}_j \geq y_j$ for all j; $\bar{y} > y$ if $\bar{y}_j \geq y_j$ for all j and $\bar{y} \neq y$; and $\bar{y} \gg y$ if $\bar{y}_j > y_j$ for all j.

and "no free lunch",

$$y \neq 0^{(m)} \Rightarrow 0^{(n)} \notin L(y).$$

The (input) distance (gauge) function, a mapping from[17]

$$Q := \{\langle x, y\rangle \in \mathbf{R}_{+}^{n+m} \mid y \neq 0^{(m)} \ \wedge \ x \neq 0^{(n)} \ \wedge \ L(y) \neq \emptyset\}$$

into the positive real line (where $0^{(n)}$ is the null vector of $\mathbf{R}_{+}^{n}$), is defined by

$$D(x, y) := \max\{\lambda \mid x/\lambda \in L(y)\}.$$

Under the above assumptions, D is well defined on this restricted domain and satisfies homogeneity of degree one, positive monotonicity, concavity, and continuity in x, and negative monotonicity in y. (See, *e.g.*, Färe and Primont [1995] for proofs of these properties and most of the duality results that follow.[18]) Assume, in addition, that D is continuously twice differentiable in x. The distance function is a representation of the technology, since (under our assumptions)

$$\langle x, y\rangle \in T \iff D(x, y) \geq 1.$$

In the single-output case ($m = 1$), where the technology can be represented by a production function, $F : \mathbf{R}_{+}^{n} \rightarrow \mathbf{R}_{+}$, $D(x, F(x)) = 1$ and the production function is recovered by inverting $D(x, y) = 1$ in y. If (and only if) the technology is homogeneous of degree one (constant returns to scale),

$$D(x, y) = \frac{F(x)}{y}.$$

The cost function, $C : \mathbf{R}_{++}^{n} \times Y \rightarrow \mathbf{R}_{+}$, where

$$Y = \{y \mid \langle x, y\rangle \in Q \text{ for some } x\},$$

is defined by

$$C(p, y) = \min_{x}\{p \cdot x \mid x \in L(y)\}$$

---

[17] We restrict the domain of the distance function to assure that it is globally well defined. An alternative approach (*e.g.*, Färe and Primont [1995]) is to define D on the entire non-negative $(n + m)$-dimensional Euclidean space and replace "max" with "sup" in the definition. See Russell [1997, footnote 12] for a comparison of these approaches.

[18] Whatever is not there can be found in Diewert [1974,] or the Fuss/McFadden [1978] volume.

or, equivalently, by

$$C(y,p) = \min_x \{p \cdot x \mid D(x,y) \geq 1\}. \tag{4.2}$$

Under our maintained assumptions, D is recovered from C by

$$D(x,y) = \inf_p \{p \cdot x \mid C(p,y) \geq 1\}, \tag{4.3}$$

and C has the same properties in p as D has in x. This establishes the duality between the distance and the cost function. On the other hand, C is *positively* monotonic in y. We also assume that C is twice continuously differentiable in p.

By Shephard's Lemma (application of the envelope theorem to (4.2), the (vector-valued, constant-output) input demand function, $\delta : \mathbf{R}^n_{++} \times Y \to \mathbf{R}^n_+$, is generated by first-order differentiation of the cost function:[19]

$$\delta(p,y) = \nabla_p C(p,y).$$

Of course, δ is homogeneous of degree zero in p. The (normalized) shadow-price vector, $\rho : Q \to \mathbf{R}_+$, is obtained by applying the envelope theorem to (4.3):

$$\rho(x,y) = \nabla_x D(x,y). \tag{4.4}$$

As is apparent from the re-writing of (4.3) (using homogeneity of C in p) as

$$D(x,y) = \inf_{p/c} \Big\{ \frac{p}{c} \cdot x \mid C(p/c, y) \geq 1 \Big\} = \inf_{p/c} \Big\{ \frac{p}{c} \cdot x \mid C(p,y) \geq c \Big\}, \tag{4.5}$$

where c can be interpreted as (minimal) expenditure (to produce output y) and the vector $\rho(x,y)$ in (4.4) can be interpreted as shadow prices normalized by minimal cost.[20] In other words, under the assumption of cost-minimizing behavior,

$$\rho(\delta(p,y), y) = \frac{p}{C(p,y)}.$$

Clearly, ρ is homogeneous of degree zero in x.

[19] $\nabla_p C(p,y := \langle\ C(p,y)/\partial p_1, \ldots, \partial C(p,y)/\partial p_n \rangle$.

[20] See Färe and Grosskopf [1994] and Russell [1997] for analyses of the distance function and associated shadow prices.

4.4 Allen and Morishima Elasticities of Substitution.

4.4.1 Allen Elasticities of Substitution (AES).

The Allen elasticity of substitution between inputs i and j, is given by

$$\sigma_{ij}^{A}(p,y) : = \frac{C(p,y)C_{ij}(p,y)}{C_i(p,y)C_j(p,y)} \tag{4.6}$$

$$= \frac{\epsilon_{ij}(p,y)}{s_j(p,y)}, \quad \forall \langle i,j\rangle \in N \times N, \tag{4.7}$$

where the subscripts on the cost function C indicate differentiation with respect to the indicated variable(s);

$$\epsilon_{ij}(p,y) := \frac{\partial \ln \delta_i(p,y)}{\partial \ln p_j} = \frac{p_j C_{ij}(p,y)}{C_i(p,y)}.$$

is the (constant-output) elasticity of demand for input i with respect to a change in the price of input j; and

$$s_j(p,y) = \frac{p_j \delta_j(p,y)}{C(p,y)} = \frac{p_j C_j(p,y)}{C(p,y)}$$

is the cost share of input j. Thus, the Allen elasticity is simply a share-weighted cross ($i \neq j$) or own ($i = j$) demand elasticity. It collapses to the Hicksian elasticity when $n = 2$.

4.4.2 Morishima Elasticities of Substitution (MES).

To define the Morishima elasticities, let p^i be the $(n-1)$-dimensional vector of price ratios with p_i in the denominator. Zero-degree homogeneity of δ in p allows us to write

$$\hat{\delta}(p^i,y) := \delta(p,y).$$

The Morishima elasticity of substitution of input i for input j is defined directly as

$$\sigma_{ij}^{M}(p,y) := \frac{\partial \ln \left(\hat{\delta}_i(p^i,y) \;/\; \hat{\delta}_j(p^i,y)\right)}{\partial \ln(p_j/p_i)}. \tag{4.8}$$

Note that any variation in a single component of p^i, say $p_{j'}/p_i$, holding other components ($j \neq j'$) constant, must be entirely manifested in variation in $p_{j'}$ alone. Hence, for all pairs $\langle i, j \rangle$,

$$\frac{\partial \hat{\delta}_i(p^i, y)}{\partial (p_j/p_i)} = \frac{\partial \delta_i(p, y)}{\partial p_j} \frac{1}{p_i}.$$

Using this fact, along with Shephard's Lemma, the MES can be re-written in the dual as

$$\sigma_{ij}^M(p, y) = p_j \left(\frac{C_{ij}(p, y)}{C_i(p, y)} - \frac{C_{jj}(p, y)}{C_j(p, y)} \right) \tag{4.9}$$

$$= \epsilon_{ij}(p, y) - \epsilon_{jj}(p, y). \tag{4.10}$$

Thus, the MES is simply the difference between the appropriate (constant output) cross price elasticity of demand and the (constant output) own elasticity of demand for the input associated with the axis along which the price ratio is being varied.

The Morishima elasticity, unlike the Allen elasticity, is non-symmetric, since the value depends on the normalization adopted in (4.8)—that is, on the coordinate direction in which the prices are varied to change the price ratio, p_j/p_i[21] Of course, if there are only two inputs, there is no difference between changing p_i and changing p_j, and the MES, like the AES, collapses to the standard Hicksian elasticity. But in the general n-input case, the interactions between these two inputs is dependent on whether it is one or the other that is varied.[22]

4.4.3 AES and MES and the Comparative Statics of Income Shares.

If $\sigma_{ij}^A(p, y) > 0$ (that is, if increasing the j^{th} price increases the optimal quantity of input i), we say that inputs i and j are Allen-Uzawa substitutes; if $\sigma_{ij}^A(p, y) < 0$, they are Allen-Uzawa complements. Similarly, if $\sigma_{ij}^M(p, y) > 0$

[21] See Blackorby and Russell [1975, 1981, 1989] for a discussion of this asymmetry, which (as pointed out by Stern [2011a]) was recognized much earlier by Pigou [1934] in his analysis of Robinson's [1933] less-formal characterization of this elasticity concept.

[22] "Own" Morishima elasticities are identically equal to zero and hence uninteresting, as one might expect to be the case for a sensible elasticity of *substitution*.

(that is, if increasing the j^{th} price increases the optimal quantity of input i relative to the optimal quantity of input j), we say that input j is a Morishima substitute for input i; if $\sigma_{ij}^M(p, y) < 0$, input j is a Morishima complement to input i. As the Morishima elasticity of substitution is non-symmetric, so is the taxonomy of Morishima substitutes and complements.[23]

The conceptual foundations of Allen-Uzawa and Morishima taxonomies of substitutes and complements are, of course, quite different. The Allen-Uzawa taxonomy classifies a pair of inputs as substitutes (complements) if an increase in the price of one causes an increase (decrease) in the quantity demanded of the other. This is the standard textbook definition of net substitutes (and complements). The Morishima concept, on the other hand, classifies a pair of inputs as substitutes (complements) if an increase in the price of one causes the quantity of the other to increase (decrease) *relative to the quantity of the input for which price has changed*. For this reason, the Morishima taxonomy leans more toward substitutability (since the theoretically necessary decrease in the denominator of the quantity ratio in (4.8) (owing to $\epsilon_{jj}(p, y) < 0$) helps the ratio to decline when the price of the input in the denominator increases).

Put differently, if two inputs are substitutes according to the Allen-Uzawa criterion, theoretically they must be substitutes according to the Morishima criterion, but if two inputs are complements according to the Allen-Uzawa criterion, they can be either complements or substitutes according to the Morishima criterion. This relationship can be seen algebraically from (4.7) and (4.10) . If i and j are Allen-Uzawa substitutes, in which case $\epsilon_{ij}(p, y) > 0$, concavity of the cost function (and hence negative semi-definiteness of the corresponding Hessian) implies that $\epsilon_{ij}(p, y) - \epsilon_{jj}(p, y) > 0$, so that j is a Morishima substitute for i. Similar algebra establishes that two inputs can be Morishima substitutes when they are Allen-Uzawa complements.

Note that, for $i \neq j$,

$$\frac{\partial \ln s_i(p, y)}{\partial \ln p_j} = \epsilon_{ij}(p, y) - s_j(p, y) = s_j(p, y)(\sigma_{ij}^A(p, y) - 1),$$

so that an increase in p_j increases the absolute cost share of input i if and

[23] These notions are referred to as "p-substitutes" and "p-complements" in much of the literature (see Stern [2011a] and the papers cited there), as distinguished from "q-substitutes" and "q-complements", which I call dual substitutes and complements in Section 4.5.

only if $\sigma_{ij}^A(p,y) > 1$; that is, if and only if inputs i and j are sufficiently strong net substitutes. Thus, the Allen-Uzawa elasticities provide immediate qualitative comparative-static information about the effect of price changes on absolute shares. To obtain quantitative comparative-static information, one needs to know the share of the j^{th} input as well as the Allen-Uzawa elasticity of substitution.

The Morishima elasticities immediately yield both qualitative and quantitative information about the effect of price changes on *relative* input shares:

$$\frac{\partial \ln(\hat{s}_i(p^i,y)/\hat{s}_j(p^i,y))}{\partial \ln(p_j/p_i)} = \epsilon_{ij}(p,y) - \epsilon_{jj}(p,y) - 1 = \sigma_{ij}^M(p,y) - 1,$$

where (with the use of zero-degree homogeneity of s_i in p) $\hat{s}_i(p^i,y) := s_i(p,y)$ for all i. Thus, an increase in p_j increases the share of input i relative to input j if and only if

$$\sigma_{ij}^M(p,y) > 1;$$

that is, if and only if inputs i and j are sufficiently substitutable in the sense of Morishima. Moreover, the degree of departure of the Morishima elasticity from unity provides immediate quantitative information about the effect on the relative factor shares.

4.4.4 Constancy of the Allen and Morishima Elasticities.

In addition to his dual reformulation of the elasticity proposed by Hicks and Allen [1934], Uzawa [1962] also extended the SMAC theorem on constancy of the elasticity of substitution to encompass more than two inputs. He conjectured (p. 293) that the "production function which extends the Arrow-Chenery-Minhaus-Solow function to the n-factor case may be the following type:"

$$F(x) = \left(\sum_{i\in N} \alpha_i\, x_i^\rho\right)^{1/\rho}, \quad \alpha_i > 0 \;\forall\, i, \quad 1 \geq \rho \neq 0. \qquad (CES)$$

This structure does indeed yield a constant Allen-Uzawa elasticity, $\sigma = 1/(1-\rho)$, for all pairs of inputs and moreover converges to

$$F(x) = \alpha \prod_{i\in N} x_i^{\beta_i}, \quad \alpha > 0, \quad \beta_i > 0 \;\forall\, i, \qquad (CD)$$

with $\sigma = 1$ as $\rho \to \infty$. Moreover, Uzawa shows that this structure is *necessary* as well as sufficient for *constancy* of the Allen-Uzawa elasticities, but it turns out not to be necessary for the elasticities to be *identical* for all pairs of inputs.

Uzawa went on to establish necessary (as well as sufficient) conditions for constancy and uniformity of the AES for a PLH production function. To explicate these conditions, consider a partition of the set of inputs into m subsets, $\mathcal{N} = \{N^1, \ldots, N^S\}$, with n_s inputs in subset s for each s. Uzawa showed that the AES are constant and identical if and only if, for some partion $\mathcal{N}$, the (PLH) production function is given by

$$F(x) = \prod_{s=1}^{S} F^s(x^s)^{\beta_s}, \quad \beta_s > 0 \;\forall\; s \quad \text{and} \quad \sum_{s=1}^{S} \beta_s = 1, \tag{4.11}$$

where, for all s,

$$F^s(x^s) = \left(\sum_{i \in N^s} \alpha_i x_i^{\rho_s} \right)^{1/\rho_s}, \quad 0 \neq \rho_s < 1, \quad \alpha_i > 0 \;\forall\; i \in N^s. \tag{4.12}$$

That is, the production function can be written as a Cobb-Douglas function of CES aggregator functions. Note that the structure (4.11)–(4.12) collapses to the (CES) case when $S = 1$ and to the (CD) case when $|N^s| = 1$ for all s. Moreover, when $n = 2$, (4.11) collapses to (CD_2) and (4.12) collapses to (CES_2).[24]

[24]McFadden [1963] showed that his direct elasticites of substitution are constant and identical if and only if

$$F(x) = \left(\sum_{s=1}^{S} \alpha_s F^s(x^s) \right)^{1/\rho}, \quad 0 \neq \rho < 1/n^*, \quad 0 \leq \rho < 1/n^*, \quad \alpha_s > 0 \;\forall\; s,$$

or

$$F(x) = \alpha_0 \prod_{s=1}^{S} F^s(x^s)^{\alpha_s / n^*}, \quad \alpha_s > 0 \;\forall\; s, \quad \textit{and} \quad n^* = \max_s \{n_s\}$$

where

$$F^s(x^s) = \prod_{i=1}^{n_s} x_i, \quad s = 1, \ldots, m;$$

that is, if and only if the production function can be written as a CES or Cobb-Douglas function of (specific) Cobb-Douglas aggregator functions.

Analogous results for constancy and uniformity of the Morishima elasticities were established by Blackorby and Russell [1975] and Kuga [1979] (generalizing the three-input case proved by Murota [1977]). Again maintaining positive linear homogeneity of the production function, the MES are constant if and only if the production function takes the form (CES) or (CD) above.

Blackorby and Russell [1989, p. 888] point to the contrast between these representations for the AES and the MES to support their view that the latter is the natural generalization of the Hicks elasticity to encompass more than two inputs: if equation (CES) "is *the* CES production function, then the MES—and not the AES—is *the* elasticity of substitution."

Using standard duality theory, constancy of the elasticity of substitution can also be characterized in terms of the structure of the cost and distance functions. For PLH production functions, the Morishima elasticities are globally constant and identical if and only if the cost function takes the form[25]

$$C(p,y) = y\left(\sum_{i=1}^{n} \hat{\alpha}_i\ p_i^{\hat{\rho}}\right)^{1/\hat{\rho}}, \quad \hat{\alpha}_i = \alpha_i^{1/(1-\rho)}\ \forall i, \quad \hat{\rho} = \rho/(\rho-1), \tag{4.13}$$

or

$$C(p,y) = y\prod_{i=1}^{n} \hat{\alpha}_i p_i^{\beta_i}, \quad \hat{\alpha}_i = \alpha_i^{\alpha_i}\ \forall\ i. \tag{4.14}$$

The distance function dual to (4.13) and (4.14), formally derived using (4.3), has the following alternative images:

$$D(x,y) = y^{-1}\left(\sum_{i=1}^{n} \alpha_i x_i^{\rho}\right)^{1/\rho} \tag{4.15}$$

or

$$D(x,y) = y^{-1}\alpha\prod_{i=1}^{n} x_i^{\alpha_i}. \tag{4.16}$$

Obviously, setting $D(x,y) = 1$ in (4.15) and (4.16) and inverting in y yields the explicit production function (CES) and (CD).

[25]Note the "self-duality" of this structure, a concept formulated by Houthakker [1965] in the context of dual consumer preferences: the cost-function structure in prices mirrors the CES/Cobb-Douglas structure of the production function in input quantities.

These dual structures, necessary and sufficient for constancy of the MES, are easily modified to encompass the cases where the production function is homogeneous but does not satisfy constant returns to scale: virtually the same proof goes through if we simply replace y by $y^{1/\alpha}$ where $\alpha = \sum_i \alpha_i$ in (4.13)–(4.16). The results are similarly extended to homothetic technologies if we replace y with $\Gamma(y)$ in (4.13)–(4.16), where Γ is an increasing function. In fact, this structure also suffices when there are multiple outputs, in which case Γ is a mapping from $\mathbf{R}^m_+$ into $\mathbf{R}_+$, increasing in each output quantity. Intuitively these simple extensions of the basic result reflect the fact that all of the structure implied by constancy of the elasticity of substitution is imbedded in the "shape" (or "curvature") of an isoquant, and as long as there is no change in its shape as we move from one isoquant to another, the basic structure is preserved.[26] The next subsection examines the possibilities when the shape of the isoquant does change when outputs change—i.e., when the technology is not homothetic.

4.4.5 Non-Homothetic Technologies.

Blackorby and Russell [1975, 1981] generalized the representation result for MES to encompass the case of non-homothetic technologies. Their proof first characterizes the constancy of the MES in terms of the structure of the the cost function. In particular, the MES are constant if and only if the cost function has the following structure:

$$C(p,y) = \Gamma(y)\left(\sum_{i=1}^{n} \alpha_i(y)\ p_i^{\hat{\rho}}\right)^{1/\hat{\rho}}, \quad 0 \neq \hat{\rho} \leq 1, \tag{4.17}$$

or

$$C(p,y) = \Gamma(y)\prod_{i=1}^{n} p_i^{\beta_i}, \quad \beta_i > 0\ \forall\ i, \quad \sum_{i=1}^{n} \beta_i = 1, \tag{4.18}$$

where, denoting the range of F by $\mathcal{R}(F)$, $\Gamma : \mathcal{R}(F) \to \mathbf{R}_{++}$, and $\alpha_i : \mathcal{R}(F) \to \mathbf{R}^n_{++},\ i = 1,\ldots,n$, are increasing functions.

Thus, the basic CES/Cobb-Douglas structure is preserved when we expand the set of allowable technologies to be non-homothetic. An important

[26] I am unaware of similar explorations of possible generalizations of the results on constancy of the Allen-Uzawa elasticities, but intuition suggests that similar results would go through there as well.

difference, however, is the dependence on output y of the "distribution coefficients," $\alpha_i(y), i = 1, \ldots, n$ in the CES structure in (4.17).[27] This additional flexibility allows the isoquants to "bend" differently for different output vectors while keeping constant the curvature of the isoquant.

The distance function dual to (4.17) is

$$D(y, p) = \Gamma(y)^{-1} \left(\sum_{i=1}^{n} \alpha_i(u)^{-1} x_i^{\rho} \right)^{1/\rho}, \quad \rho = \frac{\hat{\rho}}{\hat{\rho} - 1}. \tag{4.19}$$

Owing to the dependence of the distribution coefficients, $\alpha_i(y), i = 1, \ldots n$, on y the explicit production function cannot be derived in closed form in the case of a single output. In the Cobb-Douglas case, however, the dual distance function is

$$D(x, y) = \Gamma(y)^{-1} \alpha \prod_{i=1}^{n} x_i^{\alpha_i}, \tag{4.20}$$

which, in the case of a single output can be set equal to one and inverted in y to obtain the explicit production function:

$$y = \Gamma^{-1} \left(\alpha \prod_{i=1}^{n} x_i^{\alpha_i} \right).$$

As pointed out by Blackorby and Russell [1981], if the cost structure is given by (4.17)–(4.18), the Allen-Uzawa elasticities are constant and equal to the Morishima elasticities. As far as I know, however, necessary structural conditions for constancy of the Allen-Uzawa elasticities have not been worked out, though one might expect some variation on (4.17)–(4.18).

4.5 Dual Elasticities of Substitution.

4.5.1 Two-Input Elasticity of Substitution Redux.

Let us return briefly to the discussion of the two-variable elasticity of substitution in Section 4.2, where the elasticity is defined as the log derivative of the quantity ratio with respect to the technical rate of substitution:

[27] Blackorby and Russell [1981] proved that the dependence on y of the corresponding coefficients, $\beta_i, \ i = 1, \ldots n$, in (4.18) leads to a violation of positive monotonicity of the cost function in y. Thus, generalization to non-homothetic technologies does not expand the Cobb-Douglas technologies consistent with constancy of the MES.

$\sigma = \phi'(\ln trs_{\ell,k})$. The inverse of ϕ', with image $\sigma^d := \theta'(x_k/x_\ell)$, is the log derivative of a technical rate of substitution with respect to a quantity ratio. This is also an elasticity, one that is dual to σ. In contrast to the (direct) elasticity of substitution σ, large values of σ^d reflect "difficult" substitution, or *strong complementarity*, whereas low values reflect "easy" substitution, or *weak complementarity*. Following the lead of Sato and Koizumi [1973], several papers refer to this concept as an *elasticity of complementarity.*[28] In what follows, I use the terms "dual elasticity of substitution" and "elasticity of complementarity" interchangeably.[29]

Of course, in the two-input case σ and σ^d convey the same information about the curvature of the isoquant and the degree of substitutability (or complementarity) between the two inputs. In the remainder of this section, I extend this dual concept to technologies with more than two inputs, paralleling the development of Allen and Morishima (direct) elasticities in the previous section. The development is facilitated by the use of duality theory, since the $trs_{\ell,k}$ can be interpreted as the relative shadow price of inputs ℓ and k and in fact is equal to the market price ratio, p_ℓ/p_k, under conditions of competitive market pricing.

4.5.2 Dual Morishima and Allen Elasticities of Substitution.

The dual Morishima elasticity of substitution (Blackorby and Russell [1975, 1981]) is given by

$$\sigma_{ij}^{DM}(x,y) : = \frac{\partial \ln \Big(\hat{\rho}_i(x^i,y) \;/\; \hat{\rho}_j(x^i,y)\Big)}{\partial \ln(x_j/x_i)} \tag{4.21}$$

$$= x_j \Big(\frac{D_{ij}(x,y)}{D_i(x,y)} - \frac{D_{jj}(x,y)}{D_j(x,y)} \Big) \tag{4.22}$$

$$= \epsilon_{ij}^D(x,y) - \epsilon_{jj}^D(x,y), \tag{4.23}$$

[28] See Bertoletti [2005], Kim [2000], and Stern [2011a].

[29] Yet another possible assignation is "shadow elasticity of substitution," since this dual concept is formulated in terms of shadow prices.

where x^i is the $(n-1)$-dimensional vector of input quantity ratios with x_i in the denominator and

$$\epsilon_{ij}^D(x,y) = \frac{\partial \ln \rho_i(x,y)}{\partial \ln x_j}$$

is the (constant-output) elasticity of the shadow price of input i with respect to changes in the quantity of input j.

Analogously, Blackorby and Russell [1981] proposed the Allen elasticity of complementarity (alternatively, the dual Allen elasticity of substitution):

$$\sigma_{ij}^{DA}(x,y) = \frac{D(x,y)D_{ij}(x,y)}{D_i(x,y)D_j(x,y)} \tag{4.24}$$

$$= \frac{\epsilon_{ij}^D(x,y)}{s_j^D(x,y)}, \tag{4.25}$$

where

$$s_j^D(x,y) = \rho_j(x,y) \cdot x_j$$

is the cost share of input j (assuming cost-minimizing behavior).

If $\sigma_{ij}^{DA}(p,y) < 0$ (that is, if increasing the j^{th} quantity decreases the shadow price of input i), we say that inputs i and j are Allen-Uzawa dual substitutes; if $\sigma_{ij}^{DA}(p,y) > 0$, they are Allen-Uzawa dual complements. Similarly, if $\sigma_{ij}^{DM}(p,y) < 0$ (that is, if increasing the j^{th} quantity decreases the shadow price of input i relative to the shadow price of input j), we say that input j is a dual Morishima substitute for input i; if $\sigma_{ij}^{DM}(p,y) > 0$, input j is a dual Morishima complement to input i.

Recall from Subsection 4.5.1 that in the two-input case the elasticities of substitution and complementarity are simple inverses of one another. This is clearly not the case when $n > 2$.[30] Interestingly, since the distance function is concave in x, and hence $\epsilon_{jj}^D(x,y)$ in (4.23) is non-positive, the Morishima elasticity leans more toward dual complementarity than does the Allen elasticity (in sharp contrast to the primal taxonomy in Subsection 4.4.3). Similarly, if two inputs are dual Allen-Uzawa complements, they must be dual Morishima complements, whereas two inputs can be dual Allen-Uzawa substitutes but dual Morishima complements.

[30] Of course, the Allen and Morishima elasticities of complementarity are identical when $n = 2$, as is the case with Allen and Morishima elasticities of substitution.

There exist, of course, dual comparative-static results linking factor cost shares and elasticities of complementarity.[31] Consider first the effect of quantity changes on absolute shares (for $i \neq j$):

$$\frac{\partial \ln s_i^D(x,y)}{\partial \ln x_j} = \epsilon_{ij}^D(x,y) = \sigma_{ij}^{DA}(x,y)\ s^D(x,y),$$

so that an increase in x_j increases the absolute share of input i if and only if $\epsilon_{ij}^D(x,y) > 0$ or, equivalently, $\sigma_{ij}^{DA}(x,y) > 0$; that is, if and only if inputs i and j are dual Allen-Uzawa complements. Thus, the dual elasticities provide immediate qualitative comparative-static information about the effect of quantity changes on (absolute) shares. To obtain quantitative comparative-static information, one needs to know the share of the j^{th} input as well as the Allen-Uzawa elasticity of complementarity. Of course, the (constant-output) elasticity derived from the distance function $\epsilon_{ij}^D(x,y)$ yields the same (qualitative and quantitative) comparative-static information.

Comparative-static information about relative income shares in the face of quantity changes can be extracted from the Morishima elasticity. As the share functions, $s_i, i = 1, \ldots n$, are homogeneous of degree zero in quantities, we can re-write their images as $\tilde{s}_i(x^i, y) := s_i(x, y)$. We then obtain

$$\frac{\partial \ln\left(\tilde{s}_i(x^i,y)/\tilde{s}_j(x^i,y)\right)}{\partial \ln(x_j/x_i)} = \epsilon_{ij}^{DM}(x,y) - \epsilon_{jj}^{DM}(x,y) - 1 = \sigma_{ij}^{DM}(x,y) - 1.$$

Thus, an increase in x_j increases the share of input i relative to input j if and only if

$$\sigma_{ij}^{DM}(x,y) > 1;$$

that is, if and only if inputs i and j are sufficiently complementary in terms of the dual Morishima elasticity of complementarity. Moreover, the degree of departure from unity provides immediate quantitative information about the

[31] While shadow prices and dual elasticities are well defined even if the input requirement sets are not convex, the comparative statics of income shares using these elasticities requires convexity (as well, of course, as price-taking, cost-minimizing behavior), which implies concavity of the distance function in x. By way of contrast, convexity of input requirement sets is not required for the comparative statics of income shares using dual elasticities, since the cost function is necessarily concave in prices. See Russell [1997] for a discussion of these issues.

effect on the relative factor share. Thus, the dual Morishima elasticities provide immediate quantitative and qualitative comparative-static information about the effect of quantity changes on relative shares.

As pointed out by Blackorby and Russell [1981, p.153], constancy of the dual Morishima elasticities of substitution entails precisely the same restrictions on the production technology as does constancy of the MES elasticities. This is because the required structure of the cost function and the distance function is self dual, as can be seen by inspection of (4.13)-(4.14) and (4.19)–(4.20).

4.5.3 Symmetric Elasticity of Complementarity.

Stern [2010] points out that the dual Morishima elasticity does not reflect the curvature of the isoquant. While the log derivative in (4.23) holds output quantities y constant in assessing the effect on the shadow-price ratio of a change in the quantity ratio—changing only the jth quantity—it does not maintain $D(y, x) = 1$. Consequently, the direction of the differential change in the quantity ratio is not consistent with containment in the y-isoquant.

Stern defines the *symmetric elasticity of complementary* as follows:

$$\begin{aligned} \sigma_{ij}^{SEC}(x, y) &= \left. \frac{\partial \ln \left(\hat{\rho}_i(x^i, y) \ / \ \hat{\rho}_j(x^i, y) \right)}{\partial \ln(x_j / x_i)} \right|_{D(x,y)=1} \\ &= \frac{\Psi(x, y)}{\Gamma(x, y)} \end{aligned}$$

where

$$\Psi(x, y) = -\frac{D_{ii}(x, y)}{D_i(x, y)^2} + 2\frac{D_{ij}(x, y)}{D_i(x, y) D_j(x, y)} - \frac{D_{jj}(x, y)}{D_j(x, y)^2}$$

and

$$\Gamma(x, y) = 1/D_i(x, y)x_i + 1/D_j(x, y)x_j.$$

This elasticity is symmetric, reflecting the required (differential) movement along the isoquant. Moreover, it can be expressed as a share-weighted average of the Dual Morishima elasticities (Stern [2010]),

$$\sigma^{SEC}(x,y) = \frac{s_i(x,y)}{s_i(x,y)+s_j(x,y)}\sigma_{ij}^{DM}(x,y) + \frac{s_j(x,y)}{s_i(x,y)+s_j(x,y)}\sigma_{ji}^{DM}(x,y), \tag{4.26}$$

or the Dual Allen elasticities (Stern [2011a]),

$$\sigma^{SEC}(x,y) = \frac{s_i(x,y)}{s_i(x,y)+s_j(x,y)}\sigma_{ij}^{DA}(x,y) + \frac{s_j(x,y)}{s_i(x,y)+s_j(x,y)}\sigma_{ji}^{DA}(x,y). \tag{4.27}$$

The asymmetric elasticity of complementarity is dual to the *shadow elasticity of substitution* (McFadden's [1963] and Mundlak [1968]), which is derived by evaluating the derivative of a quantity ratio with respect to a price ratio along a constant-cost frontier:

$$\begin{aligned}\sigma_{ij}^{SES}(p,y) &= \left.\frac{\partial \ln\Big(C_i(p,y)\ /\ C_j(p,y)\Big)}{\partial \ln(p_j/p_i)}\right|_{C(p,y)=y} \\ &= \tilde{\Psi}(p,y)/\tilde{\Gamma}(p,y)\end{aligned}$$

where

$$\tilde{\Psi}(p,y) = -\frac{C_{ii}(p,y)}{C_i(p,y)^2} + 2\frac{C_{ij}(p,y)}{C_i(p,y)C_j(p,y)} - \frac{C_{jj}(p,y)}{D_j(p,y)^2}$$

and

$$\tilde{\Gamma}(p,y) = 1/C_i(p,y)p_i + 1/D_j(p,y)p_j.$$

As shown by Chambers [1988] and Stern [2011a], respectively, the shadow elasticity of substitution can be written as share-weighted averages of Morishima or Allen elasticities, as in the relationships between for dual elasticities of complementarity (4.26) and (4.27).

4.6 Gross Elasticities of Substitution.

The (primal and dual) Allen and Morishima elasticities of substitution are formulated in terms of constant-output demand functions. Their immediate usefulness in studies of the comparative statics of factor incomes is limited to firms that are output constrained or to firms with homothetic technologies (in which case the elasticities are independent of output quantities). This limitation prompted the formulation of elasticities of substitution that incorporate the effects of optimal output adjustments as input prices change. Following Blackorby, Primont, and Russell [2007], I refer to these measurement concepts as *gross* elasticities of substitution, contrasting them to the standard Allen and Morishima elasticities, which assess *net* input quantity change—that is, changes that abstract from the effects of output changes.

The gross analogue of the Allen elasticity, first formulated by Mundlak [1968] using primal production theory methods, was formulated in the dual by Lau [1978] and referred to by Bertoletti [2001, 2005] as the Hotelling-Lau elasticity in his resurrection of this concept. The gross analogue of the Morishima elasticity was proposed by Davis and Shumway [1996] and contrasted with the Hotelling-Lau elasticity by Blackorby, Primont, and Russell [2007] and Syrquin and Hollender [1982].[32] Both of these gross elasticities are most evocatively expressed in terms of the profit function (and in fact are formulated by simply substituting the profit function for the cost function in the Allen elasticity).

Let $r \in \mathbf{R}^n_{++}$ be the vector of output prices, indexed by $k, \ell = 1, \ldots, m$. The profit function, $\Pi : \mathbf{R}^m_{++} \times \mathbf{R}^m_{++}$, is defined by

$$\begin{aligned} \Pi(p,r) &= \max_{x,y}\{r \cdot y - p \cdot x \mid \langle x, y\rangle \in T\} && (4.28)\\ &= r \cdot \phi(p,r) - p \cdot \zeta(p,r), && (4.29) \end{aligned}$$

where ϕ and ζ are the (vector-valued) input-demand and output-supply functions, respectively. The profit function is nondecreasing in r, nonincreasing in p, and convex, jointly continuous, and homogeneous of degree one in $\langle r, p\rangle$. We assume in addition that it is twice continuously differentiable in all prices. The vector of supply functions and the vector of input demand functions are derived from the profit function using the envelope theorem—often termed Hotelling's Lemma in this context:

$$\phi_k(p,r) = \Pi_k(p,r) \ \forall\ k$$

and

$$\zeta_i(p,r) = -\Pi_i(p,r) \ \forall\ i,$$

where subscripts on the profit function Π indicate differentiation with respect to the indicated output or input price.

[32] See also Hicks [1970] Sato and Koizumi [1973], and Stern [2011a] for discussions of these issues.

The extension of the Allen elasticity to encompass output-quantity changes, as formulated by Lau [1978]—the *Hotelling-Lau elasticity (HLES)*—is given, for inputs i and j, by

$$\sigma_{ij}^{HL}(p,r) = \frac{\Pi(p,r)\ \Pi_{ij}(p,r)}{\Pi_i(p,r)\ \Pi_j(p,r)} = \frac{\Pi(p,r)}{p_j x_j}\frac{\partial \ln \phi_i(p,r)}{\partial \ln p_j}.$$

Blackorby, Primont, and Russell [2007] demonstrate that the HLES inherits the inadequacies of the AES, from which it is constructed by analogy. They summarize their evaluation as follows (page 206): The HLES "is not a logarithmic derivative of a quantity ratio with respect to a price ratio—allowing output to change, and it does not provide comparative static content about relative factor incomes. In fact it is not even a generalization of the AUES in any meaningful sense, since it does not reduce to the latter under the assumption of homotheticity."

Blackorby, Primont, and Russell go on to construct a gross version of the MES that rectifies the problems with the HLES. Note that the optimal input ratio can be written as

$$\ln\left(\frac{\zeta_i(p,r)}{\zeta_j(p,r)}\right) = \ln\left(\frac{-\Pi_i(p,r)}{-\Pi_j(p,r)}\right) = \ln\left(\frac{\Pi_i(p,r)}{\Pi_j(p,r)}\right), \tag{4.30}$$

where the second term is an application of Hotelling's Lemma. To differentitate the ratio in (4.30) with respect to the log of p_i/p_j, note that, using homogeneity of degree one of Π in all prices,

$$\Pi(p,r) = p_j \Pi^j(p^{-j}/p_j, r/p_j), \tag{4.31}$$

where p^{-j} is the $(n-1)$-dimensional vector of price ratios with p_j purged from p. Similarly, owing to homogeneity of degree zero of the demand functions,

$$\zeta_i(p,r) = \phi_i^j(p^{-j}/p_j, r/p_j)\ \forall i.$$

Application of Hotelling's Lemma to (4.31) yields the Morishima gross elasticity of substitution:

$$\begin{aligned}
\sigma^{MG}(p,r) &= \frac{\partial[\ln(\zeta_i^j(p^{-j}/p_j, r/p_j)/\ln \zeta_j^j(p^{-j}/p_j, r/p_j)]}{\partial \ln(p_i/p_j)} && (4.32)\\
&= p_j\Big(\frac{\Pi_{ij}(p,y)}{\Pi_j(p,r)} - \frac{\Pi_{jj}(p,r)}{\Pi_j(p,y)}\Big). && (4.33)\\
&= \epsilon_{ij}^*(p,r) - \epsilon_{jj}^*(p,r), && (4.34)
\end{aligned}$$

where $\epsilon^*_{ij}(p,r)$ is the (gross) cross elasticity of demand for input i with respect to the jth price and $\epsilon^*_{jj}(p,r)$ is the own (gross) price elasticity of demand for input j. Thus, analogously to the MES the MGES is simply the difference between the appropriate (gross) cross elasticity of input demand and the (gross) own elasticity of the input associated with the j^{th} axis, along which the price ratio is being varied.

By construction, the MGES is a derivative of an optimal input-quantity ratio with respect to the relevant input price ratio when outputs as well as inputs are allowed to adjust. This elasticity, moreover, provides immediate information about the comparative statics of factor income shares. Define the relative shares of inputs i and j:

$$s_{ij}(p,r) = \frac{p_i\ \zeta_i(p,r)}{p_j\ \zeta_j(p,r)}$$

or, in terms of price ratios,

$$\hat{s}_{ij}(p^{-j}/p_j, r/p_j) = \frac{p_i\ \hat{\zeta}_i(p^{-j}/p_j, r/p_j)}{p_j\ \hat{\zeta}_j(p^{-j}/p_j, r/p_j)}.$$

Some tedious but straightforward calculations yield

$$\frac{\partial \hat{s}_{ij}(p^{-j}/p_j, r/p_j)}{\partial \ln(p_i/p_j)} = 1 - \sigma^{MG}_{ij}(p^j, r).$$

That is, an increase in the price of input i relative to the price of input j (actually, holding p_j fixed) increases the share of input i relative to input j if and only if $\sigma^{MG}_{ij}(p^j,r) < 1$. Thus, the Morishima gross elasticity, unlike the Hotelling-Lau elasticity, yields immediate (qualitative and quantitative) comparative static information about the effect of changes in relative input prices on the relative factor income shares. As shown by Blackorby, Primont and Russell [2007], the MGES reduces to the MES when the technology is homothetic, whereas the Hicks-Lau elasticity does not collapse to the Allen elasticity under this restriction.

4.7 Elasticities of Substitution and Separability.

Technological separability—independence of technical rates of substitution of a subset of pairs of inputs or outputs from the quantities of inputs or outputs not belonging to this subset—is a powerful restriction rationalizing the existence of aggregate input or output quantities and the decentralization of optimization problems (*e.g.*, output-constrained cost minimization).[33] Application of the concept to dual representations of the technology—independence of a set of dual marginal rates of substitution (or complementarity) from price levels of inputs or outputs not in the set—have dual implications for (price) aggregation and decentralization.[34]

As was first noticed by Berndt and Christensen [1973a], certain identity restrictions on the Allen elasticities of substitution are equivalent, under some strong regularity conditions (principally homotheticity), to some corresponding separability restrictions on the technology. Russell [1975] and Blackorby and Russell [1976] generalized these results for the AES and then extended them to encompass the Morishima elasticites in Blackorby and Russell [1981]. The requisite technological restrictions take the form of separability conditions for the dual cost function (or, equivalently, the distance function).

4.7.1 Separability and Functional Structure.

Inputs $\langle i, j \rangle$, say, are separable from input k in the distance function if

$$\frac{\partial}{\partial x_k}\left(\frac{D_i(y,x)}{D_j(y,x)}\right) = \frac{\partial}{\partial x_k}\left(\frac{\rho_i(y,x)}{\rho_j(y,x)}\right) = 0. \tag{4.35}$$

That is, the technical rate of substitution of input i for input j, given the output level y, is independent of the quantity employed of input k. In the

[33] The concept was independently conceived by Leontief [1947a, 1947b] and Sono [1945, 1961]. See Blackorby, Primont, and Russell [1978] for a comprehensive development of the concept and its applications and for citations to the literature extending the Leontief-Solo concept.

[34] In fact, the concept is abstract: it can be applied to any (multiple variable) function.

single-output case, this condition is equivalent to[35]

$$\frac{\partial}{\partial x_k}\left(\frac{F_i(x)}{F_j(x)}\right) = 0. \tag{4.36}$$

Similarly, under the assumption of differentiability of the cost function, input prices $\langle i, j\rangle$, are separable from input price k in C if

$$\frac{\partial}{\partial p_k}\left(\frac{C_i(y,p)}{C_j(y,p)}\right) = \frac{\partial}{\partial p_k}\left(\frac{\delta_i(y,p)}{\delta_j(y,p)}\right) = 0. \tag{4.37}$$

That is, the ratio of constant-output demand-function images i and j is independent of input price k for given output y.

Now partition the set of input variable indices $I = \langle 1, \ldots, n\rangle$ into subsets $\mathcal{I} = \{I^1, \ldots, I^S\}$. The corresponding decompositions of the vectors x and p are $x = \langle x^1, \ldots, x^S\rangle$ and $\langle p^1, \ldots, p^S\rangle$. Define the set of triples,

$$\mathcal{I}_S = \{\langle i, j, k\rangle \mid \langle i, j\rangle \in I^r \times I^r \ \wedge \ k \in I^s, \ r \neq s\}$$

and

$$\mathcal{I}_C = \{\langle i, j, k\rangle \mid i \in I^r, \ j \in I^s, \ k \notin I^s \cup I^r\}.$$

We say that $D, \ F$, or C is separable in the partition $\mathcal{I}$ if (quantity or price) variables i and j are separable from k for all $\langle i, j, k\rangle \in \mathcal{I}_S$ and completely separable in the partition $\mathcal{I}$ if variables i and j are separable from k for all $\langle i, j, k\rangle \in \mathcal{I}_C$. That is, $D, \ F$, or C is separable in $\mathcal{I}$ if ratios of derivatives (*i.e.*, trade-offs between) variables in any set belonging to $\mathcal{I}$ are independent of values of variables outside that set, and $D, \ F$, or C is completely separable in $\mathcal{I}$ if ratios of derivatives between any two variables in I are independent of values of variables outside the sets containing the variables in the ratio.

Positing these separability conditions imposes structural restrictions on the functions representing the technology. In particular separability of F in the partition $\mathcal{I}$ holds if and only if the production-function image (in the single-output case) can be written as

$$F(x) = \hat{F}(F^1(x^1), \ldots, F^S(x^S)). \tag{4.38}$$

[35]In the case where $m = 1$, $D(F(x), x) = 1$ on the isoquant for output $F(x)$. Differentiate this identity with respect to x_i and x_j and take the ratio to obtain this equivalence.

The "aggregator" functions, $F^1, \ldots, F^S$, inherit the curvature and monotonicity properties of F and are interpreted as aggregate input quantities. The "macro" function $\hat{F}$ is increasing in its arguments.

The cost function is separable in the partition $\mathcal{I}$ if and only if

$$C(p, y) = \hat{C}(y, C^1(p, y), \ldots, C^S(p, y)) \tag{4.39}$$

and the distance function is separable in the partition $\mathcal{I}$ if and only if

$$D(x, y) = \hat{D}(y, D^1(x, y), \ldots, D^S(x, y)). \tag{4.40}$$

The sectoral cost and distance functions, $D^1, \ldots, D^S$ and $C^1, \ldots, C^S$, inherit the properties of D and C; the macro functions, $\hat{D}$ and $\hat{C}$, are increasing in the aggregator-function images. The structures, (4.39) and (4.40), are self-dual: that is, the structure (4.39) holds if and only if (4.40) holds.[36]

If the technology is homothetic, the aggregator functions in (4.38) can be normalized to be homogeneous of degree one. Moreover, the dual representations, (4.39) and (4.40), simplify as follows:

$$C(p, y) = \Pi(y)\ \tilde{C}(\Pi^1(p), \ldots, \Pi^S(p)) \tag{4.41}$$

and

$$D(x, y) = \Gamma(y)\ \tilde{D}(\Gamma^1(x), \ldots, \Gamma^S(x)), \tag{4.42}$$

where the aggregator functions are homogeneous of degree one and can be interpreted as sectoral price and quantity indexes, Π is an increasing function, and Γ is a decreasing function.

Finally, if the production function satisfies first-degree homogeneity, (4.41) and (4.42) simplify to

$$C(p, y) = y\ \tilde{C}(\Pi^1(p), \ldots, \Pi^S(p)) \tag{4.43}$$

and

$$D(x, y) = y^{-1}\tilde{D}(\Gamma^1(x), \ldots, \Gamma^S(x)). \tag{4.44}$$

Note that, on the frontier, $D(x, y) = 1$ and inversion of (4.44) in y yields the production function (4.38).

[36]Proofs of these and other results in this section can be found in Blackorby, Primont, and Russell [1978].

We say that D, F, or C is completely separable in the partition $\mathcal{I}$ if (quantity or price) variables i and j are separable from k for all $\langle i,j,k\rangle \in \mathcal{I}_C$ and completely separable in the partition $\mathcal{I}$ if variables i and j are separable from k for all $\langle i,j,k\rangle \in \mathcal{I}_C$.

Assume that the partition of the price and quantity variables $\mathcal{I}$ contains more than two groups ($S > 2$).[37] Then the (symmetrically dual) cost and distance functions have the following images if and only if they satisfy complete separability in the partition $\mathcal{I}$:[38]

$$\begin{aligned} C(p,y) &= \bar{C}\left(y, \sum_{s=1}^{S} C^s(p^s,y)\right) && (4.45)\\ &= \Gamma(y)\ \hat{C}\left(y, \sum_{s=1}^{S} C^s(p^s,y)^{\hat{\rho}(y)}\right)^{1/\hat{\rho}(y)}, \quad 0 \neq \hat{\rho}(y) \leq 1, && (4.46)\\ &\quad \text{or}\\ &\quad \Gamma(y) \prod_{s=1}^{S} C^s(p^s,y)^{\beta_s(y)}, \quad \beta(y) > 0\ \forall\ s,\\ &\qquad\qquad \sum_{s=1}^{S} \beta_s(y) = 1, && (4.47) \end{aligned}$$

and

$$\begin{aligned} D(x,y) &= \bar{D}\left(y, \sum_{s=1}^{S} \Gamma^s(x^s,y)\right) && (4.48)\\ &= \Gamma(y)^{-1}\ \hat{D}\left(y, \sum_{s=1}^{S} \Gamma^s(x^s,y)^{\rho(y)}\right)^{1/\rho(y)}, \quad 0 \neq \rho(y) \leq 1, && (4.49)\\ &\quad \text{or}\\ &\quad \Gamma(y)^{-1} \prod_{s=1}^{S} \Gamma^s(x^s,y)^{\beta_s(y)} \quad \beta_s(y) > 0\ \forall\ s,\\ &\qquad\qquad \sum_{s=1}^{S} \beta_s(y) = 1. && (4.50) \end{aligned}$$

[37] Don't ask. Or if you can't resist, I refer you to Section 4.6 of Blackorby, Primont, and Russell [1978] on "Sono independence" and additivity in a binary partition.

[38] Analogously to the case (4.13), $\rho(y) = \hat{\rho}(y)/\big(\hat{\rho}(y) - 1\big)$.

Thus, complete separability results in a dual structure for the cost and distance functions that is CES in the aggregator-function images, $C^1(p^1, y), \ldots, C^S(p^S, y)$ and $D^1(x^1, y), \ldots, D^S(p^S, y)$. These aggregator function images cannot be interpreted, however, as sectoral price and quantity indexes, since they depend on the value of the output vector as well as input-specific prices and quantities. If, however, we conjoin complete separability and the assumption of homotheticity of the technology, the above structure simplifies to

$$\begin{aligned} C(p,y) &= \Gamma(y)\left(\sum_{s=1}^{S} \Pi^s(p^s)^{\hat{\rho}}\right)^{1/\hat{\rho}}, \quad 0 \neq \hat{\rho} \leq 1, && (4.51) \\ & \text{or} \\ & \Gamma(y)\prod_{s=1}^{S} \Pi^s(p^s)^{\beta_s}, \quad \beta_s > 0 \ \forall \ s, \\ & \sum_{s=1}^{S} \beta_s = 1, && (4.52) \end{aligned}$$

and

$$\begin{aligned} D(p,y) &= \Gamma(y)^{-1}\left(\sum_{s=1}^{S} \Lambda^s(x^s)^{\rho}\right)^{1/\rho}, \quad 0 \neq \rho \leq 1, && (4.53) \\ & \text{or} \\ & \Gamma(y)^{-1}\prod_{s=1}^{S} \Lambda^s(x^s)^{\beta_s}, \quad \beta_s > 0 \ \forall \ s, \\ & \sum_{s=1}^{S} \beta_s = 1. && (4.54) \end{aligned}$$

The functions, $\Pi^s(p^s)$ and $\Gamma^s(x^s)$, $s = 1, \ldots, S$, satisfy the salient (monotonicity and homogeneity) properties of price and quantity indexes, respectively.

If $m = 1$ in (4.53) and (4.54), inversion of $D(x, y) = 1$ yields the production function,

$$\begin{aligned} y = F(x) &= \Gamma\left(\left[\sum_{s=1}^{S} \Lambda^s(x^s)^{\rho}\right]^{1/\rho}\right), \quad 0 \neq \rho \leq 1, && (4.55) \\ & \text{or} \\ & \Gamma\left(\prod_{s=1}^{S} \Lambda^s(x^s)^{\beta_s}\right), \quad \beta_s > 0 \ \forall \ s, \end{aligned}$$

$$\sum_{s=1}^{S} \beta_s = 1. \tag{4.56}$$

If, in addition, the production function is homogeneous of degree one, $\Gamma(y) = y$ and

$$\begin{aligned} F(x) &= \left[\sum_{s=1}^{S} \Lambda^s(x^s)^\rho\right]^{1/\rho}, \quad 0 \neq \rho \leq 1, && (4.57) \\ &\text{or} \\ &\prod_{s=1}^{S} \Lambda^s(x^s)^{\beta_s}, \quad \beta_s > 0 \ \forall\ s, \\ &\qquad \sum_{s=1}^{S} \beta_s = 1. && (4.58) \end{aligned}$$

4.7.2 Elasticity Identities and Functional Structure.

Berndt and Christensen [1973] were the first to notice a relationship between functional structure and certain restrictions on the values of (Allen) elasticities of substitution. Maintaining linear homogeneity of a single-output production function and $n > 2$, they showed that

$$\sigma_{ki}^{DM}(p, y) = \sigma_{kj}^{DM}(p, y) \ \forall \langle i, j, k \rangle \in \mathcal{I}_S$$

if and only if F is separable in the partition $\mathcal{I}$. This result was generalized by Diewert [1974], Russell [1975], and Blackorby and Russell [1976], and the latter results were extended to Morishima elasticities by Blackorby and Russell [1981]. These results can be summarized as follows:

The following conditions are equivalent (under the maintained assumption that $n > 2$):

(i) C is separable in the partition $\mathcal{I}$ (structure (4.41)).

(ii) D is separable in the partition $\mathcal{I}$ (structure (4.42)).

(iii) $\sigma_{ki}^{M}(p, y) = \sigma_{kj}^{M}(p, y) \ \forall \langle i, j, k \rangle \in \mathcal{I}_S$.

(iv) $\sigma_{ki}^{A}(p, y) = \sigma_{kj}^{A}(p, y) \ \forall \langle i, j, k \rangle \in \mathcal{I}_S$.

(v) $\sigma_{ki}^{DM}(p, y) = \sigma_{kj}^{DM}(p, y) \ \forall \langle i, j, k \rangle \in \mathcal{I}_S$.

(vi) $\sigma_{ki}^{DA}(p, y) = \sigma_{kj}^{DA}(p, y) \ \forall \langle i, j, k \rangle \in \mathcal{I}_S$.

That is, separability of C or D is equivalent to identity of both Allen and Morishima elasticities between all variables *within* a separable sector and all variables *outside* that sector.

The following conditions are equivalent (under the maintained assumption that $n > 2$):

(i) C is completely separable in the partition $\mathcal{I}$ (structure (4.46) if $S > 2$).

(ii) D is completely separable in the partition $\mathcal{I}$ (structure (4.49) if $S > 2$).

(iii) $\sigma^M_{ki}(p,y) = \sigma^M_{kj}(p,y)\ \forall \langle i,j,k\rangle \in \mathcal{I}_C$.

(iv) $\sigma^A_{ki}(p,y) = \sigma^A_{kj}(p,y)\ \forall \langle i,j,k\rangle \in \mathcal{I}_C$.

(v) $\sigma^{DM}_{ki}(p,y) = \sigma^{DM}_{kj}(p,y)\ \forall \langle i,j,k\rangle \in \mathcal{I}_C$.

(vi) $\sigma^{DA}_{ki}(p,y) = \sigma^{DA}_{kj}(p,y)\ \forall \langle i,j,k\rangle \in \mathcal{I}_C$.

That is, strict separability of C or D is equivalent to identity of both Allen and Morishima elasticities between all variables in *any two* sectors and all variables *outside* those sectors.

These results provide powerful tools for hypothesis testing because tests for separability—*i.e.*, for aggregate inputs or outputs—are equivalent to tests for certain equality conditions for pairs of elasticities.

4.8 Concluding Remarks.

The elasticity of substitution concept grew out of the interest of prominent English economic theorists, at the time of the Great Depression, in the distribution of income between capital and labor. The concept surged to prominence with the SMAC characterization of constant elasticity of substitution production functions and the emergence of modern growth theory in the 1960s. The elasticity of substitution between labor and capital turns out to be fundamental to many theoretical aspects of economic growth, including the possibility of perpetual growth or decline, the growth of per capita income, and the speed of convergence to an equilibrium growth path. The elasticity of substitution is especially salient in the insightful analysis of the growth process by Klump and de la Grandville [2000]. See Chirinko [2008] for a discussion of these issues and references to the relevant literature.

Most of the growth theory literature, relying originally on the historical constancy of labor and capital income shares, features the Cobb-Douglas production function.[39] But beginning about 1980, the labor share began to fall, and an accumulation of empirical evidence has indicated that the (aggregate) elasticity of substitution is significantly less than one (see Chirinko [2008] and Chirinko and Mallick [2014]).[40] Although the CES has only one more parameter than the Cobb-Douglas, its greater flexibility seems to be more attuned to the evidence.

The first empirical estimation of elasticities of substitution for more than two inputs (to my knowledge[41] is that of Griliches [1969], who estimated Allen elasticities to assess the effect of increases in human capital (reflected in years of educational attainment) on the relative wages of skilled and unskilled labor, with capital as a third important input.[42] Shortly therafter, Parks [1971] estimated a translog production function to obtain estimated Allen elasticities for five inputs (capital, labor, and three material inputs).

Another early estimation of Allen elasticities, as well as implementation of the aggregation theorems in Section 4.7 and Berndt and Christensen [1973a], is the test in Berndt and Christensen [1973b] for the existence of an aggregate capital stock comprising equipment and structures in a production technology using labor as well as the two types of capital.

Allen elasticities for technologies with more than two inputs play an important role in the research on energy economics, beginning with the classic KLEM (capital, labor, energy, and materials) paper of Berndt and Wood [1975]. Their paper employs the theorems described in Section 4.7 to test for the existence of a value-added production function—that is for separability of labor and capital inputs from material inputs. Thompson and Taylor [1995] follow up on the analysis of these issues using Morishima elasticities.

[39]In fact, as first pointed out by Antras [2004], the pre-1980 constancy of income shares does not imply unitary elasticity of substitution when one takes into account the empirical evidence of aggregate labor-saving technological change, which would tend to increase the share of capital, offsetting its declining share owing to a increasing capital intensity and an elasticity of substitution below one.

[40]Karabarbounis and Neiman [2014] estimate an elasticity of substitution greater than one, but Acemoglu [2014] argues that their use of cross-country data makes their estimates more likely to correspond to endogenous-technology elasticities.

[41]Some of this discussion is based on a working paper by Mundra and Russell [2004].

[42]Follow-ups of the Griliches study can be found in Johnson [1970], Kugler, Müller and Sheldon [1989], and Welch [1970].

Elasticities of substitution with more than two inputs have also played an important role in the assessment of the substitutability of (multiple) monetary assets. Barnett, Fisher, and Serletis [1992] estimated Allen elasticities, whereas Davis and Gauger [1996] and Ewis and Fischer [1984] each employed Morishima elasticities.

Finally, Allen elasticities play an important role in the empirical study of the effect of immigration on the relative wages of domestic and immigrant labor (Grossman [1982], Borjas [1994], and Borjas, Freeman, and Katz [1992, 1996]) and of the effect of the increase in the number of guest workers on resident and non-resident labor (Kohli [1999]).

The only empirical estimation of dual elasticities of substitution of which I am aware is in the study by Mundra [2013] of the substitutability of resident and non-resident (guest) labor (and other inputs). The paper compares estimates of primal and dual elasticities.

Without doubt, many other empirical applications of elasticities of substitution have escaped my attention.

The focus of this chapter has been on the elasticity of substitution between (or among) inputs, but the concept is equally relevant to *output* substitution. That is, for each of the elasticities defined in Sections 4.2, 4.4, and 4.5, one can formulate a corresponding (primal or dual) output elasticity of substitution—a characterization of the curvature of the output possibility curve (or surface).[43]

The elasticity of substitution also shows up in utility theory, where it reflects the ease of substitution between consumer goods (and characterizes the curvature of indifference surfaces). In fact, it was in utility theory that the CES function made its first appearance in the economics literature, when Burk (Bergson) [1936] showed that additivity of the utility function and linearity of Engel curves implies that the utility function belongs to the CES family, referred to as the "Bergson family" in consumer theory. The elasticity of substitution in intertemporal utility functions plays an important role in macroeconomic theory (Hall [1988] and in optimal growth theory (Cass [1965], and Koopmans [1965]). Finally, the CES utility function has proved useful in the study of optimal product diversity in the context of monopolistic

[43]See, *e.g.*, the analysis of the substitutability between a "good" and a "bad" output (in this case, electricity and sulphur dioxide) in Färe, Grosskopf, Noh, and Weber [2005].

competition (Dixit and Stiglitz [1977]).

Consistently with the theme of this volume, the chapter has focused primarily on the theoretical development of the measurement of substitutability: primal and dual characterizations and their close relationships to separable sectors of a production or utility function. The taxonomy for n-variable elasticities implicit in the discussions in Sections 4.4–4.6 dichotomizes elasticity-of-substitution concepts along the following lines: partial vs. ratio elasticities (Allen vs. Morishima elasticities), direct vs. dual elasticities (quantities vs. (shadow) prices as "endogenous" variables), and net vs. gross elasticities (fixed output [or technological homotheticity] vs. variable output). Other elasticity-of-substitution concepts have been proposed, but I see them as variations on these themes.[44]

References.

Acemoglu, D. [2015], "The Rise and Decline of General Laws of Capitalism," *Journal of Economic Perspectives* 29: 3–28.

Allen, R. G. D. [1938], *Mathematical Analysis for Economists*, London: Macmillan.

Allen, R. G. D., and J. R. Hicks [1934], "A Reconsideration of the Theory of Value, II," *Economica* 1, N.S. 196–219.

Antras, P. [2004], "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," *Contributions to Macroeconomics* 4: 1–34.

Arrow, K. J., H. P. Chenery, B. S. Minhaus, and R. M. Solow (SMAC) [1961], "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics* 63: 225–250.

Barnett, W. M., D. Fisher, and A. Serletis [1992], "Consumer Theory and the Demand for Money," *Journal of Economic Literature* 30: 2086–2119.

[44]This may be an unfair oversimplification: Stern [2011a], building on Mundlak [1968], proposes a related but somewhat different and more comprehensive taxonomy of the elasticities. (Nevertheless, I'm reminded of a (private) comment made by a prominent social choice theorist back in the heyday of research in his area: "The problem with social choice theory is that there are more axioms than there are ideas." Well, perhaps we have reached the point where there are more elasticity-of-substitution concepts than there are ideas.)

Berndt, E. R., and L. Christensen [1973a], "The Internal Structure of Functional Relationships: Separability, Substitution, and Aggregation," *Review of Economic Studies* 40: 403–410.

Berndt, E. R., and L. Christensen [1973b], "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929–68," *Journal of Econometrics* 1: 81–114.

Berndt, E. R., and D. O. Wood [1975], "Technology, Prices, and the Derived Demand for Energy," *The Review of Economics and Statistics* 57: 259–268.

Bertoletti, P. [2001], "The Allen/Uzawa Elasticity of Substitution as a Measure of Gross Input Substitutability, " *Rivista Italiana Degli Economisti* 6: 87–94.

Bertoletti, P. [2005], "Elasticities of Substitution and Complementarity A Synthesis," *Journal of Productivity Analysis* 24:183–196.

Blackorby, C., and R. R. Russell [1975], "The Partial Elasticity of Substitution," Discussion Paper No. 75-1, Department of Economics, University of California, San Diego.

Blackorby, C., and R. R. Russell [1976], "Functional Structure and the Allen Partial Elasticities of Substitution: An Application of Duality Theory," *Review of Economic Studies* 43: 285–292.

Blackorby, C., and R. R. Russell [1981], "The Morishima Elasticity of Substitution: Symmetry, Constancy, Separability, and Its Relationship to the Hicks and Allen Elasticities," *Review of Economic Studies* 48: 147–158.

Blackorby, C., and R. R. Russell [1989], "Will the Real Elasticity of Substitution Please Stand Up? A Comparison of the Allen/Uzawa and Morishima Elasticities," *American Economic Review* 79: 882–888.

Blackorby, C., D. Primont, and R. R. Russell [1978], *Duality, Separability, and Functional Structure: Theory and Economic Applications*, New York: North-Holland.

Blackorby, C., D. Primont, and R. R. Russell [2007], "The Morishima Gross Elasticity of Substitution," *Journal of Productivity Analysis* 28: 203–208.

Borjas, G. J. [1994], "The Economics of Immigration," *Journal of Economic Literature* 32: 1667–1717.

Borjas, G. J., R. B. Freeman, and L. F. Katz [1992], "On the Labor Market Effects of Immigration and Trade," in G. Borjas and R. Freeman (eds.), *Immigration and The Work Force*, Chicago: University of Chicago Press.

Borjas, G. J., R. B. Freeman, and L. F. Katz [1996], "Searching for the Effect of Immigration on the Labor Market," *AEA Papers and Proceedings* 8: 246–251.

Burk (Bergson), A. [1936], "Real Income, Expenditure Proportionality, and Frisch's 'New Methods of Measuring Marginal Utility'," *Review of Economic Studies* 4: 33–52.

Cass, D. [1965], "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies* 32: 233–240.

Chambers, R. [1988], *Applied Production Analysis*, Cambridge University Press, Cambridge.

Chirinko, R. S. [2008], "σ: the Long and Short of It," *Journal of Macroeconomics* 30: 671–686.

Chirinko, R. S., and D. Mallick [2017], "The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model," *American Economic Journal: Macroeconomics* 9: 225–253.

Cobb, C. W., and P. H. Douglas [1928], "A Theory of Production," *American Economic Review* 18: 23–34.

Cornes, R. [1992], *Duality and Modern Economics*, Cambridge: Cambridge University Press.

Davis, G. C., and J. Gauger [1996], "Measuring Substitution in Monetary-Asset Demand Systems," *Journal of Business and Economic Statistics* 14: 203–209.

Davis G.C, and C. R. Shumway [1996], "To Tell the Truth about Interpreting the Morishima Elasticity of Substitution," *Canadian Journal of Agricultural Economics* 44: 173–182.

de la Grandville, O. [1997], "Curvature and the Elasticity of Substitution: Straightening It Out," *Journal of Economics* 66: 23–34.

Diewert, W. E. [1974], "Applications of Duality Theory," in *Frontiers of Quantitative Economics*, Vol. 2, edited by M. Intriligator and D. Kendrick, Amsterdam: North-Holland.

Diewert, W. E. [1982], "Duality Approaches to Microeconomic Theory," in *Handbook of Mathematical Economics*, Vol. II, edited by K. Arrow and M. Intriligator, New York: North-Holland.

Dixit, A. K., and J. E. Stiglitz [1977], "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67: 297–308.

Ewis, N. A., and D. Fischer [1984], "The Translog Utility Function and the Demand for Money in the United States," *Journal of Money, Credit, and Banking* 16: 34–52.

Färe, R., S. Grosskopf, D-W Noh, and W. Weber [2005], "Characteristics of a Polluting Technology: Theory and Practice," *Journal of Econometrics* 126: 469–492.

Färe, R., and D. Primont [1995], *Multi-Output Production and Duality Theory: Theory and Applications*, Boston: Kluwer Academic Press.

Fuss, M., and D. McFadden (eds.) [1978], *Production Economics; A Dual Approach to Theory and Applications*, Amsterdam: North-Holland.

Griliches, Z. [1969], Capital-Skill Complementarity, *Review of Economic and Statistics* 51: 465–468.

Grossman, J. [1982], "The Substitutability of Natives and Immigrants in Production," *Review of Economics and Statistics* 64: 596–603.

Hall, R. E. [1988], "Intertemporal Substitution in Consumption," *Journal of Political Economy* 96: 339–357.

Hardy, G. H., J. E. Littlewood, and G. Pólya [1934], *Inequalities*, Cambridge University Press, Cambridge.

Hicks, J. R. [1932], *The Theory of Wages*, MacMillan Press.

Hicks, J. R. [1970], "Elasticity of Substitution Again: Substitutes and Complements," *Oxford Economic Papers* 22: 289–296.

Hicks, J. R., and R. G. D. Allen [1934], "A Reconsideration of the Theory of Value, Part II," *Economica* 1, N.S. 196–219.

Houthakker, H,. S [1965], "A Note on Self-Dual Preferences," *Econometrica* 33: 797–801.

Johnson, G. E. [1970], "The Demand for Labor by Educational Category," *Southern Economic Journal* 37: 190–204.

Karabarbounis, L., and B. Neiman [2014], "The Global Decline of the Labor Share," *The Quarterly Journal of Economics* 129: 61–103.

Kohli, U. [1999], "Trade and Migration: A Production Theory Approach," in *Migration: The Controversies and the Evidence*, edited by R. Faini, J. de Melo, and K. F. Zimmermann, Cambridge: Cambridge University Press.

Kim, H. Y. [2000], "The Antonelli versus Hicks Elasticity of Complementarity and Inverse Input Demand Systems," *Australian Economic Papers* 39: 245-261.

Klump, R., and O. de la Grandville [2000], "Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions," *American Economic Review* 90: 282–291.

Koopmans, T. C. [1965], "On the Concept of Optimal Economic Growth," in *The Economicric Approach to Development Planning*, Amsterdam, North Holland (for Pontificia Academic Science).

Kuga, K. [1979], On the Symmetry of Robinson Elasticities of Substitution: the General Case," *Review of Economic Studies* 46: 527–531.

Kuga, K., and T. Murota [1972], "A Note on Definition of Elasticity of Substitution," *Macroeconomica* 24: 285–290.

Kugler, P., U Müller, and G. Sheldon [1989], "Non-Neutral Technical Change, Capital, White-Collar and Blue-Collar Labor," *Economics Letters* 31: 91–94.

Lau, L. [1978], "Applications of Profit Functions," in M. Fuss and D. McFadden (eds.), *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam: North-Holland, 133–216.

Leontief, W. W. [1947a], "A Note on the Interrelation of Subsets of Independent Variables of a Continuous Function with Continuous First Derivatives," *Bulletin of the American Mathematical Society* 53: 343–350.

Leontief, W. W. [1947b], "Introduction to a Theory of the Internal Structure of Functional Relationships," *Econometrica* 15: 361–373.

Leontief, W. W. [1953], "Domestic Production and Foreign Trade: The American Capital Position Re-examined," *Proceedings of the American Philosophical Society* 97: 331–349.

Lerner A.P. [1933], "Notes on the Elasticity of Substitution II: the Diagrammatical Representation," *Review of Economic Studies* 1: 68–70.

McFadden, D. [1963], "Constant Elasticity of Substitution Production Functions," *Review of Economic Studies* 30: 73–83.

Morishima, M. [1967], "A Few Suggestions on the Theory of Elasticity" (in Japanese), *Keizai Hyoron* (*Economic Review*) 16: 144–150.

Mundlak, Y. [1968], "Elasticities of Substitution and the Theory of Derived Demand," *Review of Economic Studies* 35: 225–236.

Mundra, K. [2013], "Direct and Dual Elasticities of Substitution under Nonhomogeneous Technology and Nonparametric Distribution," *Indian Growth and Development Review* 6: 204–218.

Mundra, K., and R. R. Russell [2004], "Dual Elasticities of Substitution," Discussion Paper 01-26, Department of Economics, University of California, Riverside.

Murota, T. [1977], "On the Symmetry of Robinson Elasticities of Substitution: A Three-Factor Case," *Review of Economic Studies* 42:79–85.

Parks, R. W. [1971], "Price Responsiveness of Factor Utilization in Swedish Manufacturing, 1870–1950," *Review of Economic Studies*, 53: 129–139.

Pigou A.C. [1934], "The Elasticity of Substitution," *Economic Journal* 44: 23–241.

Robinson, J. [1933], *Economics of Imperfect Competition*, London: MacMillan.

Russell, R. R. [1976], "Functional Separability and Partial Elasticities of Substitution," *Review of Economic Studies* 42: 79–85.

Russell, R. R. [1997], "Distance Functions in Consumer and Producer Theory," Essay 1 in Färe, R., S. Grosskopf, *Index Number Theory: Essays in Honor of Sten Malmquist*, Boston: Kluwer Academic Publishers, 7–90 .

Sato, R., and T. Koizumi [1973], "On the Elasticities of Substitution and Complementarity," *Oxford Economic Papers* 25: 44–56.

Solow, R. M. [1956], "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 65: 65–94.

Sono, M. [1945, 1961], "The Effect of Price Changes on the Demand and Supply of Separable Goods," *International Economic Review* 2: 239–271 (Originally published in Japanese in *Kokumin Keisai Zasshi*, 74: 1–51.).

Stern, D. I. [2010], "Derivation of the Hicks, or Direct, Elasticity of Substitution from the Input Distance Function," *Economic Letters* 108: 349–351.

Stern D. I. [2011a], "Elasticities of Substitution and Complementarity," *Journal of Productivity Analysis* 36: 79-89.

Stern D. I. [2011b], "The Role of Energy in Economic Growth," *Annals of the New York Academy of Sciences*, 1219: 26–51.

Syrquin, M, and G. Hollender [1982], "Elasticities of Substitution and Complementarity: the General Case," *Oxford Economic Papers* 34: 515–519.

Thompson, P., and T. G. Taylor [1995], "The Capital-Energy Substitutability Debate: A New Look," *Review of Economics and Statistics* 77: 565-569.

Uzawa, H. [1962], "Production Functions with Constant Elasticities of Substitution," *Review of Economic Studies* 29: 291–299.

Welch, F. [1970], "Education in Production," *Journal of Political Economy* 78: 35–59.