Economics Working Paper Series

2019/004

# On the Benefits of Normalization in Production Functions

Iain P. Embrey

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

# On the Benefits of Normalization in Production Functions

Iain P. Embrey, *Lancaster University* *

**Abstract**

A growing literature demonstrates that production function normalization yields important benefits in both empirical and theoretical macroeconomics. Yet normalization has remained notably absent from the microeconomic literature. This paper introduces the concept of normalization within the microeconomic context, and establishes that it could provide substantial improvements in the interpretability, the tractability, and the applicability of a broad class of economic production functions. Since any production function is merely a specialized interpretation of some decision-maker's utility function, our analyses also generalize to the wider context of utility functions. We therefore conclude that the benefits of normalization may be pervasive.

**Keywords:** Normalization; Production Function; Utility Function; CES Function

**JEL Codes:** B40; C18; C50; D01; D10; D24; D33

*i.embrey@lancaster.ac.uk; Lancaster University Management School, Lancs, UK, LA1 4YX; +44(0)7885607783. ORCID: 0000-0002-9177-2916.
The author would like to thank Kim Kaivanto, Steve Bradley, David Rietzke, Kwok Tong Soo, Olivier Cardi, Geraint Johnes, and John Whittaker for their comments. Any remaining errors are entirely my own. The author would also like to thank the Economic and Social Research Council for supporting this work under grant number ES/P000665/1; there are no further interests to declare.

# 1 Introduction

A Production function is normalized by the explicit selection of a basis for its input vector space. Under the convention that inputs take values in $\mathbb{R}^+$, the set of admissible bases can be written as $\mathbf{\Theta} \boldsymbol{I_n}$, where $\mathbf{\Theta}$ is a *normalization point*, and $\boldsymbol{I}_n$ is the identity matrix of the appropriate dimension. In practice, normalization is simple to implement: we need only substitute $x_i \rightarrow \left(\frac{x_i}{\Theta_i}\right)$ for each input $i$; and it is simple to understand: we are merely selecting an appropriate unit of measurement for each input. However, over the last two decades a growing literature has demonstrated that the apparent simplicity of production function normalization is deceptive. Normalization has enabled important developments in both theoretical and empirical research.

The normalized CES production function was introduced by De La Grandville (1989), in order to prove that aggregate growth is both more rapid and more sustainable when the elasticity of substitution between labour and capital inputs is greater. Normalization now underpins a body of related contributions to the theory of growth (Klump & De La Grandville 2000, Papageorgiou & Saam 2008, Irmen & Klump 2009, Willman & McAdam 2013, Xue & Yip 2013), and it has also been shown to resolve a long-standing contention between the Real Business-Cycle model and the New Keynesian model (Cantore et al. 2014). In empirical macroeconomics, normalization has been shown to improve model convergence (Klump & Saam 2008), and robustness to misspecification (Klump, McAdam & Willman 2007, León-ledesma et al. 2010, León-Ledesma, Mcadam & Willman 2015). In light of the increasing recognition of production function normalization in key areas of macroeconomics, it is remarkable that normalization appears to be entirely absent from the microeconomic literature.

This paper establishes that the field of microeconomics could also benefit substantially from production function normalization. In Section 2 we show that, without normalization, CES share parameters are meaningless: they can vary across (0,1) as a function of the measurement units of each input. In Section 3 we derive analytic expressions for these and other CES parameters, and observe that their interpretability is entirely dependent upon normalization. We also outline some important challenges for the estimation of those parameters in the absence of normalization. In the appendices we confirm that these results for CES production functions equally apply to linear, Cobb-Douglas and Leontief forms, by extending the familiar derivations of those forms as special cases of the CES production function to allow for normalization and for nesting. We also provide a new derivation of the translog form as a second-order approximation to the CES function around its normalization point, which is substantially more useful than its familiar derivation as an approximation around the Cobb-Douglas special case, because the normalization point may be chosen freely. In Section 4 we compare the normalized form to its alternatives in the literature, to find (almost paradoxically) that the former is both more general and more tractable than those alternatives.

Our analyses are complementary to those of Klump, Mcadam & Willman (2012), who present a holistic survey of production function normalization in the macroeconomic context. In addition to recommending that normalized production functions should be used as a matter of course, those authors provide a theoretical foundation for the normalization terms $\{\Theta_i\}$: they arise as constants of integration within the derivation of a CES production function. As a corollary to that derivation, we may observe that all CES production functions are normalized, if not explicitly then implicitly at $\boldsymbol{\Theta} = \mathbf{1}$.[1] That observation demonstrates the important truth that normalization does not impose any additional assumptions. In fact normalization relaxes the implicit restriction $\boldsymbol{\Theta} = \mathbf{1}$, which is always present but seldom acknowledged in papers that use 'non-normalized' production functions.

## 2 An Illustration of the Problem

Constrained optimization is fundamental to microeconomic analysis. Indeed, economic science is often defined as the allocation of scarce resources to maximise a decision-maker's utility. Our analyses are relevant to any context in which constrained utility maximization is applied, however for specificity we will focus on the production problem of a firm.

For a simple illustration of the problem, suppose that a firm's objective is to

$$\max_{x_1,x_2} A.\Big(\gamma x_1^{\phi} + (1-\gamma)x_2^{\phi}\Big)^{\frac{1}{\phi}} \quad \Big| \quad p_1x_1 + p_2x_2 \leq R. \tag{1}$$

Here $R$ can be interpreted as the total available resource (net of fixed costs), in terms of money, time, energy, or human capital, and $p_i$ represents the price of one unit of input $x_i$ in terms of that resource. We shall derive precise interpretations for $A$ and $\gamma$ in Section 3, but for now let us follow the modal approach in the literature and interpret them loosely by $A \approx$ total-factor productivity, and $\gamma \approx$ the relative value, or 'share' of input 1. We therefore assume with minimal loss of generality that $x_1, x_2, p_1, p_2, R, A \in \mathbb{R}^+$ and $\gamma \in [0, 1]$.

This firm's CES production function is broadly applicable, since it nests linear, Cobb-Douglas, and Leontief forms as special cases (respectively at $\phi \to 1$; $0$; $-\infty$). We will now illustrate that, in general, the common practice of interpreting $\gamma$ as the share of input 1 can be meaningfully misleading.

Let us suppose that the firm maximizes its profits at some point where the ratio of inputs $x_1 : x_2$ and the ratio of prices $p_1 : p_2$ are given by

[1]It is trivial to extend this observation to any arbitrary production function, by noting that any vector-space must have a basis, and that no basis is inherently superior to any other.

$$\begin{array}{ccc} x_1 & : & x_2 \\ 100\,\text{grams} & : & 1\,\text{unit} \end{array} \qquad\qquad \begin{array}{ccc} p_1 & : & p_2 \\ \$40\,/\text{g} & : & \$4,000\,/\text{unit} \end{array} \tag{2}$$

Given the price ratio from (2) we can solve (1) for the optimal input ratio $x_1 : x_2$ in terms of $\gamma$, and so given the input ratio from (2) we can re-arrange to find the parameter $\gamma$. The first column of Table 1 calculates these values for a range of possible $\phi$. Given (2) we can also find the true factor share of input 1, which is

$$\text{share}_1 = \frac{p_1 x_1}{p_1 x_1 + p_2 x_2} = \frac{1}{2} \tag{3}$$

As expected, the first column of Table 1 shows that the parameter $\gamma$ in system (1) returns the true factor share of input 1 in the Cobb-Douglas special case. However, away from $\phi = 0$, we can see that $\gamma$ gives no directly interpretable indication as to that factor share.

This example is not contrived. Any set of values where the ratio $x_1 : x_2$ is not precisely 1 : 1 will produce the undesirable results shown in Table 1.[2] Of course, given full knowledge of $\phi$ and the input ratio it would be possible to recover the price ratio from $\gamma$ (equation (6) is useful here), and hence to calculate the true factor share, however it would be desirable for $\gamma$ itself to admit some consistent interpretation across the range of possible $\phi$. Column 1 shows that $\gamma$ alone does not even allow us to interpret the sign of the comparison $\text{share}_1$ *cf.* $\text{share}_2$ .

Table 1: *The effect of measurement units and of $\phi$ on the 'share' parameter $\gamma$ in (1).*

| Complementarity | | Using Grams | Using Tonnes |
|---|---|---|---|
| $\phi = 1$ | Linear | $\gamma = \frac{1}{101}$ | $\gamma = \frac{10{,}000}{10{,}001}$ |
| $\phi = \frac{1}{2}$ | | $\gamma = \frac{1}{11}$ | $\gamma = \frac{100}{101}$ |
| $\phi = 0$ | Cobb-Douglas | $\gamma = \frac{1}{2}$ | $\gamma = \frac{1}{2}$ |
| $\phi = -\frac{1}{2}$ | | $\gamma = \frac{10}{11}$ | $\gamma = \frac{1}{101}$ |
| $\phi = -1$ | | $\gamma = \frac{100}{101}$ | $\gamma = \frac{1}{10{,}001}$ |
| $\phi \to -\infty$ | Leontief | $\gamma$ meaningless; always optimal to use: | |
| | | 1gram : 1unit | 1tonne : 1unit |

The value of the 'share' parameter $\gamma$ in (1), for the exemplar optimal interior allocation ratios given in (2) – column 2, and (2a) – column 3. Note that in all cases the true share of input 1 is $\frac{1}{2}$ (from equation 3), but that the $\gamma$ parameter is substantially contingent upon the elasticity of substitution $\phi$ (compare rows), and upon the arbitrary measurement units of the inputs (compare columns).

[2] Our particular example models a goldsmith's production problem, with input 1 being gold and input 2 being one month's skilled labor.

The second column of Table 1 illustrates a yet more troubling result. Its values are determined by the same problem (1), now with the ratios given in (2a). However, a cursory inspection reveals that (2a) is precisely equivalent to (2) — only now the price and quantity of input 1 are measured in tonnes rather than in grams. It is reasonable to insist that $\gamma$ should be unaffected by our choice of measurement units for the inputs. However, by comparing the first and second columns of Table 1 we see that this is manifestly not the case. In fact, we prove in Appendix A that, outside of the Cobb-Douglas special case, $\gamma$ could be manipulated to any point in the open interval $(0, 1)$ by simply redefining the measurement units of one input.

$$\begin{array}{ccc} x_1 & : & x_2 \\ 10^{-4}\,\text{tonnes} & : & 1\,\text{unit} \end{array} \qquad\qquad \begin{array}{ccc} p_1 & : & p_2 \\ \$\,4\times10^{7}\,/\text{tonne} & : & \$4,000\,/\text{unit} \end{array} \tag{2a}$$

Normalization resolves these problems. If the CES production function in (1) were normalized by the appropriate input ratio, then $\gamma$ would accurately give the factor share of input 1 at that input ratio, for any choice of units and for any degree of complementarity $\phi$. In this example, Table 1 would be populated entirely by the value $\gamma = \frac{1}{2}$ until the final row. Empirically, it is not difficult to select an appropriate normalization vector: for example a policy-relevant input ratio could be given by an observed current or sample average input ratio as befits the situation. The estimated $\gamma$ could then be explicitly and correctly interpreted as a factor share at that input ratio.

The final row of Table 1 deserves particular comment. It shows that the implicitly normalized production function (1) is arbitrarily misspecified in the limit $\phi \to -\infty$. In this Leontief case, every undergraduate economics course teaches the importance of normalizing inputs (Appendix 2.2 shows that the required input ratio of (4) at $\phi \to -\infty$ is precisely that of its normalization point), however it is currently common practice to overlook normalization whenever the elasticity of technical substitution exceeds zero by any arbitrarily small amount.

## 3 Normalization in the CES Production Function

In this Section we will analyse the general (normalized) CES production function

$$f_{\text{N\_CES}} = A.\left(\sum_{i=1}^{n}\gamma_i\left[\frac{x_i}{\Theta_i}\right]^{\phi}\right)^{\frac{r}{\phi}}, \quad \text{where } \phi\in[-\infty,1]; \quad \sum_{i=1}^{n}\gamma_i=1; \quad \gamma_i\geq 0\;\forall\, i. \tag{4}$$

Here $r \geq 0$ parametrizes returns-to-scale, $A \geq 0$ parametrizes total-factor-productivity, and $\Theta_i$ can be consistently interpreted as the relative quantity requirement for $x_i$. In effect, $\Theta_i$ allows us to rescale $x_i$ from its pre-existing measurement units, which are defined in an entirely arbitrary manner, to a policy-relevant measurement unit. Our goal in this section is to establish whether an imposition of the implicit normalization $\mathbf{\Theta} = \mathbf{1}$ might be restrictive.

In considering (4) our analyses encompass a broad class of production functions. We observe that Linear, Cobb-Douglas, and Leontief production functions are strictly nested within (4), and that the translog function can be recovered as a second order approximation to (4) around its normalization point. The first three of these observations are standard results which we extend to arbitrarily nested and normalized CES functions in Appendix 2.2. The fourth result is new, and it represents a substantial improvement over the standard derivation of the translog function as a CES approximation in the neighbourhood of $\phi = 0$, because our result can be applied in the neighbourhood of any $\mathbf{\Theta} \in (\mathbb{R}^+)^n$.

We begin our analysis by noting that (4) is homogeneous of degree $r$ under any choice of normalization, implicit or otherwise. Thus the implicit normalization $\mathbf{\Theta} = \mathbf{1}$ restricts neither the interpretation nor the value of the returns-to-scale parameter $r$. Implicit normalization is similarly benign with regard to the complementarity parameter $\phi$, since under any choice of $\mathbf{\Theta}$ the elasticity of substitution between any two inputs remains

$$e_{i,j} = \frac{1}{1 - \phi} \tag{5}$$

We therefore turn our attention to the total-factor-productivity parameter $A$. It is immediate from (4) that $A$ can be interpreted as the quantity of output produced at the normalization point. Under implicit normalization, that normalization point is entirely arbitrary because it is dictated by the measurement units of each input. Thus, under implicit normalization, the relevance of $A$ will be lost whenever the point $\boldsymbol{x} = \mathbf{1}$ happens to lie outside of the consideration set of the producer. To see that $A$ is, in general, a complex function of $\{\boldsymbol{x}, \boldsymbol{\gamma},$ and $\phi\}$ we need only re-arrange (4).

In an empirical context, this problem of interpretability also impedes tractability, because there may be no relevant data with which to generate an initial value for $A$. By contrast, if $\mathbf{\Theta}$ were explicitly selected as, say, the average observed input level, then a sensible initial value for $A$ would be provided by the average observed output level. In an empirical context, it is not possible to simply substitute known values into the production function to infer an appropriate value for $A$ at $\boldsymbol{x} = \mathbf{1}$, because the parameters $r$ and $\phi$ will not be recoverable from the data until the production function itself has been estimated.

The imposition of implicit normalization poses similar challenges for the interpretation and estimation of the $\{\gamma_i\}$. These parameters are often of research interest, because they are frequently interpreted as the share of total resource that should optimally be spent on each input $x_i$. However it is not commonly acknowledged that the share interpretation of these parameters is typically only valid for the special Cobb-Douglas case where $\phi \to 0$ (as illustrated in Section 2). In general, an accurate interpretation of $\gamma_i$ is given by equation (6). This is derived in Appendix 2.1 by maximising equation (4) subject to the constraint $\sum_{i=1}^{n} p_i x_i \leq R$.

$$\gamma_i = \frac{p_i x_i \left(\frac{x_i}{\Theta_i}\right)^{-\phi}}{\sum_{j=1}^{n} p_i x_i \left(\frac{x_i}{\Theta_i}\right)^{-\phi}} \tag{6}$$

Equation (6) shows that the share of total resource that should optimally be spent on $x_i$ is generally a complex function that not only depends upon the parameter $\gamma_i$, but also upon a chosen point $\boldsymbol{x}$ in the production space, and the complementarity parameter $\phi$. From equation (6) we can see that $\gamma_i$ can only be interpreted as a universal share parameter in the Cobb-Douglas case, that is when $\phi$ happens to be precisely 0. However, in all cases $\gamma_i$ also represents the share of input $x_i$ along the specific ray in the production space where $\boldsymbol{x} = k.\boldsymbol{\Theta}$ for some scalar $k \in \mathbb{R}^+$, since here the ( ) terms in (6) 'cancel out'. Thus implicitly imposing that $\boldsymbol{\Theta} = \boldsymbol{1}$ will generally restrict the interpretability of the share parameters $\{\gamma_i\}$ to an arbitrary ray defined by the measurement units of the $x_i$.

Explicit normalization is therefore essential whenever the $\gamma_i$ are of research interest, because it allows the researcher to define $\boldsymbol{\Theta}$ at a policy-relevant point in the production space, and thereby to interpret the $\gamma_i$ as share parameters at that point. Moreover, explicit normalization affords the researcher with an opportunity to compare the share parameters experienced by specific subgroups of an empirical sample — for example between producers in developed cf. developing nations — by estimating the production function at alternative choices of $\boldsymbol{\Theta}$.[3]

Through its interpretation as a share parameter at the normalization point, $\gamma_i$ captures something of the intuitive 'value' of input $x_i$. We can make that intuition precise by calculating the output elasticity $\epsilon_i$ of (4) with respect to $x_i$. Equation (7) shows that normalization yields the intuitive interpretation of $\gamma_i$ as the relative output elasticity of $x_i$ at the normalization point, net of total-factor productivity $A$ (since at the normalisation point both bracketed terms reduce to unity). However, equation (7) also shows that the relationship between $\gamma_i$ and $\epsilon_i$ away from the normalization point is somewhat opaque, which further demonstrates that the implicit normalization $\boldsymbol{\Theta} = \boldsymbol{1}$ meaningfully restricts the interpretability of the model parameters.

$$\epsilon_i = \gamma_i.A.\left[\frac{x_i}{\Theta_i}\right]^{\phi}.\left(\sum_{i=1}^{n} \gamma_i \left[\frac{x_i}{\Theta_i}\right]^{\phi}\right)^{-1}; \qquad \epsilon_i\big|_{\boldsymbol{x}=k.\boldsymbol{\Theta}} = \gamma_i.A, \quad \text{for any } k \in \mathbb{R}^+. \tag{7}$$

[3]An analysis of this type is undertaken in Embrey (in prep.), where it provides a valuable comparison between the technologies of childhood skill formation experienced by privileged and under-privileged children respectively.

## 4 Discussion of alternative CES Production Forms

Since Dickinson (1954) first presented a production function that parametrized a constant elasticity of substitution, at least nine variants of CES production functions have been proposed in the literature. Table 2 lists each form, together the minimal parameter restrictions that it implicitly imposes upon the normalized form (3a). Whilst restrictions upon total-factor-productivity could be trivially resolved by the inclusion of a $\Lambda$ coefficient,[4] any relationships imposed between the normalization parameters $\boldsymbol{\Theta}$ and the share parameters $\boldsymbol{\gamma}$ are fundamentally restrictive. We have already seen that this is the case for the implicit normalization $\boldsymbol{\Theta} = \mathbf{1}$, and this section demonstrates that no restricted relationship between $\boldsymbol{\Theta}$ and $\boldsymbol{\gamma}$ can hold in general, because the properties of those parameters are orthogonal.

Table 2: *Alternative CES Production Forms*

| Origin | Form | Restrictions cf. (3a) |
|---|---|---|
| De La Grandville (1989) | $A\left(\gamma\left[\frac{x_1}{\Theta_1}\right]^\phi + (1-\gamma)\left[\frac{x_2}{\Theta_2}\right]^\phi\right)^{\frac{1}{\phi}}$ (3a) | — |
| Dickinson (1954) | $\left([\vartheta_1 x_1]^\phi + [\vartheta_2 x_2]^\phi + \Lambda^\phi\right)^{\frac{1}{\phi}}$ | $\gamma = \frac{1}{2}$; $A = 2^{\frac{1}{\phi}}$; additional $\Lambda$ |
| Pitchpord (1960) | $\left(\delta x_1^\phi + (1-\delta)x_2^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_1 = \Theta_2 = 1$; $A = 1$ |
| Arrow et al. (1961) | $\Lambda\left(\delta x_1^\phi + (1-\delta)x_2^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_1 = \Theta_2 = 1$ |
| David & van de Klundert (1965) | $\left([\vartheta_1 x_1]^\phi + [\vartheta_2 x_2]^\phi\right)^{\frac{1}{\phi}}$ | $\gamma = \frac{1}{2}$; $A = 2^{\frac{1}{\phi}}$ |
| Ventura (1997) | $\left(x_1^\phi + [\vartheta_2 x_2]^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_1 = 1$; $\gamma = \frac{1}{2}$; $A = 2^{\frac{1}{\phi}}$ |
| Senhadji (1997) | $\Lambda\left(\delta\left[\frac{x_1}{\delta}\right]^\phi + (1-\delta)\left[\frac{x_2}{1-\delta}\right]^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_1 = \gamma = 1 - \Theta_2$ |
| Barro & Sala-i Martin (2004) | $\Lambda\left(\delta[\vartheta_1 x_1]^\phi + (1-\delta)[(1-\vartheta_1)x_2]^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_2^{-1} = 1 - \Theta_1^{-1}$ |
| Thöni (2015) | $\delta^{\frac{\phi-2}{\phi}}\left(\delta^2\left[\frac{x_1}{\delta}\right]^\phi + (1-\delta)^2\left[\frac{x_2}{1-\delta}\right]^\phi\right)^{\frac{1}{\phi}}$ | $\Theta_1 = \gamma^{\frac{1}{2}}$; $\Theta_2 = (1-\gamma)^{\frac{1}{2}}$; $A = \delta^{\frac{\phi-2}{\phi}}$ |

Alternative CES production forms proposed in the literature. For uniformity these are presented with two inputs, with unit returns-to-scale, and with elasticity of substitution $1/(1-\phi)$. Other parameters are the 'share' of $x_1$: $\gamma, \delta \in [0,1]$, normalization parameters: $\Theta_i, \vartheta_i \in \mathbb{R}^+$, and total-factor-productivity: $A, \Lambda \in \mathbb{R}^+$.

Let us first observe that the concepts represented by $\boldsymbol{\gamma}$ and $\boldsymbol{\Theta}$ are distinct. In Section 2 we understood $\Theta_i$ as the relative resource requirement for input $x_i$. This is typically determined in advance: either theoretically, by an exogenous production process; or empirically, by observed resource utilization. In Section 3 we established that $\gamma_i$ represents the relative marginal 'value' of input $x_i$: equivalently thought of as the share of total

[4] Note that any restrictions upon $A$ would nevertheless distort the interpretability of any such coefficient.

resource expenditure optimally allocated to good $x_i$ along the normalization ray, or the relative output elasticity with respect to $x_i$ along that normalization ray. The clear distinction between these concepts suggests intuitively that neither can be subsumed into the other.

This conclusion is straightforward to demonstrate mathematically. For example, in order for $\gamma$ or $\Theta_1$ to be redundant we would require respectively:

$$\delta = \frac{\gamma}{\Theta_1^{\phi}}; \qquad \text{or} \qquad \vartheta_1 = \frac{\Theta_1}{\gamma^{1/\phi}}, \tag{8}$$

and neither of these can be consistently true across any range of elasticities $\phi$. Thus no functional form which imposes additional restrictions between $\boldsymbol{\Theta}$ and $\boldsymbol{\gamma}$ can admit the consistent parameter interpretations derived in sections 2 and 3 across any range of elasticities.

Each of the alternative forms listed in Table 2 therefore imposes meaningful restrictions upon the general, normalized form. The idiosyncratic limitations of many of these sets of restrictions have been discussed in detail by Klump & Preissler (2000) and by Klump, Mcadam & Willman (2012), however the latest contribution by Thöni (2015) deserves additional comment.

Thöni's CES variant has several desirable properties. Its parameter $\delta$ represents at once: (i) the relative resource requirement for $x_i$ across all $\phi$, (ii) the relative marginal product of $x_i$ in the linear case, and (iii) the share of $x_i$ in the Cobb-Douglas case. This is remarkable, because $\delta$ therefore arguably represents the most appropriate concept for the 'importance' of $x_i$ in all three special cases $\phi \to \{1; 0; -\infty\}$. Sadly, the form also has two overriding limitations. Firstly, $\delta$ represents neither the marginal product of $x_i$, nor the factor share of $x_i$, except in the aforementioned special cases (or in the highly improbable case that $x_i = \delta$). Secondly, in reality there is no reason that the resource requirements for the inputs $x_i$ should sum to unity. This would only be the case if the units of the $\{x_i\}$ were appropriately scaled in advance, which brings us back to the normalized functional form. The underlying problem with the Thöni variant is that it attempts to conflate two intuitively and mathematically distinct constructs into a single parameter.

This discussion has made plain the essential distinction between the share parameters $\{\gamma_i\}$ and the normalization parameters $\{\Theta_i\}$ within the general CES production function (4). We have further established that those parameters admit a consistent interpretation across the range of possible elasticities of substitution, but that this would be lost under the imposition of any additional restrictions between those parameters. The alternative functional forms proposed in Table 2 impose restrictions of this type. It is important to note that, despite restricting the freedom of the model, those additional assumptions yield no improvement in either empirical or theoretic tractability. In fact the opposite is demonstrably true in both cases (Klump & Saam 2008, Klump, Mcadam

& Willman 2012). This is because the parameters $\{\Theta_i\}$ are generally fixed a priori: in empirical applications, the normalization point is fixed prior to estimation by observed input and output levels, and in theoretical analysis the normalization point is analogously determined a priori by (algebraically) specified production and input levels.

## 5 Conclusion

Production function normalization has been accepted as an important tool in certain fields of macroeconomics, but it appears to have been overlooked by the existing microeconomic literature. Nevertheless we have shown that, in the absence of explicit normalization, key production function parameters are typically meaningless. But we have also seen that all production functions are normalized, if not explicitly then implicitly by the arbitrary measurement units of each input. The latter observation is important because it demonstrates that production function normalization imposes no additional assumptions — rather it makes an arbitrary implicit assumption both explicit and manipulable.

Explicit normalization is straightforward to implement. In some microeconomic applications, inputs are required in a ratio fixed by the production technology. In such situations it is natural to normalize production inputs in line with those relative requirements. Outside of such situations, estimated production parameters are typically of research interest in the neighbourhood of currently observed quantities of inputs and outputs; and so in practice those quantities can be used as normalization values.

The most pressing case for production function normalization arises when one considers the interpretation of share and total-factor-productivity parameters. We have seen that, outside of the Cobb-Douglas special case, those parameters admit their familiar interpretations only along the normalization ray through the production space. When normalization is implicit, that ray is arbitrary and may lie far from the producer's consideration set. In such situations, we have shown that implicit normalization can distort CES 'share' parameters to the extent that their oft-asserted interpretation is entirely lost.

Without normalization, estimated $\gamma$ can be made to vary across the open interval $(0, 1)$ simply by selecting alternative measurement units for the inputs. Normalization solves this problem, which leads us to conclude that it could be an important technique within any production context. One specialized example might be an international comparison of production with capital input, since in this situation estimated parameters will be affected by the choice of currency unit. However, nothing in our analyses is specific to that application or even to the domain of economic production. Normalization could therefore be equally important across the much broader domain of constrained utility maximization.

It is rare for existing microeconomic research to acknowledge the limitations posed by implicit normalization (one notable exception being Arrow et al. 1961). It is therefore possible that even well-established results could be misleading. For example, Embrey (in prep.) shows that implicit normalization meaningfully distorts our present understanding of the technology of childhood skill formation. In contrast, this paper has shown that explicitly normalized utility functions are at once: less restrictive, more tractable, and more interpretable than the alternative formulations proposed in the literature. We have also seen that normalization is straightforward to implement in practice. It therefore seems appropriate to suggest that normalized forms should be used as a matter of course. Certainly, where an implicit normalization is imposed, the validity of that assumption should be explicitly considered.

## References

**Arrow, K. J., H. B. Chenery, B. S. Minhas, and R. M. Solow.** 1961. "Capital-Labor Substitution and Economic Efficiency." *The Review of Economics and Statistics*, 43(3): 225–250.

**Barro, Robert J., and Xavier Sala-i Martin.** 2004. *Economic Growth.* 2nd ed. Cambridge: The MIT Press.

**Cantore, Cristiano, Miguel León-Ledesma, Peter Mcadam, and Alpo Willman.** 2014. "Shocking stuff: Technology, hours, and factor substitution." *Journal of the European Economic Association*, 12(1): 108–128.

**David, Paul A ., and Th. van de Klundert.** 1965. "Biased Efficiency Growth and Capital-Labor Substitution in the U . S ., 1899-1960." *The American Economic Review*, 55(3): 357–394.

**De La Grandville, Olivier.** 1989. "In Quest of the Slutsky Diamond." *The American Economic Review*, 79(3): 468–481.

**Dickinson, H. D.** 1954. "A Note on Dynamic Economics." *The Review of Economic Studies*, 22(3): 169–179.

**Embrey, Iain Peter.** In Prep. "Re-Estimating the Technology of Cognitive and Noncognitive Skill Formation."

**Irmen, Andreas, and Rainer Klump.** 2009. "Factor substitution, income distribution and growth in a generalized neoclassical model." *German Economic Review*, 10(4): 464–479.

**Klump, Rainer, and Harald Preissler.** 2000. "CES Production Functions and Economic Growth." *The Scandinavian Journal of Economics*, 102(1): 41–56.

**Klump, Rainer, and Marianne Saam.** 2008. "Calibration of normalised CES production functions in dynamic models." *Economics Letters*, 99(2): 256–259.

**Klump, Rainer, and Olivier De La Grandville.** 2000. "Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions." *The American Economic Review*, 90(1): 282–291.

**Klump, Rainer, Peter McAdam, and Alpo Willman.** 2007. "Factor substitution and factor-augmenting technical progress in the United States: A normalized supply-side system approach." *Review of Economics and Statistics*, 89(1): 183–192.

**Klump, Rainer, Peter Mcadam, and Alpo Willman.** 2012. "The normalized ces production function: Theory and empirics." *Journal of Economic Surveys*, 26(5): 769–799.

**León-ledesma, Author Miguel A, Peter Mcadam, Alpo Willman, Source The, American Economic, No September, Peter Mcadam, and By Miguel.** 2010. "Identifying the Elasticity of Substitution with Biased Technical Change." *The American Economic Review*, 100(4): 1330–1357.

**León-Ledesma, Miguel A., Peter Mcadam, and Alpo Willman.** 2015. "Production technology estimates and balanced growth." *Oxford Bulletin of Economics and Statistics*, 77(1): 40–65.

**Papageorgiou, Chris, and Marianne Saam.** 2008. "Two-level CES production technology in the solow and diamond growth models." *Scandinavian Journal of Economics*, 110(1): 119–143.

**Pitchpord, J.D.** 1960. "Growth and the Elasticity of Factor Substitution." *Economic Record*, 36(76): 491–504.

**Senhadji, Abdelhak S.** 1997. "Two common problems related to the use of the Armington aggregator in computable general equilibrium models." *Applied Economics Letters*, 4(1): 23–25.

**Thöni, Christian.** 2015. "A note on CES functions." *Journal of Behavioral and Experimental Economics*, 59: 85–87.

**Ventura, Jaume.** 1997. "Growth and Interdependence." *The Quarterly Journal of Economics*, 112(1): 57–84.

**Willman, Alpo, and Peter McAdam.** 2013. "Technology , Utilization , and Inflation : What Drives the New Keynesian Phillips Curve ?" *Journal of Money Credit and Banking*, 45(8).

**Xue, Jianpo, and Chong K. Yip.** 2013. *Aggregate elasticity of substitution and economic growth: A synthesis.* Vol. 38.

# A Proof that the value of CES share parameters without normalization can be arbitrarily determined by a choice of measurement units

From equation (6), which is derived in Appendix 2.1, we have that, for an interior optimal allocation under monotonic preferences and with positive prices:

$$\gamma_i = \frac{p_i x_i \left(\frac{x_i}{\Theta_i}\right)^{-\phi}}{\sum_{j=1}^{n} p_i x_i \left(\frac{x_i}{\Theta_i}\right)^{-\phi}} \tag{6}$$

We can specialize this result to the implicit normalization $\mathbf{\Theta} = \mathbf{1}$ and the two-input case of Section 2 to give:

$$\gamma = \frac{p_1 x_1 x_1^{-\phi}}{p_1 x_1 x_1^{-\phi} + p_2 x_2 x_2^{-\phi}} \tag{9}$$

Now any change in the measurement units for input $i$ can be written as $x_i \mapsto \frac{x_i}{k}$, for some $k \in \mathbb{R}^+$. Note that any such change leaves the term $p_i x_i$ unaffected, as it will also have the countervailing effect of mapping the unit prices $p_i \mapsto k.p_i$.

We proceed by fixing any production complementarity parameter $\phi$, and any current input and price ratios $x_2/x_1 =: r$, and $p_2/p_1 =: p$. Finally, fix any desired $\tilde{\gamma} \in (0, 1)$. Then the share parameter would be equal to $\tilde{\gamma}$ under the change in measurement units given by $x_1 \mapsto \frac{x_1}{k}$, where

$$k = \left( \frac{\tilde{\gamma}}{(1 - \tilde{\gamma}) p r^{1-\phi}} \right)^{\frac{1}{\phi}}. \tag{10}$$

(10) is a well-defined function for any $\phi \in (-\infty, 1] \setminus 0;\ r, p \in \mathbb{R}^+$. Thus, outside of the Cobb-Douglas case, the share parameter of any two-input CES function can be made to vary arbitrarily across (0,1) merely by redefining the measurement units of input 1. $\square$

This result generalizes to $n$ inputs, although the closed-form solution for $k$ would require solving $n$ nonlinear equations in $n-1$ unknowns, so its calculation could quickly become computationally challenging.

# B Standard Results for the CES Production Function Extended to its Nested and Normalized Form

## 2.1 Derivation of the economic interpretation of the $\gamma_i$ in a nested and normalized CES function

Consider again the normalized $n$-factor CES production function

$$f_{\text{N\_CES}} = A. \left( \sum_{i=1}^{n} \gamma_i \left[ \frac{x_i}{\Theta_i} \right]^{\phi} \right)^{\frac{r}{\phi}}, \quad \text{where } \phi \in [-\infty, 1]; \quad \sum_{i=1}^{n} \gamma_i = 1; \quad \gamma_i \geq 0 \; \forall \, i. \tag{4}$$

We first prove Lemma 1:

**Lemma 1** *Under monotonic preferences and with positive prices, the parameters $\gamma_i$ in the normalized $n$-factor CES production function (4) represent the share of resources optimally dedicated to input $x_i$ at any point along the ray in the production space that contains the normalization point.*

*Proof:*

We maximize (4) subject to the resource constraint

$$\sum_{i=1}^{n} p_i x_i \leq R. \tag{11}$$

This holds with equality when preferences are monotonic and prices $p_i$ are positive. The first order condition of this system under Lagrangian maximization is therefore given by the set of $n$ equations:

$$\frac{r\,\gamma_i}{\phi\,\Theta_i}.\phi\left[\frac{x_i}{\Theta_i}\right]^{\phi-1}.f(\cdot)^{\frac{r-\phi}{r}} = \lambda p_i,$$

where $\lambda$ is the Lagrangian multiplier. By forming the ratio of any two of these equations we obtain the well-known relationship between the optimal marginal rate of substitution between two goods and their price ratio, specifically:

$$\frac{\gamma_i\Theta_j.\,[x_i/\Theta_i]^{\phi-1}}{\gamma_j\Theta_i.\,[x_j/\Theta_j]^{\phi-1}} = \frac{p_i}{p_j},$$

which implies that

$$\gamma_i.\,\frac{[x_i/\Theta_i]^{\phi}}{x_i p_i}.\frac{x_j p_j}{[x_j/\Theta_j]^{\phi}} = \gamma_j \;\forall\, j \neq i.$$

We now sum these equations across all $j \neq i$ to obtain

$$\gamma_i.\,\frac{[x_i/\Theta_i]^{\phi}}{x_i p_i}.\sum_{j\neq i}\frac{x_j p_j}{[x_j/\Theta_j]^{\phi}} = \sum_{j\neq i}\gamma_j = 1-\gamma_i$$

rearranging gives

$$\gamma_i.\,\frac{[x_i/\Theta_i]^{\phi}}{x_i p_i}.\sum_{j=1}^{n}\frac{x_j p_j}{[x_j/\Theta_j]^{\phi}} = 1$$

$$\gamma_i = \frac{p_i x_i\left(\frac{x_i}{\Theta_i}\right)^{-\phi}}{\sum_{j=1}^{n} p_i x_i\left(\frac{x_i}{\Theta_i}\right)^{-\phi}} \tag{6}$$

Equation (6) was discussed in Section 3, where we noted that, at any point $\boldsymbol{x}$ along the ray containing $\boldsymbol{\Theta}$ we will have that $\boldsymbol{x} = k.\boldsymbol{\Theta}$ for some scalar $k \in \mathbb{R}$. Hence, along that ray, equation (6) reduces to give

$$\gamma_i = \frac{x_i p_i}{\sum x_j p_j} = \frac{x_i p_i}{R}, \tag{12}$$

which specializes to (3) in the two-input case. Thus we have demonstrated that $\gamma_i$ can be interpreted along the ray containing $\Theta$ as the share of total resources which should optimally be allocated to $x_i$. $\square$

For our main theorem we show that Lemma 1 generalizes to an arbitrary nested CES production function.

**Theorem 1** *Under monotonic preferences, with positive prices, and at any point along the ray in the production space that contains the normalization point: the share of total resources optimally allocated to input $x_i$ within an arbitrary nested and normalized CES* *production function is obtained by multiplying its coefficient $\gamma_i$ with the $\gamma$ coefficients on all nested levels that contain $x_i$.*

*Proof:*

We proceed by induction over the number of nesting levels in an arbitrary nested and normalised CES production function $g(\boldsymbol{y})$. The induction step in the construction of $g$ is to substitute the term

$$f(\boldsymbol{x}) = \gamma_i^g \left( \gamma_1^f \left[\frac{x_1}{\Theta_1^f}\right]^\phi + \gamma_2^f \left[\frac{x_2}{\Theta_2^f}\right]^\phi + ...\gamma_n^f \left[\frac{x_n}{\Theta_n^f}\right]^\phi \right)^{\frac{\varphi}{\phi}} \quad \text{for some term } \gamma_i^g \left(\frac{y_i}{\Theta_i^g}\right)^\varphi$$

within the existing $g(\boldsymbol{y})$. Let us first fix some notation for this procedure. Denote the composite, post-substitution, production function as $G := g \circ f$. Denote by $m$ the number of factors in the initial production function $g$, and let $\boldsymbol{\Theta^g} \in \mathbb{R}^m$ be its normalization point. Fix any point $\boldsymbol{y} = k.\boldsymbol{\Theta^g}$ along the ray in the production space containing $\boldsymbol{\Theta^g}$. Denote by $\boldsymbol{\Theta^f}$ the normalization point of the CES function $f$ that is to be substituted for $y_i$, and denote by denote by $\boldsymbol{x}$ the point $k.\boldsymbol{\Theta^f}$. Then denote by $\boldsymbol{\Theta^G}$ the concatenation of the first $i-1$ elements of $\boldsymbol{\Theta^g}$, followed by the $n$ elements of $\boldsymbol{\Theta^f}$, followed by the final $m-i$ elements of $\boldsymbol{\Theta^g}$, and similarly denote by $\boldsymbol{Y}$ the concatenation of the first $i-1$ elements of $\boldsymbol{y}$, followed by the $n$ elements of $\boldsymbol{x}$, followed by the final $m-i$ elements of $\boldsymbol{y}$.

Now, Lemma 1 provides the desired result for nesting level 0, and so to prove Theorem 1 it suffices to show the induction step.

By the induction assumption, the share of total resource optimally dedicated to each input $y_j$ within $g$ is given by the product of its coefficient $\gamma_j^g$ with the outer $\gamma^g$ coefficients for all nesting levels that contain $y_j$. Denote that product as $\prod_j$ for each input $y_j$. We must show that, along the ray from the origin containing the normalization point, the share of resources optimally allocated to each input $y_j$, $j \neq i$ is unchanged by the substitution $y_i/\Theta_i \mapsto f(\boldsymbol{x})$, and that the share of resources optimally allocated to each additional input $x_k$ is given by $\gamma_k^f . \prod_i$.

For the former, consider the term that has been substituted out. This had the form $\gamma_i \left[y_i/\Theta_i\right]^\varphi$, which evaluates to $\gamma_i^g \left[k\right]^\varphi$ at $\boldsymbol{y} = k.\boldsymbol{\Theta^g}$. But the substituted term

$$\gamma_i^g \left( \gamma_1^f \left[\frac{x_1}{\Theta_1}\right]^\phi + \gamma_2^f \left[\frac{x_2}{\Theta_2}\right]^\phi + ...\gamma_n^f \left[\frac{x_n}{\Theta_n}\right]^\phi \right)^{\frac{\varphi}{\phi}} = \gamma_i^g \left( \gamma_1^f \left[k\right]^\phi + \gamma_2^f \left[k\right]^\phi + ...\gamma_n^f \left[k\right]^\phi \right)^{\frac{\varphi}{\phi}} = \gamma_i^g \left[k\right]^\varphi$$

is therefore equivalent because $\sum_{k=1}^n \gamma_k^f = 1$.

Given that the share of resources optimally allocated to each input $y_j$, $j \neq i$ is unchanged by the substitution of $f$ for $y_i/\Theta_i$, and since the $\gamma_i^f$ sum to unity, we have that the share of the total resources $R^G$ available for the composite function that is optimally allocated to the newly nested function $f$ remains $\prod_i .R^G$. We may therefore apply Lemma

1 to $f$ with the total resources $R := R^f = \prod_i .R^G$ to see immediately that

$$\gamma_i^f.\prod_i = \tfrac{x_i p_i}{R^G},$$

namely that the product of the coefficient on each newly nested input $x_i$ with all coefficients on its nesting levels $\prod_i$ gives the share of total resources that should optimally be allocated to that input at any point along the ray from the origin that contains $\Theta^G$. □

Note that Theorem 1 could equivalently be rewritten as Corollary 1:

**Corollary 1** *If the normalized CES function $f$ represents the production of an intermediate good $y_i/\Theta_i$ which is an input to normalized CES production function $g$, and each function is normalized so that its share parameters are interpretable along a ray-of-interest within its respective production space, then the compound production function $G := g \circ f$, is normalized so that its share parameters are interpretable along the corresponding ray within its production space.*

## 2.2 Derivation of limiting production functions for the general normalized and nested CES form as $\{\phi\} \to \{1, 0, -\infty\}$

One should first note that, in any limit where all nested complementarity parameters converge, all nested levels are raised to the power 1, and so any nested form becomes equivalent to some non-nested form (4). It therefore suffices to consider the above limits as applied to (4).

It is trivial to see that in the limit $\phi \to 1$, (4) becomes a linear, perfect-substitutes, production function, with total-factor-productivity $A$ and returns-to-scale $r$:

$$\lim_{\phi \to 1} f(\boldsymbol{x}) = A.\left(\sum_{i=1}^{n} \gamma_i \left[\frac{x_i}{\Theta_i}\right]\right)^r. \tag{13}$$

To find the limit of $f$ as $\phi \to -\infty$, first observe that the set of values

$$\left\{\left[\frac{x_i}{\Theta_i}\right]\right\}_i$$

is finite, and so it obtains its infimum: let us denote this by $\left[\frac{x_j}{\Theta_j}\right]$. Since we are considering the limit as $\phi \to -\infty$ we can safely restrict our attention to the domain where $\phi < 0$. On that domain, $\left[\frac{x_j}{\Theta_j}\right]^\phi$ is the largest element of $\left\{\left[\frac{x_i}{\Theta_i}\right]^\phi\right\}_i$, and so, since the $\gamma_i$ are positive and sum to unity, we have that

$$\gamma_j^{\frac{1}{\phi}}\left[\frac{x_j}{\Theta_j}\right] = \left(\gamma_j\left[\frac{x_j}{\Theta_j}\right]^\phi\right)^{\frac{1}{\phi}} \leq \left(\frac{f(\boldsymbol{x})}{A}\right)^{\frac{1}{r}} = \left(\sum_{i=1}^{n}\gamma_i\left[\frac{x_i}{\Theta_i}\right]^\phi\right)^{\frac{1}{\phi}} \leq \left(\left[\frac{x_j}{\Theta_j}\right]^\phi\right)^{\frac{1}{\phi}} = \left[\frac{x_j}{\Theta_j}\right].$$

It is therefore clear that $\left(\frac{f(\boldsymbol{x})}{A}\right)^{\frac{1}{r}}$ is sandwiched between

$$\lim_{\phi\to-\infty} \gamma_j^{\frac{1}{\phi}} \left[\frac{x_j}{\Theta_j}\right] \text{ and } \left[\frac{x_j}{\Theta_j}\right],$$

and so we have that

$$\lim_{\phi\to-\infty} f(\boldsymbol{x}) = A.\left(\min_i \left\{\left[\frac{x_i}{\Theta_i}\right]\right\}\right)^r. \tag{14}$$

which is the Leontief, perfect-complements, production function, with total-factor-productivity $A$ and returns-to-scale $r$. □

To find the limit of $f$ as $\phi \to 0$, we begin by define $h(\phi)$ as the inner sum within (4)

$$h(\phi) := \sum_{i=1}^{n} \gamma_i \left[\frac{x_i}{\Theta_i}\right]^\phi. \tag{15}$$

We then apply the Taylor expansion (17) to $h(\phi)$ around the point $\phi$ to obtain

$$h(\phi) = 1 + \phi \sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + \phi^2 \sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + O(\phi^3),$$

which we can substitute back into (4) to obtain

$$f(\boldsymbol{x}) = \left(1 + \phi \sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + \phi^2 \sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + O(\phi^3)\right)^{\frac{1}{\phi}}.$$

By comparing this expression with the definition of the exponential function

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{m\to 0}(1+mx)^{\frac{1}{m}}$$

we see that

$$f(\boldsymbol{x}) = A.\left[\exp\left(\sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + \phi \sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + O(\phi^2)\right)\right]^r,$$

which simplifies as $\phi \to 0$ to (the normalization of) the familiar Cobb-Douglas form, with total-factor-productivity $A$ and returns-to-scale $r$:

$$f(\boldsymbol{x}) = \left[\exp\left(\sum_{i=1}^{n} \gamma_i \ln\left[\frac{x_i}{\Theta_i}\right]\right)\right]^r = \prod_{i=1}^{n}\left[\frac{x_i}{\Theta_i}\right]^{r.\gamma_i}. \tag{16}$$

□

### 2.3 Derivation of the translog production function as a 2nd order Taylor expansion of the CES form around its normalization point

Consider again the normalized $n$-factor CES production function given in (4). We first prove Lemma 2:

**Lemma 2** *The second order Taylor expansion of the normalized $n$-factor CES production function (4) around its normalization point has a saturated Translog form.*

*Proof:*

The (multinomial) Taylor series representation of a function $f(\boldsymbol{x})$ around the point $\boldsymbol{\Theta}$ is given by

$$f(\boldsymbol{x}) = f(\boldsymbol{\Theta}) + \sum_{i=1}^{n} f_{x_i}(\boldsymbol{\Theta}).\Delta x_i + \frac{1}{2}\sum_{i=1}^{n} f_{x_i x_i}(\boldsymbol{\Theta}).\Delta {x_i}^2 + \sum_{i=1}^{n}\sum_{j=i+1}^{n} f_{x_i x_j}(\boldsymbol{\Theta}).\Delta x_i \Delta x_j + O(\Delta {x_i}^3), \tag{17}$$

where $\boldsymbol{\Delta x} := \boldsymbol{x} - \boldsymbol{\Theta}$.

We begin by taking logs both sides of equation (4), to obtain

$$\ln f(\boldsymbol{x}) = \ln A + \frac{r}{\phi}\ln\left(\gamma_1\left[\frac{x_1}{\Theta_1}\right]^{\phi} + \gamma_2\left[\frac{x_2}{\Theta_2}\right]^{\phi} + ...\gamma_n\left[\frac{x_n}{\Theta_n}\right]^{\phi}\right) = \ln A + \frac{r}{\phi}\ln\sum_{i=1}^{n}\gamma_i e^{\phi\ln\left[\frac{x_i}{\Theta_i}\right]}$$

We then consider $\ln f$ as a function of $(\ln\left[\frac{\boldsymbol{x}}{\boldsymbol{\Theta}}\right])$, and expand around the point $\ln\left[\frac{\boldsymbol{x}}{\boldsymbol{\Theta}}\right] = \boldsymbol{0}$ which is equivalent to the point $\boldsymbol{x} = \boldsymbol{\Theta}$. The necessary terms for that Taylor expansion are:

$$\begin{aligned}
\ln f(\boldsymbol{0}) &= \ln A \\
\frac{\partial \ln f}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]}(\boldsymbol{0}) &= r.\gamma_i \\
\frac{\partial^2 \ln f}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]^2}(\boldsymbol{0}) &= r.\phi\gamma_i(1-\gamma_i) \\
\frac{\partial^2 \ln f}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]\partial \ln\left[\frac{x_j}{\Theta_j}\right]}(\boldsymbol{0}) &= -r.\phi\gamma_i\gamma_j.
\end{aligned} \tag{18}$$

Substituting these into the 2nd order Taylor series (17) yields:

$$\ln f(\boldsymbol{x}) \approx \ln A + \sum_{i=1}^{n} r.\gamma_i \ln\left[\frac{x_i}{\Theta_i}\right] + \frac{\phi}{2}\sum_{i=1}^{n} r.\gamma_i(1-\gamma_i)\left(\ln\left[\frac{x_i}{\Theta_i}\right]\right)^2 - \phi\sum_{i=1}^{n}\sum_{j=i+1}^{n} r.\gamma_i\gamma_j.\ln\left[\frac{x_i}{\Theta_i}\right].\ln\left[\frac{x_j}{\Theta_j}\right] \tag{19}$$

which has the saturated (normalized) Translog form. □

We now show that Lemma 2 generalizes to an arbitrary nested CES production function.

**Theorem 2** *The second order Taylor expansion of an arbitrary nested and normalized CES production function around its normalization point has a saturated Translog form.*

*Proof:*

As per the proof of Theorem 1, we proceed by induction over the number of nesting levels in an arbitrary nested and normalised log CES production function $\ln g(\boldsymbol{y})$. Any Such a function with nesting level $k+1$ is constructed by the substitution of a log-CES form (20) for some term $\varphi \ln \left[\frac{y_k}{\Theta_k}\right]$ within the appropriate function $\ln g(\boldsymbol{y})$ of nesting level $k$.

$$\ln f(\boldsymbol{x}) = \frac{\varphi}{\phi} \ln \left( \gamma_1^f \left[\frac{x_1}{\Theta_1}\right]^\phi + \gamma_2^f \left[\frac{x_2}{\Theta_2}\right]^\phi + ...\gamma_n^f \left[\frac{x_n}{\Theta_n}\right]^\phi \right) = \frac{\varphi}{\phi} \ln \sum_{i=1}^{n} \gamma_i^f e^{\phi \ln\left[\frac{x_i}{\Theta_i}\right]} \tag{20}$$

We denote the resultant compound production function as $G := g \circ f$, and use again the other notation defined in Appendix 2.1.

Since Lemma 2 provides the result for nesting level 0, it suffices to show that result survives the induction step described above.

Note first that each of the terms given in (18) is unaffected by the substitution whenever $i, j \neq k$, since the differential of (20) with respect to any $\ln \left[\frac{y_i}{\Theta_i}\right]$ is zero, and since (20) evaluated at $\ln \left[\frac{\boldsymbol{x}}{\boldsymbol{\Theta}}\right] = \mathbf{0}$ is also zero. By the induction assumption it is therefore possible to write these as:

$$\begin{aligned}
\ln G(\mathbf{0}) &= \ln A \\
\frac{\partial \ln G}{\partial \ln \left[\frac{y_i}{\Theta_i}\right]}(\mathbf{0}) &= a(i) \\
\frac{\partial^2 \ln G}{\partial \ln \left[\frac{y_i}{\Theta_i}\right]^2}(\mathbf{0}) &= b(i) \\
\frac{\partial^2 \ln G}{\partial \ln \left[\frac{y_i}{\Theta_i}\right] \partial \ln \left[\frac{y_j}{\Theta_j}\right]}(\mathbf{0}) &= c(i,j), \\
\text{for } i, j &\neq k,
\end{aligned} \tag{21}$$

where $a, b$, and $c$ are functions of the parameters of $g$. The additional terms relevant to the 2nd order Taylor expansion of the compound function $G$ are:

$$
\begin{aligned}
\frac{\partial \ln G}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]}(\mathbf{0}) &= \gamma_i^f.a(k)\\
\frac{\partial^2 \ln G}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]^2}(\mathbf{0}) &= \left(\gamma_i^f\right)^2.b(k)+\phi\gamma_i^f\left(1-\gamma_i^f\right).a(k)\\
\frac{\partial^2 \ln G}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]\partial \ln\left[\frac{y_j}{\Theta_j}\right]}(\mathbf{0}) &= \gamma_i^f.c(k,j),\\
\frac{\partial^2 \ln G}{\partial \ln\left[\frac{x_i}{\Theta_i}\right]\partial \ln\left[\frac{x_j}{\Theta_j}\right]}(\mathbf{0}) &= \gamma_i^f\gamma_j^f.\big(b(k)-\phi.a(k)\big),\\
\text{for } j \neq k, &
\end{aligned} \tag{22}
$$

Substituting (21) and (22) into the Taylor series (17) we obtain:

$$
\ln f(\boldsymbol{x}) \approx \ln A + \sum_{i=1}^{n+m} \tilde{a}(i,k)\ln\left[\frac{y_i}{\Theta_i}\right] + \frac{1}{2}\sum_{i=1}^{n+m}\tilde{b}(i,k)\left(\ln\left[\frac{y_i}{\Theta_i}\right]\right)^2 + \sum_{i=1}^{n+m}\sum_{j=i+1}^{n+m}\tilde{c}(i,j,k).\ln\left[\frac{y_i}{\Theta_i}\right].\ln\left[\frac{y_j}{\Theta_j}\right], \tag{23}
$$

where:

$$
\tilde{a}(i,k) := \begin{cases} a(i), & i<k \text{ or } i>k+n \\ \gamma_i^f.a(k), & k\le i\le k+n \end{cases}
$$

$$
\tilde{b}(i,k) := \begin{cases} b(i), & i<k \text{ or } i>k+n \\ \left(\gamma_i^f\right)^2.b(k)+\phi\gamma_i^f\left(1-\gamma_i^f\right).a(k), & k\le i\le k+n \end{cases}
$$

$$
\tilde{c}(i,j,k) := \begin{cases} c(i,j), & i,j<k \text{ or } >k+n \\ \gamma_i^f.c(k,j), & k\le i\le k+n,\ \text{and}\big\{j<k \text{ or } >k+n\big\} \\ \gamma_i^f\gamma_j^f.\big(b(k)-\phi.a(k)\big), & k\le i,j\le k+n \end{cases}
$$

which is again of the saturated Translog form. □