INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2005-07

Aggregate Substitutabilities between Factor Demands

E. MALINVAUD[1]

[1] INSEE-CREST. Collège de France.
Encouragement from Guy Laroque is here acknowledged.

Résumé.

Ceci constitue une présentation de, et des compléments apportés à, deux articles publiés dans les *Annales d'Economie et Statistique* sur le même sujet, qui consiste à élucider un problème d'agrégation. En effet les réponses à diverses questions posées en macroéconomie appliquée dépendent des valeurs qu'ont les élasticités de substitution entre facteurs de production. Concernant les demandes de facteurs venant de toutes les branches productives, ces élasticités reflètent non seulement les possibilités de substitution entre facteurs dans les branches mais aussi l'intensité des substitutions entre demandes des biens produits. Ces dernières substitutions sont mises en jeu par les changements dans les prix relatifs des biens, eux-mêmes induits par les changements dans les prix relatifs des facteurs. Un modèle doué d'une certaine généralité conduit à la relation qui lie directement les prix des facteurs à leurs demandes après agrégation de toutes les branches dont elles viennent. La relation est appliquée à des spécifications particulières : d'abord le modèle d'une économie dite CES (à élasticité de substitution constante dans chaque branche), puis à une économie à trois facteurs avec des fonctions de production CES emboitées. Le traitement complet du cas où n'interviennent que deux facteurs est enfin rappelé. Les compléments par rapport aux articles des *Annales* concernent premièrement un effet interne au secteur productif où, à demandes de biens données, l'agrégation des branches se traduit par une certaine atténuation des substituabilités entre facteurs, phénomène qui contrecarre plus ou moins l'effet des substitutions entre biens. Ils concernent aussi les économies CES, où l'hétérogénéité des élasticités de substitution entre branches est considérée conjointement avec l'hétérogénéité des contenus en facteurs. Enfin est examinée la conjecture selon laquelle l'agrégation n'introduirait que des effets du second ordre par rapport à l'hétérogénéité des contenus en facteurs.

Abstract.

Answers to various questions of applied macroeconomics depend on the values given to elasticities of substitution between the demands for factors, these demands being aggregated across productive sectors. Such elasticities reflect not only substitutabilities within productive processes, but also substitutabilities in the system of demand functions for goods. The latter substitutabilities show how demand for goods react to changes in relative prices of goods, which are themselves induced by changes in factor prices. A fairly general model leads to a relationship directly linking changes in factor prices to changes in the aggregate demands for these factors. The relation is applied to particular specifications : first, to a simple CES economy and, second, to a nested-CES economy with three factors, a case which points to difficulties for the establishment of general transparent properties. A full treatment is, however, provided for the case of two factors only, which may be further specified. Thus, the model exhibited here lends itself to studies about how, at the aggregate level, factor substitutabilities depend not only on the forms of production functions and demand functions for goods, but also on heterogeneity across the set of parameters characterizing these functions.

1. Introduction

Answers to various questions of applied macroeconomics depend on the values given to elasticities of substitution between the demands for various factors of production. For instance, among the determinants of changes in the skill premium in wages, the respective evolutions of the skilled and unskilled labor supplies play a major role in economies operating near full employment. How much of the increase in the skill premium can be so explained by a change in the labor market obviously depends on the elasticity of substitution between the demands for skilled and unskilled labor. But each one of these two demands comes from all industries, a fact which makes measurement of the elasticity somewhat tricky. As D. Acemoglu (2002) writes, "the elasticity of substitution σ [between skilled and unskilled labor] is important for the behavior of the skilled premium when supply changes… Unfortunately this parameter is rather difficult to estimate, since it refers to an elasticity of substitution that combines substitution both within and across industries" (p. 20).

This article reports results of an analytical investigation into the form of the combination just mentioned. The nature of the problem may be explained as follows. When the relative prices of factors of production change, the relative prices of goods also change because production of various goods uses different factor input mix : prices rise for those goods, *A* say, which use much of factors the prices of which are rising, *a* say. Because of substitutions between demands for various goods, the demands for goods *A* will tend to decrease, hence also their production. This in turn will tend to depress the demand for factor *a*. Hence, there appears to be two reasons for a reverse association between relative prices of various factors and their relative consumption levels, both a direct reason (factors *a* will be less used as inputs in production units) and an indirect reason (goods *A* will be less in demand from final users).

Indeed, the existence of these two channels was already stressed in the 1960s by several teachers of macroeconomics. They often claimed that effects of the indirect channel had to be added to those of the direct channel, so that factor substitutabilities were larger at the aggregate level than within production units. But, as we shall see, this familiar conjecture does not necessarily follow. While heterogeneity in factor contents across industries increases, the effect of the indirect channel increases, but simultaneously the effect of the direct channel is likely to decrease : substitutabilities between factors in the production sector tend to be attenuated. This effect, which is not intuitive, will be investigated and explained.

We are here faced with an aggregation problem, which may be significant in a number of contexts, for instance when estimates of factor elasticities of substitution drawn from econometric analysis of a sample of microdata are used in macroeconomic assessments. In order to avoid misunderstanding, it must be stressed from the start that aggregation is here defined as operating across goods and industries, the list of factors being untouched.

An analytical exploration of this aggregation problem makes sense because it leads us to disentangle the respective roles of various links which matter when we move from consideration of the substitution between two given factors in production units to consideration of the substitution between the same two factors at the aggregate level. Aggregation across industries, which use these two factors in different proportions and possibly with different elasticities of substitution, may appear fairly straightforward for what concerns the direct channel, except for the often ignored pitfall mentionned above. On its part the argument concerning the indirect channel brings in substitutability between goods as well as heterogeneity of the input mix across industries.

Two phases in the investigation of the problem by the author were already reported in French (E. Malinvaud, 2002 and 2005). The model explored in the first article specified a CES economy with only two factors of production. The second article, largely superseding the first one, took advantage of an approach developed by Alain Bernard[1]. It is here the object of an abridged presentation which does not repeat the analytical derivations, but only their major steps. A few important features are also more precisely described, as a result of recent deeper analyses.

The next section presents the model used for the investigation. Section 3 shows how far the analysis can go when only the first channel is considered. Section 4 leads to the general formula linking, through a square matrix, relative changes in normalized factor prices and relative changes in factor supplies within an "isoquant set" [2], so characterizing factor substitutabilities at the aggregate level. The possibility of a quadratic approximation of this matrix with respect to heterogeneity of factor contents[3] across industries is then evoked. We say speak of "the substitutability matrix".

The following sections apply the general formula to particular cases in which its interpretation turns out to be simpler and illuminating. Section 5 deals with the log-linear economy and with a familly of CES neighboring economies in which the matrix in question may or not be directly related to the covariance matrix of factor contents (across industries). Although this family is quite special, heterogeneity in elasticities of substitution across industries appears as potentially significant, besides heterogeneity in factor contents. Further comments concerning cases with more than two factors are presented in section 6. In particular the study of a nested-CES economy, exhibits complexities which may result from the multidimensionality of factor substitutabilities within industries.

Section 7 deals with the, otherwise general, case in which there are just two factors. In any industry or at the aggregate level, substitutabilities within the neighborhood of an equilibrium are characterized by a single number, an elasticity of substitution. The aggregate elasticity depends on both substitutabilities between goods and heterogeneity of factor contents across industries. Section 8 characterizes more precisely this dependence for the CES system of demands for goods which is often used in macroeconomics. This case moreover provides an opportunity for better understanding why the aggregate elasticity of substitution between the two factors may be lower than the corresponding industry elasticities. The origin of this attenuation phenomenon turns out to lie in a requirement of the equilibrium between quantities, to which the price system has to adapt. Such a requirement is overlooked in the familiar conjecture according to which the aggregate elasticity of substitution between the two factors could not be smaller than the corresponding industry elasticities. Section 9 comments on results applying for a case in which the system of demand functions for goods is not homothetic.

---

[1] Independently Alain Bernard, a French engineer economist, had previously derived the main formula given in my first article. He explained to me how to apply his more general approach to the study of the problem. This led to the second article.

[2] Definition of the isoquant set at an equilibrium will be given at the end of section 2.

[3] Factor contents of a good will be defined in the next section as the factor cost shares in the production of the good, by equations (3).

2. A model

There are n industries. Each industry h (h = 1, 2…n) produces output y^h of good h from inputs x_i^h of m factors (i = 1, 2…m) with a constant-returns-to-scale technology[4]. Markets for goods and factors are assumed perfectly competitive. At equilibrium p_i is the price of factor i and q_h that of good h. Production functions f^h are twice differentiable and involve no intermediate input.

The system of demand functions for goods is written as :

$$\ln y_h = g^h(\ln q_1, \ln q_2 ... \ln q_n\ ; \ln r) \qquad h = 1, 2...n \tag{1}$$

with income r given by :

$$r = \sum_{h=1}^{n} q_h\ y_h \tag{2}$$

The system is assumed to identically fufil this budget equation. Functions g^h are differentiable, g_k^h being the derivative with respect to $\ln q_k$ and g_r^h the derivative with respect to $\ln r$. In what follows the n-dimensional square matrix C will be understood as having elements $c_{hk} = -g_k^h$ and the column vector e as having elements $e_h = g_r^h$.

The specification so chosen deliberately assumes all buyers-users of goods to be aggregated in a single agent. Besides prices, demands for goods depend only on aggregate income r. Thus the analysis disregards distribution effects which could follow from changes in exogenous variables (the only such changes considered here will concern aggregate supplies x_i^S of factors). Together with absence of intermediate inputs, this neglect of distribution effects is a second substantial assumption. But given the object of the present investigation, both seem to be admissible. In particular neglect of distribution effects on the demand for goods concerns only the indirect channel and is unlikely to have a substantial impact on the qualitative results of this paper. The difficulties here faced for analysis of the direct channel will give a flavor of those that analysis of distribution effects on the demands for goods would involve.

Market shares and cost shares in the reference equilibrium will be denoted by v_h and w_i^h :

$$v_h = \frac{q_h y_h}{r} \qquad w_i^h = \frac{p_i x_i^h}{q_h y_h} \tag{3}$$

[4] The hypothesis of constant returns to scale is very convenient in analytical derivations. It was avoided in E. Malinvaud (2002), where decreasing returns to scale were also admitted. From that study I draw the conclusion that constant-returns-to-scale results are sufficiently informative in most respects.

The vector of the w_i^h, for *i* running from 1 to *m*, will be called *the vector of factor contents* of good *h*. The average cost share will naturally be :

$$\overline{w}_i = \sum_{h=1}^{n} v_h w_i^h \tag{4}$$

The budget identity requires on the e_h and c_{hk} :

$$\sum_h v_h e_h = 1 \tag{5}$$

and existence of a number γ such that

$$\sum_h v_h c_{hk} = \gamma v_k \qquad \text{for all } k \tag{6}$$

Since our purpose is a comparative static analysis, we shall consider changes from one equilibrium to another neighboring equilibrium. We shall focus on relative differentials denoted as :

$$X_i^h = \frac{dx_i^h}{x_i^h} \qquad P_i = \frac{dp_i}{p_i} \qquad Y_h = \frac{dy_h}{y_h} \qquad Q_h = \frac{dq_h}{q_h} \tag{7}$$

Price normalization will be equivalently given by one or the other of the following two equations :

$$\sum_{h=1}^{n} v_h Q_h = 0 \qquad \sum_{i=1}^{m} \overline{w}_i \; P_i = 0 \tag{8}$$

In the following analysis *the isoquant set of industry h* will be characterized by the equality :

$$Y_h = 0 \quad \text{or equivalently} \quad \sum_{i=1}^{m} w_i^h X_i^h = 0 \tag{9}$$

<u>3. Conditional aggregate substitutabilities</u>

Substitutabilities in industry *h* will be represented by a square *m*-dimensional matrix A^h which is such that :

$$w_i^h \; X_i^h = \sum_j a_{ij}^h \; P_j + w_i^h \; Y_h \tag{10}$$

This agrees with *the standard representation of substitutabilities in the theory of production*, which focuses on the alternative input combinations leading to a given output (hence here $Y_h = 0$). From such a beginning we shall be led to a representation of aggregate substitutabilities, which will similarly fit within common concepts in the theory of production[5].

Analysis shows that, given profit maximization in industry h, matrix A^h is derived from a matrix B^h by solution of :

$$A^h B^h + w^h \ {}^tU = I \qquad\qquad A^h U = 0 \qquad\qquad (11)$$

where w^h and U are the m-column vectors with respective components equal to w_i^h and all to 1 and ${}^tU$ is the transposed raw vector of U. Analysis also shows how the elements b_{ij}^h of B^b are derived from the first and second derivates of the production function f^h :

$$b_{ij}^h = \frac{f^h f_{ij}^h}{f_i^h f_j^h} \qquad \text{or equivalently} \qquad b_{ij}^h = \frac{x_i^h \ f_{ij}^h}{f_j^h} \bigg/ \frac{x_i^h f_i^h}{f^h} \qquad (12)$$

Equation (11) shows that matrix A^h is determined from the matrix B^h and the vector of factor contents w_i^h. Equations (12) show that matrix B^h is symmetrical and that each element b_{ij}^h may be interpreted as the ratio between two elasticities with respect to the factor input x_i^h : the elasticity of the marginal productivity f_j^h with respect to factor j and the elasticity of output, this second elasticity being equal to w_i^h in equilibrium.

Given the constant returns to scale assumption, the function f_j^h is homogeneous of degree zero, so that :

$$\sum_{j=1}^{m} f_{ij} x_j = 0 \qquad \text{implying} \qquad B^h w^h = 0 \qquad\qquad (13)$$

The singular matrix is typically of rank m – 1. Similarly[6], the matrix A^h is symmetrical, singular and of rank m – 1.

It will be worth remembering subsequently that, whereas the vector of factor contents depends on just the first derivatives of the production function f^h, definitions of matrices B^h and A^h combine first and second order derivatives of the production function.

Before we deal with the full range of factor substitutions allowed by our model, it is convenient to isolate those operating within industries only (the direct channel), i.e. to those that are conditional on given changes $Y_1 ... Y_n$ in all industries. After elementary operations, equation (10) then leads to[7] :

---

[5] The choice of the proper concepts for elasticities of substitution has been much discussed in the theory of consumption. See references in A. Bernard (1986) or D.R. Helm (1987). But it is not our concern here.

[6] See in Malinvaud (2005) how A^h is a diagonal submatrix of the inverse of a symmetrical (m+1) dimensional matrix.

[7] See the derivation of equation (23) from equation (30) in Malinvaud (2005).

$$\overline{w}_i X_i^S = \sum_j a_{ij}^S \; P_j + \sum_h v_h w_i^h Y_h \tag{14}$$

where X_i^S is the relative change in aggregate input of factor *i*

$$x_i^S = \sum_h x_i^h \tag{15}$$

and the matrix A^S is defined by :

$$A^S = \sum_h v_h A^h \tag{16}$$

It may be called "*the conditional substitutability matrix*". It does not depend on substitutabilities between goods and looks like a natural weighted average of the industry substitutability matrices. However, we shall explain in section 5 why this appearance is misleading. Formula (16) actually implies in some interesting cases a form of attenuation of industry substitutabilities.

4. The fundamental equation

Taking in (14) Y_h as given was convenient in order to provide a benchmark, but is otherwise untenable as an hypothesis in our model. Clearly Y_h is endogenous because the demand function for good *h* depends on the prices of goods as well as on income. Hence, indirectly from these two sources, Y_h also depends on factor prices. Taking these relations into account, we may derive a set of *m* equations linking directly the *m* variations in input demands X_i^S to the *m* variations in input prices P_j. But, in order to keep speaking of substitutabilities we must also take into account the hypothesis that variations in input demands are bounded to remain within the isoquant set.

In our model the system of demand functions and the fact that the price of each good must be equal to its factor cost lead to :

$$Y_h = e_h \frac{dr}{r} - \sum_{kj} c_{hk} w_j^k P_j \tag{17}$$

Since moreover *dr* = 0 results from the condition that variations take place within the isoquant set, (14) leads to :

$$\overline{w}_i X_i^S = \sum_j [a_{ij}^S - \sum_{hk} v_h w_i^h c_{hk} w_j^k] P_j \tag{18}$$

Due to (4), (6) and (8) this is equivalent to

$$\overline{w}_i X_i^S = \sum_j [a_{ij}^S - \sum_{hk} v_h z_i^h c_{hk} z_j^k] P_j \tag{19}$$

where

$$z_i^h = w_i^h - \overline{w}_i \tag{20}$$

Equation (19) can also be written in matrix form as :

$$X^S = D_{\overline{w}}^{-1}[A^S - Z\ D_v C\ {}^tZ]P. \tag{21}$$

in which $D_{\overline{w}}$ and D_v are diagonal matrices with respective elements $\overline{w}_i$ and v_h, the $m \times n$ matrix Z has elements z_i^h while ${}^tZ$ is its transposed. Such is *the fundamental equation* ruling aggregate substitutabilities.

In equation (19) the second term in the square bracket characterizes the effect of the second channel. It is a quadratic form of the deviations in factor contents from their means across industries. The *m* quadratic forms appearing in the second part of the square bracket of (21) are so structured by matrix *C*, which caracterizes, in the reference equilibrium, substitutabilities between demands for goods. Those are attractive expressions, given the purpose here. However, interpreting them as the relevant measures of the impact of heterogeneity in factor contents would be misleading, because this heterogeneity also affects the first term in the bracket : in (11) matrix A^h appears as a (non-linear) function of vector w^h and equation (16) shows that matrix A^S combines *n* such functions. Moreover in a special case examined in the next section, A^S contains, besides systematic parts depending only on some mean factor contents, nothing more than other quadratic forms of deviations in factor contents across industries. So, the two groups of quadratic forms have to be combined in order to give the effects of heterogeneity in factor contents.

It is worth noticing that the effect of heterogeneity, in the second term of the bracket in (21), and even in the whole formula for some special cases, so turns out to be second-order small with respect to disparities between industries. Small disparities would so be negligible. We shall ask the question of knowing how general could that property be.

5. Multiple factors in a family of homothetic economies. The attenuation property

No restriction was imposed so far on the reference equilibrium, except implicitly for the condition that the *n* systems (11) had unique solutions. For the definition of special cases it will now be convenient to impose further restrictions on the economies to which the model is applied. In section 7 the number of dimensions of heterogeneity will be drastically reduced by assuming there are just two factors (m = 2).

Now, on the contrary, we drastically simplify the forms of substitutabilities within industries and in the system of demand functions for goods. *We assume the production function f^h to have the CES form (a constant and universal elasticity of substitution σ^h).* We similarly assume that the system of demand functions has a CES form with the elasticity of substitution σ^C. Moreover *income elasticities e_h are all assumed to equal 1,* so that the final demand system is homothetic and the matrix *C* is equal to σ^C multiplying the unit matrix of order n. The family of economies now discussed is tightly restricted. It contains, however, all homothetic log-linear economies as particular cases for which $\sigma^C = \sigma^h = 1$ for all h.

Analytical derivations then lead to :

$$b_{ij}^h = \frac{1}{\sigma^h}\left[1 - \frac{\delta_{ij}}{w_j^h}\right] \qquad a_{ij}^h = \sigma^h\left[w_i^h w_j^h - \delta_{ij} w_j^h\right] \tag{22}$$

where δ_{ij} is the Kronecker indicator equal to 1 if i=j and to 0 otherwise. Application of (16) leads to :

$$a_{ij}^S = -\delta_{ij} \sum_h v_h \; \sigma^h \; w_j^h + \sum_h v_h \; \sigma^h w_i^h \; w_j^h \tag{23}$$

Interpretation of this formula is easy thanks to convenient notations.

Quite naturally we define the mean elasticity of substitution in production as :

$$\sigma^P = \sum_{h=1}^{n} v_h \; \sigma^h \tag{24}$$

where v_h is the market share of good h, as defined by (3) and already used to define by (4) $\overline{w}_i$ the average cost share of factor i. Looking at (23) shows, however, the appearance of another mean value $\hat{w}_i$ of cost shares w_i^h of factor i, in which the weights v_h are replaced by $v_h \; \sigma^h / \sigma^P$:

$$\sigma^P \hat{w}_i = \sum_{h=1}^{n} v_h \; \sigma^h w_j^h \tag{25}$$

The weight given to the cost share w_i^h is all the higher not only as good h is more in demand but also as the elasticity of substitution σ^h is higher in industry h. In the particular case where all σ^h are equal to the same value σ^P, the mean value $\hat{w}_i$ is, of course, equal to $\overline{w}_i$.

In order to interpret equation (23) it is convenient to consider the corresponding mean value of the following product of deviations : $(w_i^h - \hat{w}_i)(w_j^h - \hat{w}_j)$. We shall then use the notation $C\hat{o}v(w_i, w_j)$ defined by :

$$\sigma^P C\hat{o}v(w_i, w_j) = \sum_h v_h \sigma^h (w_i^h - \hat{w}_i)(w_j^h - \hat{w}_j) \tag{26}$$

and write for convenience $V\hat{a}r(w_i) = C\hat{o}v(w_i, w_i)$. Equation (23) may then be written as :

$$a_{ii}^S = -\sigma^P[(1 - \hat{w}_i)\hat{w}_i - V\hat{a}r(w_i)] \tag{27}$$

$$a_{ij}^{S} = \sigma^{P}[\hat{w}_i \hat{w}_j + C\hat{o}v(w_i, w_j)] \qquad \text{if} \quad i \neq j \tag{28}$$

Let us finally note that the second part in the square bracket of (19) is now equal to $\sigma^{C}\ Cov(w_i, w_j)$ with the common definition :

$$Cov(w_i, w_j) = \sum_h v_h z_i^h z_j^h \tag{29}$$

Interpretation of those formulas is interesting in several respects. *Let us begin with the special case where the* σ^h *are all equal to the same* σ^P so that $\hat{w}_i = \overline{w}_i$ and $C\hat{o}v(w_i, w_j) = Cov(w_i, w_j)$. Then the fundamental equation (19) becomes :

$$X_i^S = \frac{\sigma^P - \sigma^C}{\overline{w}_i} \sum_j \ \mathrm{C}ov(w_i, w_j)P_j - \sigma^P P_i \tag{30}$$

We note immediately that, when $\sigma^P = \sigma^C$, hence in particular in the log-linear economies, heterogeneity of factor contents plays no role in the aggregate demands for factor. Relative variations in the demand for factor *i* are proportional to relative variation in the normalized price of this factor, with an elasticity just equal to the elasticity of substitution, which is common to all industries. The somewhat familiar conjecture according to which elasticities of substitution should be higher at the aggregate level than at the industry level is invalidated (we shall presently discuss the explanation of this result).

Generally, except for *i = j*, the matrix of cross-elasticities of the aggregate demands for factors has the same structure as the covariance matrix of factor contents across industries. As soon as *m > 2* description of the matrix of bilateral elasticities of substitution can give rise to as many developments as those that took place, in mathematical statistics, about the description of covariance matrices.

Let us now turn attention to the general case, in which the σ^h differ, and to the interpretation of formulas (27) and (28). In formula (22) when *j = i* the square bracket is equal to $-w_i^h(1-w_i^h)$, hence negative, as soon as $0 < w_i^h < 1$. Equation (16) implies that a_{ii}^S is also negative. In (27) we then see that $0 < V\hat{a}r(w_i) < (1-\hat{w}_i)\hat{w}_i$, as soon as the *n* values w_i^h are not all equal. So, increasing heterogeneity of the w_i^h implies a decrease in the absolute value $-a_{ii}^S$. *This may be interpreted as meaning an attenuation in the effect of a decrease in* p_i *on the input* x_i^S *along the direct channel.*

For $j \neq i$ formula (28) shows that there will be such an attenuation of the negative cross-effect of the same decrease in p_i on the input x_j^S if and only if $C\hat{o}v(w_i, w_j)$ is negative. There is no reason for this to be always the case as soon as *m*>2. Indeed, many combinations of cross-correlations of the w_j^h are conceivable. But, from :

$$\sum_j w_j^h = 1 \qquad \text{hence} \qquad \sum_j \hat{w}_j = 1 \tag{31}$$

we directly conclude :

$$\sum_{j \neq i} C\hat{o}v(w_i, w_j) = -V\hat{a}r(w_i) < 0 \tag{32}$$

In this weak sense attenuation must dominate also in cross-effects along the direct channel.

Heterogeneity in the industry elasticities of substitution σ^h *leads to a remark concerning the idea according to which the effects of heterogeneity in factor contents would be second-order small*. Such is certainly the case when all σ^h are equal to the same number σ^P and equation (30) applies. As it is defined, the set of the $Cov(w_i, w_j)$ is indeed a natural caracterization for the heterogeneity in factor contents. But the set of the $C\hat{o}v(w_i, w_j)$ mixes in its definition heterogeneity of factor contents with heterogeneity of elasticities of substitution. A more transparent expression of a_{ij}^S, for $i \neq j$ for instance, would be equivalent to (28), namely :

$$\sigma^P[\overline{w}_i\overline{w}_j + Cov(z_i, z_j)] + \overline{w}_i\, Cov(\sigma, z_j) + \overline{w}_j Cov(\sigma, z_i) + \sum_h v^h(\sigma^h - \sigma^P) z_i^h\, z_j^h \tag{33}$$

Given this expression, $a_{ij}^S - \sigma^P \overline{w}_i \overline{w}_j$ can be said to be second-order small if the 3*n* variables $\sigma^h - \sigma^P$, z_i^h and z_j^h are jointly small. But if some of the $\sigma^h - \sigma^P$ are not small, $Cov(\sigma, z_i)$ for instance is first-order, not second-order, small in terms of the z_i^h.

This remark also draws our attention *on a limitation of the interpretation given above of the attenuation phenomenon*. Interpreting an increase in $V\hat{a}r(w_i)$ as an increase in heterogeneity makes sense when we focuss attention on matrix A^S. But when there is heterogeneity in the elasticities σ^h, a better measure of heterogeneity in factor contents would be $Var(w_i)$ defined in accordance with (29). The attenuation phenomenon might then disappear, even for the effect of a decrease in p_i on the input x_i^S along the direct channel. This might occur in case of a high negative covariance between the σ^h and the w_i^h.

6. Further comments for the case when there are more than two factors

Limiting attention to CES production functions, as we just did, cannot do justice to the complexity of the set of matrices A^S which may be interesting in applications. So, before we turn to a full investigation of cases involving just two factors, a few additional comments are in order.

Properties of A^S reflect properties of the *n* matrices A^h and, further behind, those of the *n* matrices B^h. These 2*n* + 1 square matrices are all symmetrical. They have rank *m* – 1 with, associated to the zero root, vector *U* for A^S and A^h, and vector w^h for B^h, as shown by the second equation in respectively (11) and (13). Given these equations and the vectors of factor contents w^h, each matrix is determined by the *m*(*m-1*)/2 elements above its diagonal. Whereas the vector w^h is obtained from the first derivatives of the production function f^h, knowledge of the second derivatives is also necessary for the determination of these off-diagonal elements and, through (11) or (13), of the diagonal elements as well.

When realizing that there are *m*(*m-1*)/2 degrees of liberty in the determination of relevant characteristics of industry production functions, we may suspect that, in order to reach strong properties, strongly restrictive hypotheses have to be added to our general model, as was done in the foregoing section. Here we shall limit our further comments about possible difficulties to those suggested by a particular specification. For the purpose we shall assume that m = 3 and *all production functions are nested-CES functions* of the following form, in which, for simplicity, we omit notation of the industry *h*.

$$f(x_1, x_2, x_3) = \left\{\beta_1 x_1^{-\omega} + \left[\beta_2 x_2^{-\chi} + \beta_3 x_3^{-\chi}\right]^{\omega/\chi}\right\}^{-1/\omega} \tag{34}$$

where $\beta_1, \beta_2, \beta_3, \omega$ and χ are positive parameters. Two elasticities of substitution are introduced :

$$\sigma = \frac{1}{1+\omega} \qquad \rho = \frac{1}{1+\chi} \tag{35}$$

The analytical treatment of this case (sketched in the appendix) leads to the following value of matrix A^h :

$$A^h = \begin{bmatrix} -\sigma\, w_1^h (1 - w_1^h) & \sigma\, w_1^h\, w_2^h & \sigma\, w_1^h\, w_3^h \\ \sigma\, w_1^h\, w_2^h & -[\sigma\, w_1^h\, w_2^h + a_{23}^h] & a_{23}^h \\ \sigma\, w_1^h\, w_3^h & a_{23}^h & -[\sigma\, w_1^h\, w_3^h + a_{23}^h] \end{bmatrix} \tag{36}$$

with

$$a_{23}^h = \frac{(\rho - \sigma\, w_1^h) w_2^h\, w_3^h}{1 - w_1^h} \tag{37}$$

where the same values ρ and σ of the two elasticities of substitution are assumed to apply to all industries.

We note that, whereas in (22), all a_{ij}^h were quadratic functions of the factor contents, this is no longer the case here for a_{23}^h, hence also for a_{22}^h and a_{33}^h (except if $\rho = \sigma$, which would give a particular case of that studied in section 5). These values lead to a complex definition of the aggregates a_{ij}^S. For going on with the study of this nested-CES case, it is natural to replace the expression of a_{23}^h by a quadratic approximation in terms of $\overline{w}_1, \overline{w}_2, \overline{w}_3$ and z_1^h, z_2^h, z_3^h as defined by equations (4) and (20). Computations sketched in the appendix lead to expressions such as :

$$\tilde{a}_{11}^S = -\sigma\left[\overline{w}_1(1 - \overline{w}_1) - Var(z_1)\right] \tag{38}$$

$$\tilde{a}_{12}^S = \sigma[\overline{w}_1\overline{w}_2 + Cov(z_1, z_2)] \tag{39}$$

$$\tilde{a}_{23}^S = \sigma[\overline{w}_2\overline{w}_3 + Cov(z_2, z_3)] + \frac{(\rho - \sigma)\alpha_{23}}{(1-\overline{w}_1)} \tag{40}$$

$$\tilde{a}_{22}^S = -\sigma[\overline{w}_2(1-\overline{w}_2) - Var(z_2)] - \frac{(\rho - \sigma)\alpha_{23}}{(1-\overline{w}_1)} \tag{41}$$

where :

$$\alpha_{23} = \overline{w}_2\overline{w}_3 - \left(\frac{\overline{w}_2\overline{w}_3}{\overline{w}_2 + \overline{w}_3}\right)^2 Var\left[\frac{z_2}{\overline{w}_2} - \frac{z_3}{\overline{w}_3}\right] \tag{42}$$

We shall not now indulge into a discussion of these formulas, but just point to *the complexity that crops up as soon as more than two factors are taken into account with less constrained substitutabilities than those used in section 5*. In particular, if there is no prior information on substitutabilities, detecting the effects of heterogeneity in substitutabilities *(m(m-1))/2* degrees of liberty) is likely to be definitely more difficult than detecting the effects of heterogeneity in the (*m*-1) independent factor contents, which are easily measurable.

7. The cases of just two factors

Let us now turn to the family of cases in which $m = 2$, but the reference equilibrium is not subject to any other constraint than those specified in the model. Then it can be proved that system (21) reduces to :

$$\left[\frac{dx_1^S}{x_1^S} - \frac{dx_2^S}{x_2^S}\right] = \sigma^S\left[\frac{dp_2}{p_2} - \frac{dp_1}{p_1}\right] \tag{43}$$

which, together with the price normalization (8), determines the values of P_1 and P_2 which have to follow from any couple of changes in the supplies x_1^S and x_2^S occurring along the isoquant line

$$\overline{w}_1\frac{dx_1^S}{x_1^S} + \overline{w}_2\frac{dx_2^S}{x_2^S} = 0 \tag{44}$$

The number σ^S, clearly the aggregate elasticity of substitution, is then defined by :

$$\overline{w}_1\overline{w}_2\sigma^S = \sum_h v^h w_1^h w_2^h \sigma^h + \sum_{hk} v_h z_1^h c_{hk} z_1^k \tag{45}$$

(Since $z_1^h + z_2^h = 0,$ the double sum could as well be written with z_2 instead of z_1).

Note that in (45) the numbers σ^h and c_{hk} are not taken as parameters but rather as local values of industry elasticities of substitution and derivatives of the demand functions in the reference equilibrium.

Note also that the multipliers of σ^h in the first sum add up to less than $\overline{w}_1\overline{w}_2,$ the difference being $Var(w_1),$ the variance of w_1^h across industries defined as in (29) : such is the pitfall signaled in the introduction and again at the end of section 3. We may say that *aggregation of industries in the competitive general equilibrium has an attenuation effect on their elasticities of substitution.* This, not so intuitive, result is the crux for the solution of the puzzle posed by the familiar conjecture.

In order to stress this point, let us also write the first term in (45) as :

$$[\overline{w}_1\overline{w}_2 - Var(\overline{w}_1)]\sigma^{PS} \tag{46}$$

where σ^{PS} may be interpreted as the conditional aggregate elasticity of substitution in industries.

8. Two factors in a CES economy

An interesting particular case combining the assumptions of sections 5 and 7 will now bring simplicity in the results. It specifies $\sigma^h = \sigma^P$ for all *h*, as well as $e_h = 1$ for all *h* and $C = \sigma^C I,$ so that (45) in (43) is replaced by :

$$\sigma^S = \sigma^P + (\sigma^C - \sigma^P)\frac{Var(w_1)}{\overline{w}_1\overline{w}_2} \tag{47}$$

It shows how substitutabilities in productions (σ^P) and in demands for goods (σ^C) combine their effect with heterogeneity of factor contents : $Var(w_1)/\overline{w}_1\overline{w}_2.$

In order to better explain the error in the familiar conjecture, looking precisely at a still more particular case will suffice. Let us then assume *n = m = 2,* $\sigma^P = 1$ and $\sigma^C = 0$ (assumed complementarity in demands for goods is a limit case to which (47) applies). Let us moreover take $v_1 = v_2 = 0.5$ and particularize the two Cobb-Douglas production functions with :

$$w_1^1 = 0.8, \qquad w_2^1 = 0.2, \qquad w_1^2 = 0.2, \qquad w_2^2 = 0.8$$

hence $\overline{w}_1 = \overline{w}_2 = 0.5,$ and $Var(w_1) = Var(w_2) = 0.09.$

Let us look at a neighborhood of the equilibrium defined by :

$$x_1^S = x_2^S = 50 \qquad y_1 = y_2 = 50 \qquad p_1 = p_2 = q_1 = q_2 = 1$$

Application of (46) in that case, where *C* is the null matrix, leads to $\sigma^S = 0.64.$ The attenuation effect amounts to –0.36. How can we explain it, taking account of the fact that no substitution can occur between demands for goods ?

An infinitesimal change in factor supplies along the isoquant containing the equilibrium in question is given by :

$$\frac{dx_1^S}{x_1^S} = -\varepsilon \qquad \frac{dx_2^S}{x_2^S} = \varepsilon$$

Clearly :

$$\frac{dp_1}{p_1} = \frac{-dp_2}{p_2} = \frac{\varepsilon}{0.64} = \frac{25\varepsilon}{16}$$

Not surprisingly, if ε is positive, the price of the scarcer factor 1 increases. Indeed, it has to increase all the more so as σ^S is smaller than 1. Both industries are led, in the production of fixed outputs, to substitute factor 2 to factor 1 with an elasticity equal to 1. The result is :

$$\frac{dx_1^1}{x_1^1} = \frac{-5}{8}\varepsilon \qquad \frac{dx_2^1}{x_2^1} = \frac{5}{2}\varepsilon \qquad \frac{dx_1^2}{x_1^2} = \frac{-5}{2}\varepsilon \qquad \frac{dx_2^2}{x_2^2} = \frac{5}{8}\varepsilon$$

In order to maintain unchanged outputs in both industries the relative change in the input of factor 1 has to be smaller in sector 1, which has the higher content of this factor. Such is the explanation of the attenuation effect of aggregation across industries. It is implied by equilibrium of input quantities and by the condition that the same input prices apply to all industries.

In this case with $\sigma^C = 0,$ the aggregate elasticity of substitution is equal to the conditional elasticity directly following from application of section 3. In cases where 0 < $\sigma^C < 1,$ the other characteristies of the reference equilibrium remaining unchanged, substitutions between the demand for goods indeed leads to an aggregate elasticity of substitution higher than 0.64 while smaller than 1. The intuition in the familiar conjecture was correct in as much as increases in substitutabilities between goods along the indirect channel would boost aggregate substitutability. The fault of the conjecture was to ignore the attenuation effect, which affects the results of the direct channel as soon as are aggregated industries with heterogeneous factor contents.

9. A non-homothetic system of demands for goods

Stopping the analytical investigation at this point would have made me uneasy : the fully homothetic economies specified in sections 5 and 8 are so particular. Embarrassment would not have come from the assumption of constant returns to scale in production, which was relaxed in E. Malinvaud (2002) and was there not found to be crucial for the main results of the article. But the assumption that Engel curves in the system of demand functions for goods were straight lines from the origin was difficult to accept without reservation as illustrating all implications of the fundamental equation (21) : the distinction between necessities and luxuries is unavoidable in any application of consumption theory.

This is why I wanted to discuss also in E. Malinvaud (2005), for the case of two factors, another specification in which the vector *e* of income elasticities and the matrix *C* of price elasticities would be less special than in the preceding section. I selected for the purpose an analytically convenient generalization of the Stone-Geary model used since the 1950s for a number of econometric estimations of systems of consumer demand functions. According to this specification Engel curves still are straight lines, but not from the origin (R. Stone, 1954).

Analytical derivations require in my French article a five page appendix. The main outcome states that, for application of (43), a more complex formula than (47) gave the aggregate elasticity of substitution. Considering again for simplicity the case in which all σ^h are equal to σ^P, this formula writes :

$$\overline{w}_1\overline{w}_2\sigma^S = \sigma^P[\overline{w}_1\overline{w}_2 - Var(w_1)] + g\sigma^C\left\{M(ew^2) - [M(ew)]^2\right\} \quad (48)$$

where g is a positive number, usually smaller than 1 in applications of the Stone-Geary model, and :

$$M(ew) = \sum_h v_h e_h w_1^h \qquad M(ew^2) = \sum_h v_h e_h (w_1^h)^2 \quad (49)$$

Two dimensions of heterogeneity then appear, concerning respectively factor contents in industries and income elasticities of goods produced by these industries, a fact which is not surprising in itself. The aggregate elasticity of substitution is a quadratic function of the w_1^h for given values of the e_h. But when the two dimensions of heterogeneity are simultaneously considered, σ^S turns out to be a fourth degree polymonial function.

A precise discussion of the formula (48) requires that *M(ew)* and $M(ew^2)$ are developed in terms of centered moments, account being taken of equation (5). The multiplier of $g\sigma^C$ in the right-hand member of (48) then turns out to be :

$$Var(w_1) + Cov(e, w_1^2) - [Cov(e, w_1)]^2 \quad (50)$$

where the covariances are defined as in (29).

Clearly, estimation of the value of this expression, for instance when the two factors would be skilled and unskilled labor whereas a dozen of goods, say, would be distinguished, would require a specific investigation. However, E. Malinvaud (2005) ventures to conclude that the effect on the aggregate elasticity of substitution through the indirect channel is likely to be smaller, but not much smaller, than what would be found by application of (47) with the same values of σ^P and σ^C.

Appendix

Treatment of nested-CES production functions (section 6)

Given the production function (34) and the notation (35) :

$$b_{11} = -(1-w_1)/\sigma w_1 \qquad\qquad b_{12} = b_{13} = b_{21} = b_{31} = 1/\sigma$$
$$b_{22} = -(\sigma w_3 + \rho w_1 w_2)/\sigma\rho w_2(1-w_1)$$
$$b_{23} = b_{32} = (\sigma - \rho w_1)/\sigma\rho(1-w_1)$$
$$b_{33} = -(\sigma w_2 + \rho w_1 w_3)/\sigma\rho w_3(1-w_1)$$

Solution of system (11) leads to :

$$a_{11} = -\sigma w_1(1-w_1) \qquad a_{12} = a_{21} = \sigma w_1\ w_2 \qquad a_{13} = a_{31} = \sigma w_1\ w_3$$
$$a_{22} = -w_2(\rho w_3 + \sigma w_1\ w_2)/(1-w_1)$$
$$a_{23} = a_{32} = (\rho - \sigma w_1)w_2\ w_3/(1-w_1)$$
$$a_{33} = -w_3(\rho w_2 + \sigma w_1 w_3)/(1-w_1)$$

Presence of $(1-w_1)$ in denominators of four elements of the matrix *A* has the effect that these elements are not polynomial functions of w_1, w_2, w_3. But a quadratic approximation of $(1-w_1)^{-1}$ in a neighborhood of $w_1 = \overline{w}_1$ leads to convenient approximations for the elements of matrix *A*. Indeed :

$$(1-w_1)^{-1} \sim 1 + z_1/(1-\overline{w}_1) + z_1^2/(1-\overline{w}_1)^2$$

where $z_1 = w_1 - \overline{w}_1$ is the deviation, in conformity with (20).

Introducing now the index *h* of the industry and considering (36) and (37) of the main text, we write a quadratic approximation of a_{23}^h as :

$$\tilde{a}_{23}^h = C_{23} + L_{23}^h + Q_{23}^h$$

where the common constant C_{23} is given by :

$$C_{23} = \frac{\rho - \sigma\,\overline{w}_1}{1-\overline{w}_1}\overline{w}_2\overline{w}_3$$

the linear term is :

$$L_{23}^h = C_{23}\left[\frac{z_1^h}{1-\overline{w}_1} + \frac{z_2^h}{\overline{w}_2} + \frac{z_3^h}{\overline{w}_3}\right] - \sigma\, \overline{w}_2\overline{w}_3 \frac{z_1^h}{1-\overline{w}_1}$$

and the quadratic term is :

$$Q_{23}^h = \frac{(\rho-\sigma)\overline{w}_2\overline{w}_3}{1-\overline{w}_1}\left[\frac{(z_1^h)^2}{(1-\overline{w}_1)^2} + \frac{z_1^h z_2^h}{(1-\overline{w}_1)\overline{w}_2} + \frac{z_1^h z_3^h}{(1-\overline{w}_1)\overline{w}_3}\right] + C_{23}\frac{z_2^h z_3^h}{\overline{w}_2\overline{w}_3}$$

Applying aggregation (16) to this quadratic approximation is essentially simple, all the more so as aggregation of the linear terms L_{23}^h leads to zero. This is, of course, a result of our assumption that the same two values σ and ρ apply to all industries. If that had not been assumed, the remarks made in the last part of section 5 would apply. They indeed are generally valid, as soon as heterogeneity of elasticities of substitution matters.

Taking into account the fact that $z_1^h + z_2^h + z_3^h = 0$ for all *h*, the most transparent formulas for the quadratic approximations turn out to be given by expressions such as :

$$\tilde{a}_{11}^S = -\sigma\left[\overline{w}_1(1-\overline{w}_1) - Var(z_1)\right]$$

$$\tilde{a}_{12}^S = \sigma\left[\overline{w}_1\overline{w}_2 + Cov(z_1, z_2)\right]$$

$$\tilde{a}_{23}^S = \sigma\left[\overline{w}_2\overline{w}_3 + Cov(z_2, z_3)\right] + \frac{(\rho-\sigma)\alpha_{23}}{(1-\overline{w}_1)}$$

$$\tilde{a}_{22}^S = -\sigma\left[\overline{w}_2(1-\overline{w}_2) - Var(z_2)\right] - \frac{(\rho-\sigma)\alpha_{23}}{(1-\overline{w}_1)}$$

where :

$$\alpha_{23} = \overline{w}_2\overline{w}_3 - \left(\frac{\overline{w}_2\overline{w}_3}{\overline{w}_2+\overline{w}_3}\right)^2 Var\left[\frac{z_2}{\overline{w}_2} - \frac{z_3}{\overline{w}_3}\right]$$

References

D. Acemoglu (2002), « Technical change, inequality and the labor market » , *Journal of Economic Literature*, March, p. 7-72.

A. Bernard (1986), « Mesure et représentation de la substituabilité » , Communication au Séminaire René Roy, Paris.

D.R. Helm (1987), « Elasticity of substitution », *The New Palgrave Dictionary of Economics,* Macmillan, London.

E. Malinvaud (2002), « Sur l'agrégation des demandes de travail non-qualifié », *Annales d'Economie et de Statistique,* N° 66, p. 42-80.

E. Malinvaud (2005), « Substituabilités entre demandes de facteurs induites par les demandes de biens », *Annales d'Economie et de Statistique.*

R. Stone (1954), « Linear expenditure systems and demand analysis : an application to the pattern of British demand », *Economic Journal,* vol. 64, p. 511-27.