Available online at www.sciencedirect.com

SCIENCE DIRECT®

Ecological Economics 51 (2004) 217–223

ECOLOGICAL ECONOMICS

www.elsevier.com/locate/ecolecon

ANALYSIS

Facing the truth about separability: nothing works without energy

Manuel Frondel*, Christoph M. Schmidt

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Hohenzollernstr. 1-3, D-45128 Essen, Germany

Received 15 February 2004; received in revised form 22 April 2004; accepted 24 May 2004
Available online 2 November 2004

Abstract

Separability is a pivotal theoretical and empirical concept in production theory. While the standard definition of separability is primarily motivated by the desire to conceptualize production decisions as a sequential process, the principal purpose of an appropriate concept of separability in empirical work is to justify the omission of variables for which data are either of poor quality or unavailable. This paper demonstrates that this empirical concept needs to be more restrictive than the classical notion of separability is. Therefore, we suggest a novel definition of separability based on cross-price elasticities that has clear empirical content. Because there is ample empirical reason to even doubt the assumption that energy is separable from all other production factors in the relatively mild form of classical separability, energy seems to be an indispensable production factor under separability aspects.

Keywords: Substitution; Cost and production functions

JEL classification: C3; D2

1. Introduction

When modelling factor substitution, e.g., the substitutability of capital and labor, it is generally impossible to focus on the bi- or multivariate relationship between the variables of interest. Two situations may arise, however, which could justify the isolated analysis of these factors. First, the omitted variable might be of limited quantitative relevance to the production process. Energy, for instance, accounted for a negligible share of production cost during the "the golden years" of economic growth in the early post-WW II era (Crafts and Toniolo, 1996). The analysis of the possibilities to substitute capital (K), labor (L), and material inputs (M) would not have been altered by the inclusion or exclusion of energy and its cost. Yet, in the aftermath of the energy crises of the seventies, the production factor energy (E) became nonnegligible and consequently gained prominence in empirical studies.

Since then, a large number of studies have appeared analyzing the issues of substitution and

* Corresponding author.
E-mail address: frondel@rwi-essen.de (M. Frondel).

0921-8009/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolecon.2004.05.005

separability of energy and nonenergy inputs, which are crucial for understanding the macroeconomic impacts of energy price shocks and for evaluating energy policies, including a step-by-step energy tax increase. The urgency of these problems rises and falls with the price of oil, but the question still remains important. Berndt and Wood (1975, 1979), Griffin and Gregory (1976), and Pindyck (1979) are seminal studies; more recent examples are Yuhn (1991) and Ramaiah and Dalal (1996), with the cost shares of energy varying between 1% and 10% for a survey. (See Frondel and Schmidt, 2002, who provide ample empirical evidence that capital-energy complementarity, specifically, appears to be simply a matter of cost shares of capital and energy).

Second, whether a nonnegligible variable, such as energy after the oil crises, is included might be irrelevant for inferences about the ease of substitution between nonenergy inputs. For example, despite omitting the factor energy, estimates of cross-price elasticities, measuring the ease of substitution between nonenergy inputs, may still remain correct. Such a notion of separability is particularly important when the data do not provide information on a nonnegligible input factor, but interest is on the substitutability relations of observable factors. For German manufacturing, for instance, empirical studies investigating the issue of factor substitution typically do not incorporate the factor energy (see, e.g., Rutner, 1984; Stark, 1988; Kugler et al., 1989).

As their justification for this omission, these authors typically invoke a standard notion of separability that has been researched thoroughly in economic production theory, there serving the principal purpose to form a conceptual basis for the idea of sequential decision making. Inadvertently, though, these studies implicitly build on an assumption of $[(K, L, M), E]$ separability, a property focusing on the ease of substitution among nonenergy inputs rather than on sequential decision processes. Such a separability assumption incorporates stronger requirements do than those implied by the standard notion of separability, making it questionable whether this severe restriction on the process is palatable.

Moreover, there is ample reason to doubt the applicability of even the relatively mild standard form of separability. In their classic study, Berndt and Wood (1975), e.g., provide empirical evidence that the similar assumption of $[(K,L), (M,E)]$ separability, i.e., the standard assumption of separability of K and L on the one hand and M and E on the other, is violated for U.S. manufacturing (1947–1971). Because $[(K,L), (M,E)]$ separability is an assumption necessarily required for value-added studies, which exclusively employ the inputs K and L, Berndt and Wood (1975, p. 266) "call into question the reliability of [. . .] factor demand studies for U.S. manufacturing based on [. . .] value-added specification[s]". It clearly transpires from the strata of the discussion that, if we want to understand the role of energy in production, specifically under what conditions it can be omitted safely from the empirical analysis, we need a clear notion of the precise restrictions involved in assuming separability.

This paper theoretically investigates the concept of separability. Throughout, the intuition about separability pursued is that the ease of substitution between two factors should be unaffected by a third factor, from which those factors are assumed to be separable (see Hamermesh, 1993, p. 34). In our theoretical analysis, we provide clarification of the rigid nature of the assumption of separability in empirical applications: First, we discuss theoretically in which sense the ease of substitution between two factors i and j is unaffected by the factor k if both are separable from k according to the classical definition of separability, introduced independently by Leontief (1947) and Sono (1945).

The structure of the classical Leontief–Sono separability conditions is identical in both primal and dual contexts. However, it is demonstrated here that these conditions lead to quite different implications regarding substitution issues. In contrast to the previous literature, we thus distinguish *primal* from *dual* separability: Two factors i and j are primally (dually) separable from factor k if and only if their marginal rate of substitution (their input proportion x_i/x_j) is unaffected by the input level of k (the price of factor k). Yet, characterizing substitution relationships between two factors in such ways is rather unusual in empirical studies. Therefore, we suggest a novel and more restrictive definition of separability that has clear empirical content.

The following section theoretically discusses the standard notion of separability and illuminates the relationship between separability and restrictions on substitution elasticities. In Section 3, we present a novel, more restrictive definition of separability. Section 4 concludes.

2. Separability and substitutability

In the received literature, considerations regarding the separability of production factors have their principal motivation in a theoretical issue, the possibility to conceptualize the optimization of production decisions by stages. The standard concept of separability, introduced by Leontief (1947) and Sono (1945), was employed by Strotz (1957) to analyze two-stage optimization: If separability holds according to the classical definition of separability (given in detail below), factor intensities can first be optimized within each separable subset. Then, optimal intensities can be attained by holding fixed the within-subset intensities and optimizing the between-subset intensities.

For empirical work, though, one would like to determine whether the omission of nonnegligible variables for which data are unavailable is justified. As substitution is the center piece of any empirical study (see, e.g., Messner, 2002), the natural intuition of separability for empirical work is that the ease of substitution among observable factors should be unaffected by the variable omitted. In contrast to the previous literature, we distinguish primal from dual separability due to the different implications for the ease of substitution among the production factors[1]. Neither of these types of separability lends itself to empirical application, though. Therefore, we proceed to develop a more restrictive concept of separability with clear empirical content.

2.1. Primal separability

Quite naturally, empirical studies investigated substitution issues by estimating production functions. Consequently, the notion of primal separability, which is based on production functions, was historically defined first. Along the lines of, for instance Goldmann and Uzawa (1964) and Berndt and Christensen (1973, p. 404), henceforth BC73, two factors i and j of a twice differentiable production function $Y=F(x_1, x_2,\ldots, x_n)$ with nonvanishing first and second partial derivatives are defined here to be primally separable from factor k if and only if

$$\frac{\partial}{\partial x_k}\left(\frac{\frac{\partial F(x_1,..,x_n)}{\partial x_i}}{\frac{\partial F(x_1,..,x_n)}{\partial x_j}}\right)=0$$

$$\Leftrightarrow \frac{\partial^2 F}{\partial x_i \partial x_k}\frac{\partial F}{\partial x_j}-\frac{\partial^2 F}{\partial x_j \partial x_k}\frac{\partial F}{\partial x_i}=0. \quad (1)$$

This Leontief–Sono separability condition, which may locally hold at a point or globally, reads alternatively

$$\frac{\partial}{\partial x_k}\left(\frac{\partial x_j}{\partial x_i}\right)=0, \quad (2)$$

with $-\partial x_j/\partial x_i$ denoting the marginal rate of substitution between i and j. Thus, classical primal separability of factor k from i and j implies that factor k's intensity does not affect the ease of substitution—measured in terms of the marginal rate of substitution—between i and j. However, for empirical work, there are more relevant candidate concepts for measuring the ease of substitution than the marginal rate of substitution: cross-price elasticities, e.g., or Allen's elasticities of substitution (AES), the most prominent measures of substitution.

Moreover, for a linear-homogeneous production function F definition (1) is equivalent to

$$\mathrm{AES}_{ik}:=\frac{\frac{\partial F}{\partial x_i}\frac{\partial F}{\partial x_k}}{F\frac{\partial^2 F}{\partial x_i \partial x_k}}=\frac{\frac{\partial F}{\partial x_j}\frac{\partial F}{\partial x_k}}{F\frac{\partial^2 F}{\partial x_j \partial x_k}}=:\mathrm{AES}_{jk}. \quad (3)$$

That is, the primal separability of factor k from i and j means that the ease of substitution between k and i—in terms of AES—equals that between k and j. However, it does not imply, in terms of this concept, that k does not affect the substitution of i for j or vice versa.

In practice, due to the fact that inputs of production functions may be endogenous, and therefore estimators may be inconsistent, the classical way to overcome these endogeneity problems has been to apply dual cost function approaches (Mundlak, 1996, p. 431). Hence, in addition to definition (2) of primal separability, a dual definition is indispensable.

[1] Contrary to the literature, we do not distinguish strong from weak separability, a terminology coined by Strotz (1957), because the intuition regarding substitution issues is perfectly the same behind both concepts.

2.2. Dual separability

On the basis of a twice differentiable cost function $C(Y, p_1, p_2..., p_u)$ with nonvanishing first and second partial derivatives, two factors i and j are defined to be dually separable from factor k along the lines of, e.g., Blackorby and Russell (1976 p. 286) and BC (73, p. 405) if and only if the Leontief–Sono separability condition is valid:

$$\frac{\partial}{\partial p_k}\left(\frac{\frac{\partial C(Y,p_1,..,p_n)}{\partial p_i}}{\frac{\partial C(Y,p_1,..,p_n)}{\partial p_j}}\right) = 0 \Leftrightarrow$$

$$\frac{\partial^2 C}{\partial p_i \partial p_k}\frac{\partial C}{\partial p_j} - \frac{\partial^2 C}{\partial p_j \partial p_k}\frac{\partial C}{p_i} = 0. \qquad (4)$$

Because the structure of Leontief–Sono condition (4) and primal condition (1) is identical, at first glance, the primal separability of i and j from input k seems to imply dual separability of i and j from k and vice versa. However, these two types of separability are not equivalent. According to Blackorby and Russell (1976, p. 287), the conjunction of both types of separability would require the production function $F(x_1, x_2..., x_n)$ to be homothetic.

Furthermore, with particular respect to the interpretation of the ease of substitution between i and j, both separability definitions differ: Using Shephard's Lemma, $\partial C/\partial p_i = x_i$, the Leontief–Sono separability condition (4) equals

$$\frac{\partial}{\partial p_k}\left(\frac{x_i(p_1, p_2, .., p_n)}{x_j(p_1, p_2, .., p_n)}\right) = 0. \qquad (5)$$

That is, two inputs i and j are dually separable from factor k if and only if the their input proportion x_i/x_j is independent of changes of factor k's price. The ease of substitution between i and j in virtually every empirical study measured in terms of their cross-price elasticities or by AES or the Morishima elasticities of substitution (MES) is not at issue, though.

Moreover, the implications of dual separability so defined for cross-price elasticities and for AES and MES are similar to those of primal separability for AES given by Eq. (3): First, differentiating and multiplying condition (5) by p_k yields

$$\eta_{x_i p_k} := \frac{p_k}{x_i}\frac{\partial x_i}{\partial p_k} = \frac{p_k}{x_j}\frac{\partial x_j}{\partial p_k} =: \eta_{x_j p_k}. \qquad (6)$$

That is, under dual separability condition (5), substitution reactions between i and k on the one hand and j and k on the other are restricted to being equal when captured by cross-price elasticities $\eta_{x_i}p_k$. Second, AES and MES are related to cross-price elasticities by

$$\mathrm{AES}_{x_i p_k} := \eta_{x_i p_k}/s_k \quad \text{and}$$
$$\mathrm{MES}_{x_i p_k} := \eta_{x_i p_k} - \eta_{x_k p_k}, \qquad (7)$$

where s_k denotes the cost share of k. By Eq. (6), AES and MES thus obey similar restrictions under dual separability of i and j from k (see also Blackorby and Russell, 1976, p. 288):

$$\mathrm{AES}_{x_i p_k} = \mathrm{AES}_{x_j p_k} \text{ and } \mathrm{MES}_{x_i p_k} = \mathrm{MES}_{x_j p_k}. \qquad (8)$$

In sum, if two factors i and j are dually separable from factor k, the resulting condition (5) merely implies that the ease of substitution between i and j is unaffected by factor k when this ease is measured on the basis of the input proportion x_i/x_j. However, condition (5) generally does not imply that the ease of substitution is unaffected by factor k when this ease is measured by cross-price elasticities, AES, or MES, i.e., those measures that are employed in empirical substitution studies almost without exception. As a consequence, when empirical analysts, as in numerous studies, invoke the assumption of classical separability to justify the omission of a nonnegligible input factor from their empirical analysis, but then proceed to express their results in terms of, say, AES, they are inadvertently losing their empirical work on an insufficient assumption. Therefore, this paper suggests a new definition of separability with clear empirical content.

3. An empirically oriented approach

We define two factors i and j to be empirically dually separable from factor k if and only if

$$\frac{\partial}{\partial p_k}\eta_{x_i p_j} = 0 \quad \text{and} \quad \frac{\partial}{\partial p_k}\eta_{x_j p_i} = 0, \qquad (9)$$

i.e., if and only if the ease of substitution between i and j, measured by the cross-price elasticities involving both factors, is not affected by the price of factor k. While there was some choice of specific approach, we decided to build our separability definition (9) on the basis of cross-price elasticities, because alternative definitions based on AES or MES are more restrictive than our definition[2]: When substitution relationships are intended to be measured by AES, the requirements given by definition (9) do not assure that the ease of substitution concerning i and j is independent of the price of factor k because

$$\frac{\partial}{\partial p_k}\mathrm{AES}_{x_ip_j} = \frac{\partial}{\partial p_k}(\eta_{x_ip_j}/s_j) = \frac{1}{s_j}\frac{\partial}{\partial p_k}\eta_{x_ip_j} - \eta_{x_ip_j}\frac{1}{s_j^2}\frac{\partial s_j}{\partial p_k} = 0 - \eta_{x_ip_j}\frac{1}{s_j^2}\frac{\partial s_j}{\partial p_k} \neq 0, \quad (10)$$

in general. Rather, beyond both conditions of definition (9), changes in the price of factor k must also not affect both cost shares s_i and s_j to guarantee

$$\frac{\partial}{\partial p_k}\mathrm{AES}_{x_ip_j} = 0 \quad \text{and} \quad \frac{\partial}{\partial p_k}\mathrm{AES}_{x_jp_i} = 0. \quad (11)$$

It is difficult to imagine that this is possible in actual applications. Correspondingly, even if the requirements given by definition (9) do hold, it is

$$\frac{\partial}{\partial p_k}\mathrm{MES}_{x_ip_j} = \frac{\partial}{\partial p_k}\eta_{x_ip_j} - \frac{\partial}{\partial p_k}\eta_{x_jp_j} = 0 - \frac{\partial}{\partial p_k}\eta_{x_jp_j} \neq 0, \quad (12)$$

in general. Similar conditions to those of definitions (9) and (11) will hold for MES only if definition (9) is valid and if, additionally, own-price elasticities of both factors i and j are unaffected by the price of factor k.

Using Shephard's Lemma, $\partial C/\partial p_i = x_i$, and the definitions of $\eta_{x_ip_j}$ and $\eta_{x_jp_i}$ given in Eq. (6), definition (9) may be written alternatively as

$$\frac{\partial^2 C}{\partial p_i \partial p_k}\frac{\partial^2 C}{\partial p_i \partial p_j}\Big/\frac{\partial C}{\partial p_i} = \frac{\partial^3 C}{\partial p_i \partial p_j \partial p_k} \quad \text{and} \quad \frac{\partial^2 C}{\partial p_j \partial p_k}\frac{\partial^2 C}{\partial p_i \partial p_j}\Big/\frac{\partial C}{\partial p_j} = \frac{\partial^3 C}{\partial p_i \partial p_j \partial p_k}. \quad (13)$$

When both these conditions hold, the Leontief–Sono separability condition (4) of standard dual separability is fulfilled, but not vice versa. Hence, the standard Leontief–Sono separability condition (4) is necessary but not sufficient for our dual separability definition (9) to hold and, hence, represents a weaker requirement.

In consequence, omitting the factor energy, e.g., because energy data are unavailable, from any empirical analysis might be unjustified even when classical separability conditions are satisfied. Yet, conversely, the violation of the classical Leontief–Sono separability conditions implies that energy is not empirically dually separable from all other inputs and would already put the issue at rest. In sum, when addressing the question of whether energy can safely be omitted in any empirical analysis of substitution relationships based on cross-price elasticities, researchers are, in fact, forced to assume empirical dual separability, our more restrictive notion of separability.

4. Summary and conclusion

With particular respect to substitution issues, the natural intuition of two factors i and j being separable from a factor k is that this factor k should not affect the ease of substitution among the former (see Hamermesh, 1993, p. 34). According to the classical Leontief–Sono separability conditions, primal (dual) separability of factor i and j from factor k implies that, in primal (dual) approaches, their marginal rate of substitution (their input proportion x_i/x_j) is unaffected by the input level of k (the price of factor k). However, rather than by marginal rates of substitution or input

[2] Furthermore, Frondel (2003) concludes that AES and MES do not provide any economically meaningful information beyond that given by cross-price elasticities, which form the common basis of both AES and MES (see definition (7)).

proportions, the overwhelming majority of empirical substitution studies analyzes the ease of substitution between two factors on the basis of cross-price elasticities, AES, or MES.

In consequence, when empirical analysts, as in numerous studies, invoke the standard assumption of separability to justify the omission of a nonnegligible input factor from their empirical analysis, but then proceed to express their results in terms of, say, AES, these researchers base their empirical work inadvertently on an insufficient assumption. This paper therefore criticizes the standard notion of separability to be of limited relevance for empirical studies– notwithstanding its important role in the conceptual justification of stepwise optimizing decisions in production theory—and suggests a practically more important definition of separability based on cross-price elasticities, which we call *empirical dual separability*.

We define two factors i and j to be empirically dual separable from factor k if and only if both the cross-price elasticities $\eta_{x_i p_j}$ and $\eta_{x_i p_i}$ are unaffected by the price of factor k. This definition incorporates the standard definition of dual separability, but is more restrictive. That means that even if, say, capital (K) and labor (L) were separable from the factor energy (E) according to the standard notion, this would nevertheless not imply that the ease of substitution between K and L in terms of cross-price elasticities remains unaffected by E. Therefore, even if K and L were separable from E in the classical meaning, omitting energy from the data base might be unjustified under empirical aspects.

In sum, when addressing the question of whether energy can safely be omitted in any empirical analysis of substitution relationships among all other production factors, researchers are forced to invoke the assumption of empirical dual separability, our more restrictive notion of separability. Otherwise, researchers generally risk finding, for instance, incorrect cross-price elasticities $\eta_{\mathrm{K}P\mathrm{L}}$ and $\eta_{\mathrm{L}P\mathrm{K}}$ when omitting relevant factors like energy from the analysis, unless K and L are empirically dually separable from E. Finally, because there is ample empirical reason to even doubt the assumption that energy is separable from all other production factors in the relatively mild standard form of separability, energy seems to be an indispensable production factor—amongst other things under the aspect of empirical dual separability.[3]

Acknowledgement

We are grateful to Boris Augurzky, Malte Faber, and an anonymous referee for their helpful comments.

References

Berndt, E.R., Christensen, L.R., 1973. The internal structure of functional relationships: separability, substitution and aggregation. Review of Economic Studies 40, 403–410.

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

Berndt, E.R., Wood, D.O., 1979. Engineering and econometric interpretations of energy–capital complementarity. American Economic Review 69 (3), 342–354.

Blackorby, C., Russell, R., 1976. Functional structure and the Allen partial elasticities of substitution: an application of duality theory. Review of Economic Studies 43, 285–292.

Crafts, N., Toniolo, G., 1996. Economic Growth in Europe Since 1945. Cambridge University Press, New York.

Frondel, M., 2003. Empirical assessment of energy-price policies: the case for cross-price elasticities. Energy Policy 32 (8), 989–1000.

Frondel, M., Schmidt, C.M., 2002. The capital–energy controversy: an artifact of cost shares. The Energy Journal 23 (3), 53–79.

Goldmann, S.M., Uzawa, H., 1964. A note on separability in demand analysis. Econometrica 32, 387–398.

Griffin, J.M., Gregory, P.R., 1976. An intercountry translog model of energy substitution responses. American Economic Review 66, 845–858.

Hamermesh, D.S., 1993. Labor Demand. Princeton University Press, New Jersey.

Kugler, P., Müller, U., Sheldon, G., 1989. Non-neutral technical change, capital, white-collar and blue-collar labor. Economics Letters 31, 91–94.

Leontief, W.W., 1947. Introduction to a theory of the internal structure of functional relationships. Econometrica 15, 361–373.

Messner, F., 2002. Material substitution and path dependence: empirical evidence on the substitution of copper for aluminum. Ecological Economics 42 (1–2), 259–271.

[3] The discipline of thermodynamics emphasizes the absolute necessity of energy for the performance of work: Perpetual motion machines, i.e., devices using a limited amount of energy to continually perform work forever are explicitly precluded by the Second Law of Thermodynamics.

Mundlak, Y., 1996. Production function estimation: reviving the primal. Econometrica 64 (2), 431–438.

Pindyck, R.S., 1979. Interfuel substitution and the industrial demand for energy: an intertemporal comparison. Review of Economics and Statistics 61, 169–179.

Ramaiah, K.C., Dalal, A.J., 1996. Functional separability and the existence of consistent aggregates in U.S. manufacturing: comment. International Economic Review 37, 253–258.

Rutner, D., 1984. Faktorsubstitution in den Produktionssektoren der Bundesrepublik Deutschland: Eine Ökonometrische Analyse des Translog-Modells. Schriften Zur Angewandten Ökonometrie, Band Vol. 12.

Sono, M., 1945. The effect of price changes on the demand and supply of separable goods (in Japanese). Kokumin Keizai Zasshi 74, 1–51. (translated into English in International Economic Review, Vol. 2, 1–51.).

Stark, E., 1988. Faktoreinsatzverhalten im Verarbeitenden Gewerbe: Eine sektorale Analyse unter Verwendung eines Translog-Modells. DIW-Vierteljahresheft. 1/2 - 88.

Strotz, R.H., 1957. The empirical implications of a utility tree. Econometrica 25, 269–280.

Yuhn, K., 1991. Functional separability and the existence of consistent aggregates 12 in U.S. manufacturing. International Economic Review 32, 229–250.