

Knoblach, Michael; Stöckl, Fabian

Working Paper

What determines the elasticity of substitution between capital and labor? A literature review

DIW Discussion Papers, No. 1784

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Knoblach, Michael; Stöckl, Fabian (2019) : What determines the elasticity of substitution between capital and labor? A literature review, DIW Discussion Papers, No. 1784, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/191911>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

1784

Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2019

What Determines the Elasticity of Substitution between Capital and Labor? A Literature Review

Michael Knoblach and Fabian Stöckl

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2019

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

What determines the elasticity of substitution between capital and labor? A literature review*

Michael Knoblach[†]

TU Dresden

Fabian Stöckl[‡]

DIW Berlin and TU Berlin

January 14, 2019

Abstract

This paper provides the first comprehensive review of the empirical and theoretical literature on the determinants of the elasticity of substitution between capital and labor. Our focus is on the two-input constant elasticity of substitution (CES) production function. By example of the U.S., we highlight the distinctive heterogeneity in empirical estimates of σ at both the aggregate and industrial level and discuss potential methodological explanations for this variation. The main part of this survey then focuses on the determinants of σ . We first review several approaches to the microfoundation of production functions, especially the CES production function. Second, we outline the construction of an aggregate elasticity of substitution (AES) in a multi-sectoral framework and investigate its dependence on underlying sectoral elasticities. Third, we discuss the influence of the institutional framework on the determination of σ . The concluding section of this review identifies a number of potential empirical and theoretical avenues for future research. Overall, we demonstrate that the effective elasticity of substitution (EES), which is typically estimated in empirical studies, is generally not an immutable deep parameter but depends on a multitude of technological, non-technological and institutional determinants.

JEL classification: D24; E23; O14; O40.

Keywords: Elasticity of substitution; aggregate elasticity; capital; labor; economic growth; microfoundation; Cobb-Douglas and CES production function.

*We thank Claudia Kemfert, Alexander Kemnitz, Patrick Zwerschke, and the participants of the Brown Bag seminar at the TU Dresden for valuable comments and intensive discussions that improved the paper significantly.

[†]Corresponding author. Technische Universität Dresden, Faculty of Business and Economics, Dresden, Germany. E-mail: michael.knoblach@tu-dresden.de

[‡]German Institute for Economic Research (DIW Berlin) and Technische Universität Berlin, Berlin, Germany. E-mail: FStoeckl@diw.de

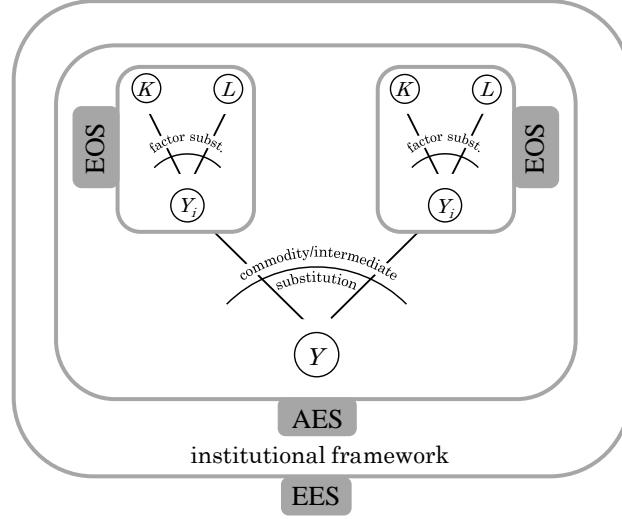
1 Introduction

The principle of substitution is a core element of neoclassical theory, especially production theory. As an implicit or explicit parameter of the production function, the elasticity of substitution represents the envelope of all technically feasible combinations of input factors that produce a certain amount of output. Intuitively, the elasticity of substitution can be regarded “as a measure of the efficiency of the productive system” to transform inputs into output (de La Grandville, 1989, p. 479). That is, the more interchangeable two inputs in production are, i.e., the higher the elasticity of substitution is, the better an economy can transform an increase in the relative abundance of an input factor into further output.

In macroeconomics, the elasticity of factor substitution between capital and labor, σ , has been shown to be of critical importance for a broad range of topics. The degree of substitutability influences, for instance, the response of business investments to variation in the interest rate (Chirinko, 2002), the returns of productive factors in an open-economy context (Jones and Ruffin, 2008), the relation between technology shocks and hours worked (Cantore et al., 2014, 2017), and the degree of sectoral transformation (Alvarez-Cuadrado et al., 2017). Additionally, Piketty’s (2014) recent explanation of the observed decline in the share of labor in total income over the past 20 years crucially depends on the assumption that σ exceeds unity. The elasticity of substitution is even more important in the theory of economic growth. As shown in the growth models of Solow (1956) and Pitchford (1960), for a sufficiently high value of σ , necessarily greater than unity, perpetual growth is possible, even in the absence of technological progress. Further aspects of economic growth, such as the level of per capita income (Klump and de La Grandville, 2000), the speed of convergence to the steady state (Turnovsky, 2002, 2008, Klump and Saam, 2008), the sensitivity of cross-country income differences (Caselli, 2005), and the direction of technological change (Acemoglu, 2003), are also strongly related to the magnitude of σ .

Although the contributions mentioned above highlight the fundamental importance of the elasticity of substitution between capital and labor for a broad range of theoretical and empirical research, much less is known about what underlies and determines σ itself. Moreover, no agreement exists in the literature regarding the precise value of σ for different countries and specific sectors and industries in either the short-run or over very long periods of time (Chirinko, 2008, McAdam, 2016). Moreover, there is no consensus on the economic framework the elasticity considered in (Jones, 1965, Miyagiwa and Papageorgiou, 2007). In addition, little knowledge exists of the technological and non-technological reasons that could lead to a change in the elasticity of substitution over time (Klump and Preissler, 2000). This gap of research appears to be puzzling, as some initial conjectures about potential determinants of σ can already be found in Hicks (1932, 1936, 1963). Hicks also suggested that, in contrast to a one-sector economy, in a multi-sectoral framework, the aggregate degree of substitutability between capital and labor is determined not only by the properties of the production technology itself but also by changes in consumption. According to Hicks, the aggregate elasticity of substitution is determined by

Figure 1: Different concepts of factor substitution



The aggregate elasticity of substitution (AES) of an economy comprises both factor substitution (EOS) at the industrial or sectoral level and commodity or intermediate substitution at the aggregate level. On both levels, substitution is potentially influenced by the institutional framework. Combined, these concepts establish the effective elasticity of substitution (EES), which is typically estimated in empirical studies.

i) intrasectoral substitution of known methods of production, ii) technological innovations that augment this set of methods, and iii) intersectoral substitution due to commodity substitution. Moreover, at any level of aggregation, factor substitution is potentially influenced by the institutional framework.

On the basis of Hicks' conjectures, the purpose of this paper is to summarize and critically review the existing literature on the determinants of the elasticity of substitution between capital and labor. We attempt to answer the question of whether σ should be regarded as an exogenously given, immutable deep parameter or as an endogenous variable (Palivos, 2008, p. 688). As illustrated by figure 1, we distinguish three concepts of factor substitution. The first concept treats the elasticity of substitution (EOS) as a purely technological parameter as prevalent on the one-sector firm level. The second concept is based on the idea of an economy-wide aggregate elasticity of substitution (AES) that also accounts for several non-technological determinants, e.g., consumption preferences. Finally, the third concept takes into account the role the institutional framework plays in factor substitution. This final concept describes the effective elasticity of substitution (EES), which is typically estimated in empirical studies. The EES includes technological substitution and (non-technological) commodity/intermediate substitution, which are both influenced by the institutional framework.

The rest of this paper is structured as follows. The investigation of the role of σ in macroeconomics has focused on its use in the two-input CES production function. Thus, section 2 provides a short summary of not only the history and properties of the elasticity of substitution but also the CES production function. By example of

the U.S., in section 3, we highlight the large heterogeneity in empirical estimates of the elasticity of substitution and discuss potential methodological reasons that might explain part of the observed variation. In section 4, different approaches to the microfoundation of production functions are discussed to better understand what underlies the EOS. In section 5, we outline the construction of an aggregate elasticity of substitution comprising multiple sectors. The influence of the institutional framework on σ is discussed in section 6. Finally, section 7 concludes and presents some suggestions for future research.

2 Background: σ and the CES production function

The concept of an elasticity of substitution between factors of production was developed by John R. Hicks in *The Theory of Wages* (1932) and Joan Robinson in *The Economics of Imperfect Competition* (1933). Interested in the effect of changes in the supply of productive factors on the distribution of factor income, Hicks obtained his definition as a byproduct of his analysis. In his setting, he considered a production function with constant returns to scale (CRS), $Y = F(K, L)$, that produces output, Y , as a combination of the two input factors capital, K , and labor, L , in a framework of perfect competition. On the basis of these assumptions, Hicks (1932, p. 244) defines the elasticity of substitution as follows:

$$(1) \quad \sigma_{KL} = \frac{p_L p_K}{\underbrace{p_Y^2 F_{LK} Y}_{\text{Hicks' def.}}} = \frac{F_L F_K}{F_{LK} Y},$$

where p_i with $i \in \{K, L\}$ denotes the price of either input factor, p_Y is the price of the output good, and $F_i = \partial Y / \partial i$ and $F_{ij} = \partial^2 Y / (\partial i \partial j)$ with $j \in \{K, L\}$ are the first and second/cross derivatives of the production function with respect to the inputs.¹ As noted by Kahn (1933) and Lerner (1933), a major drawback of definition (1) is that it lacks an intuitive economic interpretation. Thus, attention was directed to a second formulation that was independently developed by Robinson (1933) in her analysis of firm behavior under imperfect competition. To obtain an explicit measure of the technologically determined difficulty of substituting factors of production for each other, Robinson (1933, p. 256) defines the elasticity of substitution along a production isoquant with output fixed at \bar{Y} , as

¹Given perfect competition and a CRS production function, for cost-minimizing production, the marginal product of each input factor must be equal to its price, i.e., $F_i = p_i$. Moreover, since in general equilibrium theory only relative prices matter, the price of the final good can be normalized to one, i.e., $p_Y = 1$. Together, this establishes the equality of the two representations presented in equation (1).

$$(2) \quad \sigma_{KL} = \frac{\frac{d(K/L)}{K/L}}{\frac{d(F_L/F_K)}{F_L/F_K}} \bigg|_{Y=\bar{Y}} = \frac{\frac{d(K/L)}{K/L}}{\frac{d(p_L/p_K)}{p_L/p_K}} \bigg|_{Y=\bar{Y}},$$

where the same assumptions as mentioned above in the context of Hicks' definition apply. According to equation (2), the elasticity of substitution is defined as the percentage change in the capital-labor ratio due to a one percent change in the ratio of the marginal products of inputs, i.e., the marginal rate of technical substitution, along a given production isoquant (Helm, 1987).² Thus, following Robinson, the elasticity of substitution describes the shape of a production function. Soon after the introduction of these two seemingly competing definitions, it became clear that they are equivalent under the following assumptions: only two input factors, a constant returns to scale production function, and perfect competition.³

In the following years, production theory mostly maintained the assumption of a Cobb-Douglas production function, which implies a unitary elasticity of substitution that is reconcilable with the observed constancy of factor income shares for many developed economies. However, based on the empirical observation that this constancy does not hold in general, especially at the industrial level, Arrow et al. (1961) developed the more flexible CES production function:

$$(3) \quad Y = \gamma \left[\delta K^{\frac{\sigma-1}{\sigma}} + (1-\delta) L^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

where $\gamma > 0$ is a Hicks-neutral efficiency parameter, $\delta \in (0, 1)$ denotes the distribution parameter, and $\sigma \in [0, \infty]$ is the elasticity of substitution. As implied by its name, the CES production function is characterized by a constant elasticity of substitution along and across any of its production isoquants.⁴ It contains the Leontief ($\sigma \rightarrow 0$), Cobb-Douglas ($\sigma \rightarrow 1$) and linear ($\sigma \rightarrow \infty$) production function as special cases. Moreover, the CES production function in (3) is homothetic and exhibits constant returns to scale. Finally, following Acemoglu (2002), inputs are considered gross complements if $\sigma < 1$ and gross substitutes if $\sigma > 1$. This distinction also captures the difference in the asymptotic behavior of the marginal product of factor inputs:

²Since its introduction, a multitude of variations and generalizations of the elasticity of substitution have been developed, especially with respect to the case of more than two input factors. The most prominent versions are the *Allen-Uzawa* and *Morishima* elasticities of substitution. Stern (2011) presents a useful classification scheme of the various definitions and discusses how they are related.

³The equivalence of the two definitions under the conditions mentioned above is shown in Kahn (1933) and Hicks (1933). A formal and compact proof is presented in the second edition of Hicks' *The Theory of Wages* (1963, p. 373).

⁴Another, more general class of production functions comprises those with a variable elasticity of substitution (VES) that is positively or negatively dependent on the input factor ratio.

$$(4) \quad \lim_{K \rightarrow \infty} \frac{\partial Y}{\partial K} = \begin{cases} 0 & \text{if } \sigma < 1 \\ 0 & \text{if } \sigma = 1 \\ \gamma \delta^{\frac{\sigma}{\sigma-1}} & \text{if } \sigma > 1. \end{cases}$$

While, in the limit, the marginal product of capital approaches zero for gross complements, it does not fall below a certain lower bound if the inputs are gross substitutes.⁵ Intuitively, the higher the elasticity of substitution is, the more interchangeable are the inputs in production and the less pronounced is the diminishment of marginal returns (Brown, 1966). Additionally, if inputs are gross substitutes, they become inessential; that is, production is possible with only one of the two input factors: $\lim_{K \rightarrow 0} Y(K, L) = \gamma(1 - \delta)^{\frac{\sigma}{\sigma-1}} L > 0$ for $L > 0$ and $\lim_{L \rightarrow 0} Y(K, L) = \gamma \delta^{\frac{\sigma}{\sigma-1}} K > 0$ for $K > 0$. By contrast, for gross complements, production is possible only if both inputs are used in positive amounts: $\lim_{K \rightarrow 0} Y(K, L) = \lim_{L \rightarrow 0} Y(K, L) = 0$.

3 Heterogeneity in empirical estimates: a short summary

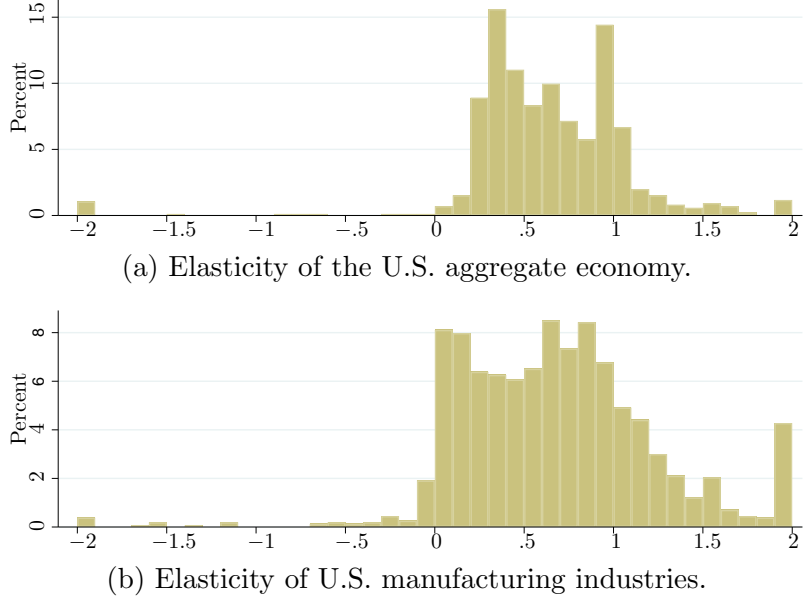
Since the introduction of the CES production function by Arrow et al. (1961), the elasticity of substitution between capital and labor has been estimated by a multitude of empirical studies at the aggregate, sectoral, and industrial level. The results of these estimates vary considerably. Figure 2a provides an illustration of this variation, depicting the distribution of 852 estimates of σ for the U.S. aggregate economy gathered from 49 studies published between 1961 and 2017.⁶ Most of the estimates cluster within the broad range of zero and somewhat above unity. Although a peak exists near 0.9 and 1, the mass of the results scatter in the neighborhood of 0.3 and 0.7. This pattern changes only slightly for the distribution of 1566 estimates of σ for U.S. manufacturing industries collected from 34 studies published between 1961 and 2017. As shown by figure 2b, the range of estimates broadens and clusters in the neighborhood of 0.1 and between 0.6 and 0.8. The diversity in estimation results, as documented by figure 2, is remarkable and explains the difficulty of obtaining agreement on an empirically backed consensus value for the elasticity of substitution. Furthermore, both the large variety of the estimation results and their clustering around values below unity stand in strong contrast to the frequent application of the Cobb-Douglas production function in theoretical models. To explain the observed heterogeneity in estimation results, past research has focused mainly on regularities that can be traced back to the specification of the estimation equation and to the appropriate modeling of technological change.⁷

⁵Note that a non-zero asymptotic marginal product violates one of the Inada conditions, according to which both the average and the marginal product of capital tend to zero in the limit.

⁶For an outline of empirical studies of σ for other countries, see Klump et al. (2007b).

⁷For a detailed discussion of further potential sources of heterogeneity in the estimation results, e.g., the quality and structure of the underlying data and the performance of various estimation techniques, see Berndt (1976), Hamermesh (1993), Antras (2004), Klump et al. (2007a), Chirinko (2008), Léon-Ledesma et al. (2010), Léon-Ledesma et al. (2015), and Knoblach et al. (2016).

Figure 2: Open-ended histogram with borders -2 and 2 of the elasticities of substitution of the U.S. aggregate economy and U.S. manufacturing industries, respectively.



Due to the non-linearity of its functional form, the CES production function does not permit a suitable analytical linearization, as is possible in the Cobb-Douglas case. Thus, standard linear regression techniques are not applicable to estimate the parameters of the CES production function. However, two alternative types of linear single-equation estimators exist for the elasticity of substitution.⁸ The first type was introduced by Arrow et al. (1961). Assuming constant returns to scale and fully competitive goods and factor markets, σ can be estimated by employing one of the two first-order conditions (FOC) of profit maximization. Based on equation (3), in log form, the FOCs can be written as follows:

$$(5) \quad \log \left(\frac{Y_t}{L_t} \right) = \sigma \log \left(\frac{1}{1 - \delta} \right) + (1 - \sigma) \log \gamma + \sigma \log \left(\frac{w_t}{p_t} \right),$$

$$(6) \quad \log \left(\frac{Y_t}{K_t} \right) = \sigma \log \delta + (1 - \sigma) \log \gamma + \sigma \log \left(\frac{r_t}{p_t} \right),$$

where w_t is the wage rate, r_t is the rental rate of capital, p_t is the price of the

⁸Additionally, non-linear least squares techniques can be applied to estimate CES functions. However, these approaches can suffer from convergence problems and local extrema (Henningsson and Henningsson, 2012).

output good, and t is an index representing time.⁹ Alternatively, the two first-order conditions can also be combined to construct a third estimation equation,

$$(7) \quad \begin{aligned} \log \left(\frac{K_t}{L_t} \right) &= \sigma \log \left(\frac{\delta}{1 - \delta} \right) + \sigma \log \left(\frac{w_t}{r_t} \right) \\ &= \sigma \log \left(\frac{\delta}{1 - \delta} \right) + \sigma \log \left(\frac{w_t L_t}{Y_t - w L_t} \frac{K_t}{L_t} \right), \end{aligned}$$

that builds upon cost minimization as it relates the capital-labor ratio to the corresponding relative factor prices.¹⁰ In previous research, (5) to (7), and combinations thereof, have extensively been applied to estimate σ .¹¹ Although estimations of σ should, theoretically, be independent of the estimation equation used, the results of empirical estimates typically differ. For instance, across several studies based on data for the U.S. economy (e.g., Eisner and Nadiri, 1968, Kalt, 1978, Léon-Ledesma et al., 2010, Young, 2013), the elasticity of substitution estimated from the FOC with respect to labor consistently exceeds that estimated with respect to capital.¹²

A second approach was proposed by Kmenta (1967), who suggested to expand the CES production function in a second-order Taylor series around the initial point $\sigma = 1$.¹³ This approximation allows for a linear estimation procedure. Theoretically, the major advantage of the Kmenta approximation is that it does not require any assumptions about the reward of factors and can simply be fitted to observable data on output and inputs. However, as demonstrated by Kmenta (1967), the error caused by using a Taylor approximation of the CES function can be large, especially if the input factor ratio or the elasticity of substitution is either very high or very low.¹⁴ To avoid problems commonly related to the single-equation approaches presented above, more recent estimates (e.g., Klump et al. (2007b), Young (2013), Herrendorf et al. (2015)) often rely on the so-called supply-side system approach. To estimate σ , a system of a linearized or non-linear production function and one or two FOC variants is established. Compared to single-equation approaches, the main

⁹Both equations can easily be transformed to account for non-constant returns to scale. For instance, an adjusted variant of the first-order condition with respect to labor may be written as $\log \left(\frac{Y_t}{L_t} \right) = \left(\frac{\nu}{\nu - \rho} \right) \log \left(\frac{1}{\nu \gamma (1 - \delta)} \right) + \left(\frac{\nu}{\nu - \rho} \right) \log \left(\frac{w_t}{p_t} \right) + \left(\frac{\rho(1 - \nu)}{\nu - \rho} \right) \log L$, where $\rho = \frac{\sigma - 1}{\sigma}$ refers to the substitution parameter and ν captures returns to scale. In the case of constant returns to scale, i.e., $\nu = 1$, the last term on the right-hand side drops out, and the equation reduces to (5).

¹⁰The rental rate of capital is typically not reported. Thus, the transformation in the second line of equation (7) exploits Euler's theorem on linear homogeneous functions, according to which $r = (Y - wL)/K$.

¹¹Early attempts to estimate σ typically apply (5), see, e.g., Solow (1964), Ferguson (1965), Sveikauskas (1974) and Takayama (1974). The most recent estimates based on (6) and (7) can be found in Chirinko and Mallick (2017). For a direct comparison of approaches (5) to (7), see Berndt (1976), Antras (2004), and Young (2013).

¹²Fuss (1977) as well as Berndt (1991) attribute these differences mainly to the putty-clay structure of capital stock.

¹³For a full derivation of the Kmenta approximation, see Henningsen and Henningsen (2011).

¹⁴Additional Monte Carlo results by Maddala and Kadane (1967), Thursby and Lovell (1978), and Léon-Ledesma et al. (2010) confirm that the Kmenta procedure usually does not provide reliable estimates of σ .

advantage of system estimation is that it can exploit variation in both optimization behavior (expressed by the FOCs) and technology (expressed by the production function). Combined with cross-equation parameter constraints, the resulting two- or three-equation system facilitates the identification of the structural parameters, e.g., technological progress or the elasticity of substitution (Klump et al., 2007b).¹⁵ Applied to the aggregate U.S. economy, equation systems typically generate estimates of σ considerably below unity.

In terms of technological dynamics, a modification of the standard CES production function that allows for biased technological change was proposed by David and Van de Klundert (1965). However, the majority of early studies neglect technological progress entirely or incorporate only a Hicks-neutral representation. This simplification may systematically bias the estimation results. For instance, as Antras (2004) revealed, the restriction to Hicks-neutral technological change appears to have biased the influential estimation results of Berndt (1976) towards unity.¹⁶ To account for this issue, Antras (2004) incorporated factor-augmenting technological change and found robust evidence that, for the aggregate U.S. economy, σ is significantly below unity. Contemporary studies allow for a more general functional form of technological progress. The application of the flexible Box and Cox (1964) transformation enables the incorporation of exponential, logarithmic, and hyperbolic growth patterns of factor efficiency and, hence, allows the data to select the appropriate functional form of technological dynamics. Using the Box-Cox transformation, σ has recently been estimated as approximately 0.7 for the Euro Area (Klump et al., 2008) and between 0.6 and 0.7 for the U.S. (Klump et al., 2007a,b).

Despite the multitude of observable regularities in the pattern of estimated elasticities, a thorough evaluation of potential biases in the estimation results was long neglected. Léon-Ledesma et al. (2010) were the first to systematically evaluate the performance of different estimation approaches to identify σ by performing extensive Monte Carlo simulations.¹⁷ The authors revealed that the three-equation system (production function plus two FOCs) is superior to two-equation systems (e.g., Berthold et al., 2002) and single-equation approaches. Furthermore, Léon-Ledesma et al. (2015) found that misspecification of technological progress can result in substantial biases in the estimation results of σ . On the basis of a comprehensive meta-regression analysis, these issues have been confirmed by Knoblauch et al. (2016). The authors reveal that heterogeneity in previously reported estimates for the U.S. aggregate economy is driven primarily by the different modeling decisions for technological dynamics. Based on extensive recent research effort, the assump-

¹⁵Contradictory evidence is obtained in Stewart and Li (2018). The authors note that the system approach does not surmount the Diamond impossibility theorem (Diamond et al., 1978), which proved the impossibility to identify the particular effects of σ and factor-augmenting technological change simultaneously for Canadian data.

¹⁶Unaware of this issue, for more than one-quarter of a century, the latter has been the major reference to justify the assumption of a Cobb-Douglas production function. Moreover, the confidence in the validity of the results provided by Berndt (1976) led to a sudden end of estimates based on CES production functions for many years.

¹⁷Notable forerunners include Maddala and Kadane (1966), Kumar and Gapinski (1974), Thursby and Lovell (1978), and Thursby (1980).

tion of a Cobb-Douglas production function appears to be rejected with considerable certainty, at least for U.S. aggregate production. Rather, the empirical evidence increasingly favors a below-unity elasticity of substitution in the broad range of 0.40 - 0.70 (Chirinko, 2008, Knoblich et al., 2016).

4 Microfoundation of production functions

After having reviewed the empirical literature on the extent of factor substitution, we now turn to what underlies σ from a theoretical perspective. The following section takes a close look at the literature on the microfoundation of production functions. This strand of literature derives production functions, specifically the Cobb-Douglas and CES, from deeper microeconomic principles.¹⁸ In such a framework, the elasticity of substitution is determined as a purely technological parameter. In the following, we discuss two of the most important approaches, the idea-based “endogenous technology choice” and the “mechanization” framework.¹⁹

4.1 Idea-based endogenous technology choice

The idea-based “endogenous technology choice” framework was proposed in Jones (2005) and has been further developed in Growiec (2008a,b, 2013, 2017), Matveenko (2010, 2011), and Matveenko and Matveenko (2015), among others. In this approach, production techniques are regarded as ideas that are discovered over time. A representative profit-maximizing firm is equipped with a specific production technique i , henceforth called the local production function (LPF). The LPF i combines capital, K , and labor, L , to produce output, Y , and is parameterized by its capital- and labor-augmenting parameters, $b_i \geq 0$ and $a_i \geq 0$, respectively. The LPF is defined as

$$(8) \quad Y = \tilde{F}(b_i K, a_i L),$$

where $\tilde{F}(\cdot, \cdot)$ is characterized by constant returns to scale and complementarity be-

¹⁸Note that most of these approaches are related to the “aggregation problem”, which casts doubt on the feasibility of obtaining a single production function from a set of micro-production functions. See García Molina (2005) and Felipe and McCombie (2013) for a profound discussion. In this context, an explicit discussion of the CES production function can be found in Schefold (2008).

¹⁹Additional approaches include models of search frictions (Lagos, 2006), the assignment of workers to tasks (Rosen, 1978, Dupuy, 2012), differential games of bargaining (Matveenko, 2013), and the concept of stochastic macro-equilibrium (Hiraguchi, 2015).

tween productive factors.²⁰ Due to this complementarity, the (specific) LPF i offers little or no possibility to react to a substantial change in the relative endowment of factors. Therefore, the firm is dependent on a new production technique (idea), defined as a pair (a, b) of factor-augmenting parameters, to combine both factors in a more “appropriate” manner in response to a change in the capital-labor ratio.²¹ The set of ideas to which the firm has access is characterized by a technology frontier function:

$$(9) \quad H(a, b) = N,$$

where the first derivatives are given by $H_a > 0$ and $H_b > 0$, respectively. This technology frontier captures the trade-off between the two factor-augmenting parameters. Techniques that are more efficient in using capital are less efficient in using labor, and vice versa. The parameter $N > 0$ determines the location of the frontier in the (a, b) space. A higher value of N , e.g., induced by Hicks-neutral technological progress, supports higher levels of a and b . Moreover, the shape of the technology frontier can either be deterministically given, as in Caselli and Coleman (2006) and Léon-Ledesma and Satchi (forthcoming), or it can emerge as the result of stochastic draws of ideas from certain distributions, as in Jones (2005) and Growiec (2008a,b, 2013). Finally, the endogenous technology choice can be modeled as an optimization problem subject to equations (8) and (9). The resulting global production function (GPF),

$$(10) \quad Y = F(K, L; N) = \max_{b,a} \tilde{F}(bK, aL),$$

is then described as the maximum amount of output, Y , a firm is able to produce from any given set of factor endowments, K and L , if it is free to choose between any production technique out of the stock of available ideas. Graphically, as a result of a smooth approximation to the sequence of linear combinations of the efficient subset of ideas, the GPF represents the convex hull of non-dominated local production techniques. In macroeconomics, this efficient subset is typically considered simply as “the technology”. However, based on the framework presented above, a production function should not be treated as a single technology but as an assembly

²⁰More precisely, the elasticity of substitution of the LPF, σ_{LPF} , is assumed to lie below unity. An extreme but reasonable assumption, particularly on the firm level, is that of a Leontief LPF with $\sigma_{LPF} = 0$, where the factors of production are used in fixed, predetermined proportions. This view is consistent with the “recipe” understanding of the LPF in Jones (2005), where a specific production technique is interpreted as a strictly defined set of instructions of how to transform inputs into output. However, if the LPF is understood as a sector-wide or country-wide production function, as in Caselli and Coleman (2006), higher values of σ_{LPF} might be more appropriate.

²¹The concept of “appropriate technologies”, according to which different factor endowments induce the use of specific production techniques, originated in Atkinson and Stiglitz (1969), where it is called “localized technology”, and has been studied further in Basu and Weil (1998) and Acemoglu and Zilibotti (2001).

of a multiplicity of local production techniques.²² The possibility of firms to switch between different known and available production techniques, and therefore the resulting elasticity of substitution of the GPF, can thus be studied with respect to i) the shape of the LPFs, ii) the distribution of ideas in the (a, b) space, and iii) the temporal and monetary adjustment costs of transition.

In his seminal contribution, Jones (2005) showed that independent of the shape of the local production function, the GPF converges to a Cobb-Douglas production function in the long-run if both factor-augmenting parameters, a and b , are randomly drawn from independent Pareto distributions.²³ Empirical evidence of the existence of Pareto distributions for scientific productivity can already be found in Lotka (1926) and has in the following been confirmed within a multitude of different economic contexts, including the upper tail of the income distribution, firm size (sales and employees), total factor-productivity, and innovation size (citations of patents and financial returns).²⁴ However, empirical evidence of the pattern of dependence between the two efficiency levels a and b is lacking. In a generalized model, Growiec (2008a) allows both parameters to be correlated according to the Clayton family of copulas.²⁵ For this assumption on the correlation, a particularly interesting case emerges when both Pareto distributions share the same shape parameter $\alpha > 0$. This case implies a CES result for the GPF with a global elasticity of substitution equal to

$$(11) \quad \sigma_{GPF} = \frac{\alpha\delta - \theta}{\alpha\delta - \theta - \alpha\delta\theta}$$

where $\delta \geq -1$ captures the degree and sign of dependence between the two Pareto distributions, $\alpha\delta$ refers to the curvature of the technology frontier, and $\theta = \frac{\sigma_{LPF}-1}{\sigma_{LPF}}$ is the substitution parameter of the LPF.²⁶ Equation (11) reveals that the global elasticity of substitution σ_{GPF} is driven by the difference between the curvature of

²²This perspective is consistent with Yuhn (1991, p. 344), where the elasticity of substitution is considered as “a menu of choice available to entrepreneurs”. The transition between different ideas or techniques has also been referred to as “microscale technological change” (Grübler et al., 1999, p. 547) and “technological substitution” (Sue Wing, 2006, p. 542).

²³Similar results can be found in Houthakker (1955-56) and Levhari (1968), where Leontief coefficients are allocated across “production cells”, e.g., firms or machines, based on a Pareto distribution. In the presence of the capacity constraints of individual firms, the Cobb-Douglas production function appears as a result of aggregation. For a useful generalization and more elaborate discussion of the results obtained by Jones (2005), see Matveenko (2010, 2011) and Growiec (2017). Specifically, Matveenko (2011) shows that a non-asymptotic derivation of a Cobb-Douglas production function can be obtained under the assumption of an exponential distribution of ideas.

²⁴See, for example, Reed (2001), Piketty and Saez (2003), Newman (2005), Luttmer (2007), Silverberg and Verspagen (2007), Gabaix (2009), Toda and Walsh (2015), and Gabaix et al. (2016).

²⁵For an introduction to the theory of copulas, see Nelsen (2006).

²⁶For $\delta < (>) 0$, the two parameters a and b are negatively (positively) correlated, whereas $\delta = 0$ refers to the case of independence, as analyzed in Jones (2005). To guarantee that σ is positive, $\alpha\delta - \theta - \alpha\delta\theta > 0$ is assumed. Furthermore, the assumption that the curvature of the technology frontier is greater than the curvature of the LPF, i.e., $\alpha\delta > \theta$, ensures an interior solution of the optimization problem described in (10).

the technology frontier $\alpha\delta$ and the curvature of the local production function θ : the greater the difference is, the lower is the value of σ_{GPF} . This approach nests the Leontief, Cobb-Douglas, and linear production function as special cases. As an alternative to the Pareto distribution assumption, Growiec (2008b, 2013) shows that a CES production function can emerge based on independent Weibull distributions for a and b . On the basis of these results, further empirical investigation of the distribution and dependence of factor productivities for different countries, sectors, and industries would be extremely valuable, as it has the potential to solve some of the empirical puzzles (e.g., heterogeneity in industrial estimates) described in the previous section.

In contrast to the local-global distinction outlined above, another strand of literature focuses on the distinction of the short-run and long-run substitutability between capital and labor. Early contributions rely on the putty-clay model of production, which was introduced by Johansen (1959) and has subsequently been further developed by Phelps (1963) and Solow (1962). The putty-clay model describes technologies for which factor proportions are variable *ex ante*, that is, before capital has been committed to concrete form, but for which variability disappears *ex post* (Solow, 1967). For example, in Caballero and Hammour (1998), the *ex ante* technology frontier is characterized by a CES production function with an elasticity of substitution considerably greater than unity. However, once a particular technique is chosen, substitutability disappears due to investment irreversibility, and the *ex post* technology reduces to a Leontief production function. Likewise, Gandolfo (2008) argues that production is more accurately described by a putty-clay structure. On the contrary, by imposing monetary friction on the choice of technology, Léon-Ledesma and Satchi (forthcoming) offer a novel method to obtain an increase in the elasticity of substitution over time. In the short-run, adjustment costs, e.g., due to the costly acquisition of advanced machinery and equipment, force the firm to maintain the current technique, which is characterized by gross complementarity. However, based on a continuous log-linear technology frontier, the long-run production function that emerges from technology choice is Cobb-Douglas. The speed of transition between the short- and long-run production function depends on the costs to switch between different techniques.

4.2 Capital accumulation and mechanization

A second approach to endogenously derive a production function is the mechanization framework developed in Nakamura and Nakamura (2008) and Nakamura (2009, 2010). This framework is based on the pioneering work of Zeira (1998, 2008), who explains economic growth via industrialization, during which machines replace workers in a growing number of tasks. Similar to the idea-based technology choice framework outlined above, the long-run production function is dynamically derived as the envelope of short-run production functions. The technology determining trade-off in the mechanization framework is modeled through the degree of mechanization, where the difficulty of its implementation is related to the elasticity of substitution of the corresponding production function.

More precisely, in this type of model, a representative profit-maximizing firm produces a final good by combining a continuum of intermediate goods, the total number of which is normalized to unity. Following Zeira (1998), each intermediate good $z(i)$, $i \in [0, 1]$, can be produced by two different techniques, where only one technique can be chosen at each point in time. The first is a manual technique that utilizes only labor with a factor-productivity of $\eta(i) > 0$. The second is an industrial technique that produces intermediate goods using capital with a factor-productivity of $\theta(i) > 0$. Mechanization occurs when more and more inputs are produced with capital rather than labor. On the basis of the two factor productivities $\eta(i)$ and $\theta(i)$, all intermediate goods $z(i)$ can be ranked starting from those in which capital is relatively more productive than labor to those in which capital is relatively less productive. The function

$$(12) \quad \Psi(i) \equiv \frac{\theta(i)}{\eta(i)}$$

indicates the relative productivity of capital to labor for all intermediates $z(i)$, where $\Psi(i) \geq 0$, $\partial\Psi(i)/\partial i < 0$, and $\Psi_0(i) > 0$. Thus, the shape of (12) captures the difficulty of mechanization. A relatively flat function $\Psi(i)$ indicates that industrial techniques are easy to implement; therefore, capital will be widely used in the production of intermediates, resulting in a high degree of mechanization. In comparison, mechanization is difficult in sectors or industries where $\Psi(i)$ is relatively steep; thus, the production of intermediates uses mainly manual techniques. In each period, based on the specific functional form of $\Psi(i)$ and the ratio of the rental rate of capital to the wage rate (r_t/w_t), the firm chooses the profit maximizing share of intermediates that are produced with the industrial technique, i.e., the optimal degree of mechanization. As the economy grows and more and more capital is available relative to labor, the ratio of factor prices decreases, leading to a growing number of intermediates becoming mechanized. As an illustration, in the basic model of Nakamura and Nakamura (2008) and Nakamura (2010), a final good is produced by a Cobb-Douglas aggregator, $\ln Y_t = \int_0^1 \ln[z(i)]di$, over all intermediates $z(i)$. That is, for each degree of mechanization, there exists a distinctive Cobb-Douglas production function. Moreover, the productivity of capital relative to labor is assumed to fall log-linearly over the continuum of intermediates according to:

$$(13) \quad \Psi_i = \left(\frac{b}{1-b} \right)^{1+\gamma} \left(\frac{1-i}{i} \right)^\gamma,$$

where $0 > b > 1$. The shape parameter $\gamma > 0$ indicates the difficulty of mechanization. A smaller γ represents easier mechanization. On the basis of these assumptions, a long-run production function of the CES type can be derived as an envelope over all Cobb-Douglas functions representing different degrees of mechanization:

$$(14) \quad Y_t = C \left(bK_t^{\frac{1}{1+\gamma}} + (1-b)L_t^{\frac{1}{1+\gamma}} \right)^{1+\gamma},$$

with $C = e^{-\gamma}$ and an elasticity of substitution equal to $\sigma = 1 + 1/\gamma$.²⁷ Equation (14) reveals the formal link between the difficulty of mechanization and the elasticity of substitution. Industries or sectors with a small γ , i.e., those that allow for a relatively easy implementation of mechanization, are characterized by a relatively high σ , and vice versa. In the past, it has been difficult to empirically identify industries or sectors with either a consistently high or low elasticity of substitution (see, e.g., Morawetz, 1976, Herrendorf et al., 2015, Chirinko and Mallick, 2017). The mechanization framework provides a sound theoretical explanation of why differences might exist among industries and sectors and stimulates deeper investigation of the structure of the underlying production process steps.

5 Aggregate elasticity of substitution

Thus far, we have focused on the elasticity of substitution as a purely technological parameter. Based on definition (2), the technological elasticity of substitution for the one-sector, two input framework is defined as the percentage change in the input factor ratio along a production isoquant in response to a one percent change in the ratio of marginal factor productivities. As shown in the last section, a production function comprising such a concept of the elasticity of substitution can be derived from a distribution of local production techniques. However, when extending the analysis to a multi-sectoral, two-input framework, a purely technological definition of the elasticity of substitution is not sufficient. As illustrated in figure 1, in addition to intra-sectoral factor substitution, in a general equilibrium setting, demand-side-induced inter-sectoral substitution of consumption or intermediate goods arises as a second channel through which the input factor ratio and the ratio of input factor prices are related. Specifically, demand-side-induced inter-sectoral substitution comprises two effects. First, *ceteris paribus*, a change in input factor prices, e.g., due to a change in factor endowment, changes the relative prices of produced goods. This change, in turn, motivates optimizing consumers to inter-sectorally adjust their consumption plans in favor of goods that have become relatively cheaper. These adjustments in goods consumption then have a feedback effect on factor demand and, thus, on factor prices. This feedback effect establishes the influence of commodity substitution on the relation between a change in the input factor ratio and the ratio of input factor prices, i.e., the elasticity of substitution. Second, for non-homothetic preferences, an increase in factor endowment can result in an income effect that changes the distribution of consumption expenditures and stimulates further inter-sectoral adjustment. Analogous to the first effect, the feedback effect of changing consumption patterns on input factor demand results in a change of the aggregate elasticity of substitution. Accordingly, a sole focus on technological substitutability might be misguided in a multi-sectoral economy and should therefore be replaced by a broader concept of factor substitution, i.e., the aggregate elasticity of substitution,

²⁷In Nakamura (2009), a long-run CES production function in which σ takes any positive value is dynamically derived from short-run Leontief production functions. Furthermore, Nakamura (2010) implements a learning effect that offset the decline in the productivity of capital to derive a long-run VES production function.

which comprises both technological and non-technological substitution.

The first comprehensive and formal analysis of the AES in a multi-sectoral framework was provided by Jones (1965). Two separate sectors are included in his general equilibrium framework, each producing a specific commodity good by means of both capital and labor. The two sectors are characterized by perfect competition, constant returns to scale, and their own, sector-specific technological elasticity of factor substitution. The model is closed by a generic, homothetic utility function. Homotheticity implies that there are no income effects and that the composition of consumption is dependent on only relative commodity prices. Assuming full mobility of factors, Jones (1965) showed that the aggregate elasticity of factor substitution can be written as the weighted arithmetic mean of each of the two technological elasticities of factor substitution and the elasticity of commodity substitution:²⁸

$$\begin{aligned}
 (15) \quad \sigma_{\text{AES}} &= Q_1\sigma_1 + Q_2\sigma_2 + Q_C\sigma_C \\
 &= Q_1\sigma_1 + Q_2\sigma_2 + (1 - Q_1 - Q_2)\sigma_C,
 \end{aligned}$$

where σ_i with $i \in \{1, 2, C\}$ are the two sector-specific technological elasticities and commodity substitution, respectively. Moreover, Q_i are the endogenously determined factor endowment-dependent weights, with $\sum_i Q_i = 1$.²⁹ McManus (1988) confirms the result of Jones (1965) and shows that $\sum_i Q_i = 1$ is a direct implication of the assumption of constant returns to scale in production, which does not hold for increasing or decreasing returns. In particular, if both sectors of production exhibit increasing (decreasing) returns to scale, the weights sum to less (more) than one. However, the AES remains a linear combination of the two technological elasticities and of commodity substitution. Moreover, the weights, Q_i , in general, vary as an economy's factor endowment changes. Thus, the aggregate elasticity of substitution, σ_{AES} , is also generally not constant. This result remains true if the three primary elasticities, σ_i , are constant but not identical. In addition, as noted by Jones (1965), even if both production functions are Leontief with $\sigma_1 = \sigma_2 = 0$, the AES can still be positive due to commodity substitution. Finally, a version of the AES incorporating a non-homothetic Stone-Geary-type utility function is presented in Osumi (2015). In that case, the AES is again a linear combination of the three primary elasticities but is no longer a weighted average of them, i.e., $\sum_i Q_i \neq 1$.

Although Arrow et al. (1961, p. 241) already conjectured that “the of economic development itself might shift the over-all elasticity of substitution”, the first systematic analysis of the relation between economic development, measured in terms of an increasing capital-labor ratio, and the aggregate elasticity of substitution was performed by Miyagiwa and Papageorgiou (2007). The authors build upon the

²⁸Note that in a general equilibrium framework, it is not changes in the ratio of the input factor prices that change factor employment but rather changes in input factor endowment that cause a change in input factor prices. This reversed interpretation of the elasticity of substitution can be expressed by $1/\sigma_{\text{AES}}$.

²⁹See Jones (1965) for a detailed derivation of the AES and the composition of the endogenous, factor endowment-dependent weights Q_i .

framework of Jones (1965) and assume a CES function for the commodity aggregator (utility) and Cobb-Douglas production functions for the two manufacturing sectors (CES-CD-CD).³⁰ Embedded in a classical Solow (1956) growth framework, the authors generally find a positive relation between capital intensity, k , and the AES. This result not only confirms the Arrow et al. (1961) conjecture but is also supported by empirical evidence. For instance, Duffy and Papageorgiou (2000) find that the AES is higher in richer countries than in poorer countries. Recently, Xue and Yip (2013) have developed a unifying framework that includes the full factor mobility, CES-CD-CD specification of Miyagiwa and Papageorgiou (2007) as a special case. In general, the monotone positive relation between capital intensity and the aggregate elasticity of substitution is confirmed. However, Xue and Yip (2013) also show that this positive relation becomes ambiguous for the case of sector-specific factor inputs, that is, where either capital or labor alone is used in one but not the other of the two manufacturing sectors while the other factor is used in and is fully mobile across both sectors. In that case, the relation between capital intensity and the AES can be hump-shaped, U-shaped, or positive/negative and depends on both the relative size of the two primary elasticities and their directions, i.e., whether the sectoral production or consumption inputs are (gross) complements or substitutes. Technically, in the specific-factor case, the AES is determined as the weighted harmonic mean of the elasticity of substitution in the sector employing both inputs and of commodity substitution:

$$(16) \quad \sigma_{\text{AES}}^{\text{harmonic}} = \left[Q_1 \frac{1}{\sigma_1} + Q_C \frac{1}{\sigma_C} \right]^{-1}$$

where $\sum_i Q_i = 1$. Finally, with respect to the size of the aggregate elasticity of substitution and its underlying primary elasticities, the following relation applies:

$$(17) \quad \min(\sigma_i) \leq \underbrace{\sigma_{\text{AES}}^{\text{harmonic}}}_{\substack{\text{"specific} \\ \text{capital/labor"}}} \leq \underbrace{\sigma_{\text{AES}}^{\text{arithmetic}}}_{\text{Jones (1965)}} \leq \max(\sigma_i).$$

The main insight from the framework of Jones (1965) is that in a general equilibrium setting, the AES is always a (weighted) linear combination of technological intra-sectoral substitution and of non-technological commodity substitution. Moreover, the overview above highlights the crucial difference between the purely technological elasticity of factor substitution and the aggregate elasticity of substitution. While the first concept is strongly linked to the problem of the single firm, as analyzed by Robinson (1933), the AES is essential for an understanding of the relation between economic growth and the distribution of factor income shares in a general equilibrium framework. However, as highlighted above, the relation between the AES and economic development is generally ambiguous. Thus, the identification of the conditions necessary for either a positive or negative relation remains a challenging

³⁰Actually, in Miyagiwa and Papageorgiou (2007), two intermediate goods are aggregated to a final good via an additional production function. However, mathematically, this is equivalent to the use of an aggregating utility function with two final goods, as in Jones (1965).

task for both theoretical and empirical research. Moreover, given recent theoretical work on the positive impact of σ on growth (de La Grandville, 1989, Klump and de La Grandville, 2000), there could arise a self-accelerating or self-decelerating feedback effect. However, the impact of a larger σ on growth, as studied in these models, relies on a single representative CES production function, whereas the effect of economic development on substitutability, as presented in this section, has thus far been analyzed in only an AES framework. Thus, an extension of the analysis of the relation between the AES and economic growth might substantially improve our understanding of growth, especially of possible convergence or divergence dynamics.

6 Institutional framework

In the previous sections, we treated the elasticity of substitution as independent of the institutional framework in which a technology is used. However, institutions can have an important effect on the degree to which technologically feasible substitution possibilities can actually be carried out. With respect to the different concepts of substitutability distinguished in this paper, institutional characteristics are able to explain part of the difference between technological and non-technological substitution, as represented by the EOS or AES concept and the effective elasticity of substitution (EES) typically observed in empirical studies. In this section, we provide an overview of some institutional characteristics that are often suggested as determinants of the elasticity of substitution. We illustrate how these characteristics are supposed to affect σ and evaluate which of these institutional characteristics are supported empirically.

6.1 Openness to trade

Hicks (1936) already mentioned international trade as an institutional factor that may influence the aggregate elasticity of substitution of an economy. Early contributions that formalize the link between openness to trade and the effective elasticity of substitution include Azariadis (1996) and Ventura (1997). In Ventura (1997), a final good is produced with a CES production function,

$$(18) \quad Y = \left(Y_1^{\frac{\sigma-1}{\sigma}} + Y_2^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

that combines two intermediates, Y_1 and Y_2 , where σ refers to the elasticity of substitution between inputs. Each intermediate is produced by a linear technology, where one unit of capital generates one unit of intermediate 1, while one worker produces A units of intermediate 2. Both intermediates can be traded internationally at prices p_1 and p_2 , respectively. By adapting the framework to a small open economy with fixed terms of trade (p_1/p_2) and without any discrimination between domestic and foreign intermediates, free international trade mimics a de facto linear aggregate technology, $Y = p_1 AL + p_2 K$, where the effective elasticity of substitution between

capital, K , and labor, L , is infinite. Although highly stylized, the framework provided by Ventura (1997) demonstrates the primal influence of the state of trade liberalization on the AES, as it allows indirect substitution of both input factors by each other.³¹ Generally, in a growing economy, international trade enables the conversion of the production of capital-intensive intermediates into labor-intensive imports such that an increasing capital-labor ratio does not necessarily translate into capital-deepening in production. Thus, international trade can help to alleviate the decreasing marginal returns to capital that are usually associated with capital accumulation. However, trade liberalization does not necessarily increase the AES. Saam (2008) develops a Heckscher-Ohlin-Samuelson model with two large countries that trade intermediate goods under incomplete specialization.³² Both countries are equipped with identical CES technologies and differ solely in their rate of capital accumulation. In such a setting, it can be shown that while both countries gain from trade, openness to trade increases the AES of the country with higher growth in capital but decreases the AES of the country with lower growth in capital. If, on the other hand, the same growth rates of capital are assumed for both countries, a switch from autarky to trade would not have any influence on the AES. However, as Irmen (2008) notes, openness to trade appears in Saam (2008) merely as a co-determinant that affects the AES only implicitly through differences in the rate of capital accumulation but does not change the AES by itself. Generally, in all the models presented above, the technological elasticity of substitution is completely unaffected by trade. However, as demonstrated by Eaton and Kortum (1999), trade and foreign direct investment (FDI) can be an important source of technology diffusion. Further theoretical considerations should, for instance, investigate the modeling of technology transfers from industrialized to developing countries.

6.2 Central planning

Another institutional factor often presented as a potential determinant of the elasticity of substitution is “inclination to socialist ideas” (Mallick, 2012, p. 683). This consideration is rooted in an old debate on the reasons for the decline in Soviet GDP growth in the period after 1950.³³ According to the “extensive growth” hypothesis (Ofer, 1987, p. 1786), a low elasticity of substitution causes returns to capital to diminish quickly and therefore growth to slow. With respect to possible reasons for the low Soviet elasticity of substitution, Easterly and Fischer (1995) speculate that centrally planned economies might not be able to develop the broad variety of capital goods necessary to effectively substitute for labor. Moreover, countries with a high degree of state intervention and planning might miss “such market-oriented

³¹See Klump (2001) for a discussion and generalization of the results provided by Ventura (1997).

³²In some respects, this setup places the AES of the closed economy studied in Miyagiwa and Papageorgiou (2007) into an open economy setting. Furthermore, Miyagiwa (2008) applies a Ricardo-Viner model with land as an additional input factor to show that openness to trade increases the aggregate elasticity of substitution.

³³As Weitzman (1970) notes, the puzzling characteristic of the retardation in Soviet output growth is that it was not accompanied by a sufficiently large slowdown in the growth of input factors.

types of physical and human capital as entrepreneurial skills, marketing and distributional skills, and information-intensive physical and human capital” (Easterly and Fischer, 1995, p. 363) so that both the purely technological and the effective elasticity of substitution are low. Similarly, Nakamura (2015) argues that due to the administrative command organization of the Soviet economy, prices were distorted and thus did not provide the information and incentives necessary for an efficient substitution between capital and labor. Moreover, the author argues that Soviet companies might simply not have had the authority to independently substitute capital for labor; thus, the low Soviet effective elasticity of substitution might simply have been the result of rigid economic policies. Although all these arguments appear to be reasonable, to the best of our knowledge, no systematic analysis exists of whether centrally planned economies or countries with a high degree of government intervention generally exhibit a significantly lower elasticity of substitution between capital and labor than do market-oriented economies. Moreover, as Nakamura (2015) notes, it is not even clear whether the elasticity of substitution estimated for the Soviet Union was actually low compared to that of other countries. Estimates of the elasticity of substitution for the Soviet Union range from as low as 0.04 up to 0.40.³⁴ In comparison, for the time period 1950 - 2000, Mallick (2012) estimates a mean elasticity of 0.34 for the sample of OECD countries, with values as high as 1.19 (Sweden) and 0.64 (U.S.) and as low as 0.09 (Denmark) and 0.22 (France). Given the large variation in the estimation results of the Soviet elasticity of substitution and the heterogeneous results for Western (OECD) economies, the idea of a general (and unambiguous) link between a country’s economic system and the elasticity of substitution between input factors should be treated with caution.

6.3 Labor market regulations and trade unions

Substitution might also be limited by labor market regulations and institutional factors, such as the degree of unionization. This argument applies to both intra-sectoral and inter-sectoral substitution. As de La Grandville (2016, p. 153) notes, economies with “strong customary or regulatory barriers to large changes in capital-labor ratios” are likely to be characterized by an effective elasticity of substitution that is lower than what would be expected based on the underlying EOS or AES.

At the intra-sectoral level, Freeman and Medoff (1982) and Maki and Meredith (1987) provide empirical evidence of a negative effect of strong labor unions on the elasticity of substitution for the manufacturing sector in the U.S. and Canada, respectively. However, the authors themselves also note that their results potentially suffer from reverse causality, as labor unions might self-select into industries with a low technological elasticity of substitution, where wage gains need not be traded-off with large losses of employment. In a recent study relying on an instrumental variable approach to control for the potential endogeneity of unionization, Shahiri and Osman (2017) reject the hypothesis of a negative effect of unionization on the elasticity of substitution for the U.S. postal service industry. The mixed evidence presented above casts doubt on the idea of a general negative effect of unionization

³⁴We refer to Nakamura (2015) for an extensive comparison of these estimations.

on the effective elasticity of substitution. To the best of our knowledge, no analysis has been performed on the aggregate level or for a broad range of industries. We suggest that future research should take into account the fact that the effect of labor unions on the elasticity of substitution is probably asymmetric, i.e., labor unions probably only restrict the substitution of capital for labor and not the substitution of labor for capital. A proper incorporation of this peculiarity might substantially improve the empirical analyses.

6.4 Political intervention and the financial system

In an investigation of the economic development of South Korea in the second half of the last century, Yuhn (1991) identified the relatively high elasticity of substitution as one of the main drivers of the impressive growth performance. Moreover, the author argues that South Korea's high elasticity is a direct consequence of political intervention with the intention to keep the price of capital input artificially low. However, since σ , as defined by equation (2), measures the change in the optimal input factor ratio in reaction to a change in input factor prices along a production isoquant, we do not regard South Korea's expansive monetary policy as a determinant of either the technological or the effective elasticity of substitution. More precisely, interventions with respect to the interest rate can have an effect on the elasticity of substitution only if, at the same time, capital supply is restricted such that it does not meet demand. Technically, this implies that the marginal (value) product of capital is above the interest rate. Only in this case do the technological and effective elasticity of substitution not coincide. However, although we reject the interest rate as a direct determinant of the elasticity of substitution, we agree with Yuhn (1991) that policies, such as low interest rate loans, foreign exchange controls, low tariffs on imported raw materials, longer tax holidays, shorter write-off periods, and accelerated depreciation on capital goods, can affect the elasticity of substitution, either through their impact on the actual use of available substitution possibilities (effective elasticity) or through their influence on the incentives shaping the range of technologies available to firms (technological elasticity). For instance, cheap capital might induce the development of new technologies that allow for easier substitution of capital for labor or simply reduce the effective transition costs between techniques. However, for none of the above mentioned policies does there exist a sound theoretical foundation or an empirical evaluation of their potential impact on the elasticity of substitution.

Another channel through which political intervention may change the elasticity of substitution was proposed by Klump and Preissler (2000), where the authors speculate that efficient factor substitution requires a well-functioning monetary and financial system. Although the stage of development and efficiency of the financial system have no direct influence on the technological elasticity of substitution, they may well affect the effective elasticity of substitution, e.g., by providing liquidity to finance costly adjustments associated with substitution processes. However, there exists no study that theoretically analyzes or empirically assesses the relationship between the performance of the fiscal and monetary system on the one side and the

effective elasticity of substitution on the other side.

7 Conclusion and future research

The elasticity of substitution between capital and labor plays a crucial role in a broad variety of economic topics. However, despite its importance and extensive empirical research, no consensus has emerged in the literature in terms of the factors that underlie and determine σ . This lack of insight is even more puzzling as empirical evidence suggests large variation in the value of σ for different countries, specific sectors, and industries, as well as over time. Motivated by this observation, the purpose of this paper was to provide an overview of the existing literature on the determinants of the elasticity of substitution between capital and labor. On the basis of this review, we showed that the effective elasticity of substitution of an economy is influenced by several technological, non-technological, and institutional determinants and therefore should not be treated as a immutable deep parameter but rather as an endogenous variable.

The existing literature can be extended in several directions. For future research we propose the following, incomplete list of possible extensions and improvements:

- I. *Empirical considerations:* Although considerable effort has been made in estimating the CES production function for the U.S., other countries, especially developing countries, are less well investigated. Further investigations are recommended, as some evidence exists for significant cross-country differences in σ (Duffy and Papageorgiou, 2000, Masanjala and Papageorgiou, 2004, Mallick, 2012). Possible explanations for that finding are scarce. The same is true for cross-industrial and cross-sectoral differences in σ .³⁵ Furthermore, although the elasticity of substitution is traditionally considered to be time-invariant, some evidence of a time-dependent variation exists. For instance, de La Grandville and Solow (2017) identify a slight positive drift in the elasticity of substitution over time for a panel of 16 OECD countries. Further evidence of a positive time trend in σ for the aggregate U.S. economy can be found in Knoblach et al. (2016) and Cantore et al. (2017). However, from an empirical perspective, it remains unclear what the driving force behind such an increase could be. With respect to the estimation approach, despite their advantages, the flexible Box-Cox transformation and the three-equation system have been applied only to selected industrialized countries. Additionally, although almost universally applied, the assumptions of perfect competition and constant returns to scale appear to be invalid for the product and factor markets in most industries and countries.
- II. *Microfoundation:* We view the literature on the microfoundation of production functions as the basis for investigation of “sigma-augmenting technological change” (Klump et al., 2012, p. 793). A good starting point could be the

³⁵See, for instance, the inverted sectoral ranking of σ in Herrendorf et al. (2015) and Chirinko and Mallick (2017).

development of models that explicitly microfound the functional form of the technology frontier as, for instance, applied in Growiec (2008a, 2013). A refinement of the technology concept that accounts for different sector-specific characteristics might be another important task for future research. Finally, the elasticity of substitution could also be modeled as being directly dependent on past investments in sigma-augmenting research. Such a reduced-form investment approach has been used in the context of resource economics (Growiec and Schumacher, 2008, Fenichel and Zhao, 2015) but lacks a microfoundation of research investment similar to to endogenous efficiency enhancing technological change (Romer, 1990, Grossman and Helpman, 1991, Aghion and Howitt, 1992)

- III. *Aggregate elasticity of substitution:* The concept of an aggregate elasticity of substitution has been analyzed for only very specific conditions. For instance, the assumption of homothetic consumer preferences was relaxed for the first and only time by Osumi (2015), who derived the AES for a Stone-Geary-type utility function. Additional research on the impact of non-homothetic preferences is highly desirable to better understand how economic development might influence the AES through changing consumption patterns. Similarly, in the existing theoretical literature, the effect of non-constant returns to scale has been mentioned only once (McManus, 1988) and should receive further attention to capture the specific characteristics of different markets and aggregation levels. Moreover, the AES has thus far been derived for only the two-input case. Given the increasing importance of additional inputs, e.g., materials and energy and different types of labor, multi-input definitions of the elasticity of substitution might be a beneficial generalization. Similarly, a generalization of the AES to a framework with more than two sectors of production would be an important improvement, especially with respect to the growing literature on sectoral transformation. Finally, it would be interesting to investigate how the elasticity of substitution calculated from an AES framework differs from that estimated with a single CES production function.³⁶
- IV. *Institutional aspects:* With respect to institutional determinants of σ , a thorough empirical assessment of their impact is probably the most important task for future research. In this context, the major difficulty is to properly separate the elasticity of substitution into its technological, non-technological, and institutional components. For instance, in the case of labor unions, a possible solution might be to explicitly model the influence of institutions on factor employment. Additionally, an analysis of the effect of partial trade liberalization on the AES in the framework provided by Saam (2008) would be an interesting extension. Finally, for both the empirical and theoretical analysis of the influence of the institutional framework on the elasticity of substitution, we stress the importance of clearly distinguishing between those cases where the technological elasticity of substitution, i.e., the set of available techniques

³⁶To some extent, this exercise has been performed in Chirinko and Mallick (2017) based on U.S. data.

itself, is influenced and those cases where only the possibility of applying existing technological substitution is affected. In this context, the explicit modeling of the effect of institutional characteristics on the development of new techniques and the employment of already available techniques is a promising task for future research.

To conclude our review, we remind the researchers of the multi-dimensional character of the elasticity of substitution between capital and labor. That is, the effective elasticity of substitution, as estimated at the level of an economy as a whole, is always a combination of technological, non-technological, and institutional determinants. Awareness of this fact is important for several reasons. First, one must be careful when using the estimated value of σ in the context of theoretical models that are based on a purely technological notion of the elasticity of substitution, e.g., the calibration of numerical simulations of growth models. Similarly, the estimated elasticity of substitution may be greater than unity, even if the technological elasticity of substitution is well below unity in all producing sectors. Thus, an estimated elasticity greater than unity does not necessarily imply that production is possible with capital alone. Second, as we have shown, the effective elasticity of substitution of an economy is generally not constant but varies over time. These theoretical results should also find greater support in the empirical literature, e.g., by the (re)consideration of more flexible approaches, such as the variable elasticity of substitution (VES) production function.

References

- Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, 69(4):781–809.
- Acemoglu, D. (2003). Labor- and Capital-Augmenting Technical Change. *Journal of the European Economic Association*, 1(1):1–37.
- Acemoglu, D. and Zilibotti, F. (2001). Productivity Differences. *The Quarterly Journal of Economics*, 116(2):563–606.
- Aghion, P. and Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2):323–351.
- Alvarez-Cuadrado, F., Van Long, N., and Poschke, M. (2017). Capital-Labor Substitution, Structural Change and Growth. *Theoretical Economics*, 12(3):1229–1266.
- Antras, P. (2004). Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution. *Contributions to Macroeconomics*, 4(1).
- Arrow, K. J., Chenery, H. B., Minhas, B. S., and Solow, R. M. (1961). Capital-Labor Substitution and Economic Efficiency. *The Review of Economics and Statistics*, 43(3):225–250.
- Atkinson, A. B. and Stiglitz, J. E. (1969). A New View of Technological Change. *The Economic Journal*, 79(315):573–578.
- Azariadis, C. (1996). The Economics of Poverty Traps Part One: Complete Markets. *Journal of Economic Growth*, 1(4):449–486.
- Basu, S. and Weil, D. N. (1998). Appropriate Technology and Growth. *The Quarterly Journal of Economics*, 113(4):1025–1054.
- Berndt, E. R. (1976). Reconciling Alternative Estimates of the Elasticity of Substitution. *The Review of Economics and Statistics*, 58(1):59–68.
- Berndt, E. R. (1991). *The Practice of Econometrics: Classic and Contemporary*. Addison-Wesley.
- Berthold, N., Fehn, R., and Thode, E. (2002). Falling Labor Share and Rising Unemployment: Long-Run Consequences of Institutional Shocks? *German Economic Review*, 3(4):431–459.
- Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.
- Brown, M. (1966). *On the Theory and Measurement of Technological Change*. Cambridge University Press.

- Caballero, R. J. and Hammour, M. L. (1998). Jobless Growth: Appropriability, Factor Substitution, and Unemployment. *Cornegie-Rochester Conference Series on Public Policy*, 48:51–94.
- Cantore, C., Ferroni, F., and Léon-Ledesma, M. A. (2017). The Dynamics of Hours Worked and Technology. *Journal of Economic Dynamics and Control*, 82:67–82.
- Cantore, C., Léon-Ledesma, M. A., McAdam, P., and Willman, A. (2014). Shocking Stuff: Technology, Hours, and Factor Substitution. *Journal of the European Economic Association*, 12(1):108–128.
- Caselli, F. (2005). Accounting for Cross-Country Income Differences. In Durlauf, P. A. and Steven, N., editors, *Handbook of Economic Growth*, volume 1, Part A, pages 679–741. Elsevier.
- Caselli, F. and Coleman, II, W. J. (2006). The World Technology Frontier. *American Economic Review*, 96(3):499–522.
- Chirinko, R. S. (2002). Corporate Taxation, Capital Formation, and the Substitution Elasticity Between Labor and Capital. *National Tax Journal*, 55(2):339–355.
- Chirinko, R. S. (2008). σ : The Long and Short of it. *Journal of Macroeconomics*, 30(2):671–686.
- Chirinko, R. S. and Mallick, D. (2017). The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model. *American Economic Journal: Macroeconomics*, 9(4):225–253.
- David, P. A. and Van de Klundert, T. (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, O. (1989). In Quest of the Slutsky Diamond. *The American Economic Review*, 79(3):468–481.
- de La Grandville, O. (2016). *Economic Growth: A Unified Approach*. Cambridge University Press, second edition.
- de La Grandville, O. and Solow, R. M. (2017). Capital-Labour Substitution and Economic Growth. In *Economic Growth - A Unified Approach*, pages 114–155. Cambridge University Press, Cambridge.
- Diamond, P., McFadden, D., and Rodriguez, M. (1978). *Production Economics: A Dual Approach to Theory and Applications*, volume 2, chapter Measurement of the Elasticity of Substitution and Bias of Technical Change, pages 125–147. McMaster University Archive for the History of Economic Thought, Amsterdam.
- Duffy, J. and Papageorgiou, C. (2000). A Cross-Country Empirical Investigation of the Aggregate Production Function Specification. *Journal of Economic Growth*, 5(1):87–120.

- Dupuy, A. (2012). A Microfoundation for Production Functions: Assignment of Heterogenous Workers to Heterogeneous Jobs. *Economica*, 79(315):534–556.
- Easterly, W. and Fischer, S. (1995). The Soviet Economic Decline. *The World Bank Economic Review*, 9(3):341–371.
- Eaton, J. and Kortum, S. (1999). International Technology Diffusion: Theory and Measurement. *International Economic Review*, 40(3):537–570.
- Eisner, R. and Nadiri, M. I. (1968). Investment Behavior and Neo-Classical Theory. *The Review of Economics and Statistics*, 50(3):369–382.
- Felipe, J. and McCombie, J. S. L. (2013). *The Aggregate Production Function and the Measurement of Technical Change*. Edward Elgar Publishing.
- Fenichel, E. P. and Zhao, J. (2015). Sustainability and Substitutability. *Bulletin of Mathematical Biology*, 77(2):348–367.
- Ferguson, C. E. (1965). Time-Series Production Functions and Technological Progress in American Manufacturing. *Industry Journal of Political Economy*, 73(2):135–147.
- Freeman, R. B. and Medoff, J. L. (1982). Substitution Between Production Labor and Other Inputs in Unionized and Nonunionized Manufacturing. *The Review of Economics and Statistics*, 64(2):220–233.
- Fuss, M. A. (1977). The Structure of Technology Over Time: A Model for Testing the "Putty-Clay" Hypothesis. *Econometrica*, 45(8):1797–1821.
- Gabaix, X. (2009). Power Laws in Economics and Finance. *Annual Review of Economics*, 1:255–294.
- Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The Dynamics of Inequality. *Econometrica*, 84(6):2071–2111.
- Gandolfo, G. (2008). Comment on C.E.S. Production Functions in the Light of the Cambridge Critique. *Journal of Macroeconomics*, 30(2):798–800.
- García Molina, M. (2005). Capital Theory and the Origins of the Elasticity of Substitution (1932-35). *Cambridge Journal of Economics*, 29(3):423–437.
- Grübler, A., Nakicenovic, N., and Victor, D. G. (1999). Modeling Technological Change: Implications for the Global Environment. *Annual Review of Energy and the Environment*, 24:545–569.
- Grossman, G. M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *The Review of Economic Studies*, 58(1):43–61.
- Growiec, J. (2008a). A New Class of Production Functions and an Argument Against Purely Labor-Augmenting Technical Change. *International Journal of Economic Theory*, 4(4):483–502.

- Growiec, J. (2008b). Production Functions and Distributions of Unit Factor Productivities: Uncovering the Link. *Economics Letters*, 101(1):87–90.
- Growiec, J. (2013). A Microfoundation for Normalized CES Production Functions with Factor-Augmenting Technical Change. *Journal of Economic Dynamics and Control*, 37(11):2336–2350.
- Growiec, J. (2017). Factor-Specific Technology Choice. *Journal of Mathematical Economics*, 77(1):1–14.
- Growiec, J. and Schumacher, I. (2008). On technical change in the elasticities of resource inputs. *Resources Policy*, 33(4):210–221.
- Hamermesh, D. S. (1993). *Labor Demand*. Princeton University Press.
- Helm, D. R. (1987). Elasticity of Substitution. *The New Palgrave: A Dictionary of Economics*, 2:127–128.
- Henningsen, A. and Henningsen, G. (2011). Econometric Estimation of the "Constant Elasticity of Substitution" Function in R: Package micEconCES. IFRO Working Paper, Institute of Food and Resource Economics, University of Copenhagen.
- Henningsen, A. and Henningsen, G. (2012). On Estimation of the CES Production Function - Revisited. *Economics Letters*, 115(1):67–69.
- Herrendorf, B., Herrington, C., and Valentinyi, A. (2015). Sectoral Technology and Structural Transformation. *American Economic Journal: Macroeconomics*, 7(4):104–133.
- Hicks, J. R. (1932). *The Theory of Wages*. MacMillan, London, 1st edition.
- Hicks, J. R. (1933). IV.—A Note on Mr. Kahn's Paper. *The Review of Economic Studies*, 1(1):78–80.
- Hicks, J. R. (1936). Distribution and Economic Progress: A Revised Version. *The Review of Economic Studies*, 4(1):1–12.
- Hicks, J. R. (1963). *The Theory of Wages*. MacMillan, London, 2nd edition.
- Hiraguchi, R. (2015). On the Stochastic Macro-Equilibrium and a Microfoundation for the Production Function. *RIETI Discussion Paper Series 15-E-040*.
- Houthakker, H. S. (1955–1956). The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis. *The Review of Economic Studies*, 23(1):27–31.
- Irmen, A. (2008). Comment on "On the Openness to Trade as a Determinant of the Macroeconomic Elasticity of Substitution". *Journal of Macroeconomics*, 30(2):703–706.

- Johansen, L. (1959). Substitution Versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis. *Econometrica*, 27(2):157–176.
- Jones, C. I. (2005). The Shape of Production Functions and the Direction of Technical Change. *The Quarterly Journal of Economics*, 120(2):517–549.
- Jones, R. W. (1965). The Structure of Simple General Equilibrium Models. *Journal of Political Economy*, 73(6):557–572.
- Jones, R. W. and Ruffin, R. J. (2008). Trade and Wages: A Deeper Investigation. *Review of International Economics*, 16(2):234–249.
- Kahn, R. F. (1933). III.—The Elasticity of Substitution and the Relative Share of a Factor. *The Review of Economic Studies*, 1(1):72–78.
- Kalt, J. P. (1978). Technological Change and Factor Substitution in the United States: 1929-1967. *International Economic Review*, 19(3):761–775.
- Klump, R. (2001). Trade, Money and Employment in Intertemporal Optimizing Models of Growth. *The Journal of International Trade and Economic Development*, 10(4):411–428.
- Klump, R. and de La Grandville, O. (2000). Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions. *American Economic Review*, 90(1):282–291.
- Klump, R., McAdam, P., and Willman, A. (2007a). Factor Substitution and Factor-Augmenting Technical Progress in the United States: A Normalized Supply-Side System Approach. *The Review of Economics and Statistics*, 89(1):183–192.
- Klump, R., McAdam, P., and Willman, A. (2007b). The Long-Term SucCESs of the Neoclassical Growth Model. *Oxford Review of Economic Policy*, 23(1):94–114.
- Klump, R., McAdam, P., and Willman, A. (2008). Unwrapping Some Euro Area Growth Puzzles: Factor Substitution, Productivity and Unemployment. *Journal of Macroeconomics*, 30(2):645–666.
- Klump, R., McAdam, P., and Willman, A. (2012). The Normalized CES Production Function: Theory and Empirics. *Journal of Economic Surveys*, 26(5):769–799.
- Klump, R. and Preissler, H. (2000). CES Production Functions and Economic Growth. *The Scandinavian Journal of Economics*, 102(1):41–56.
- Klump, R. and Saam, M. (2008). Calibration of Normalised CES Production Functions in Dynamic Models. *Economics Letters*, 99(2):256–259.
- Kmenta, J. (1967). On Estimation of the CES Production Function. *International Economic Review*, 8(2):180–189.
- Knoblauch, M., Roessler, M., and Zwerschke, P. (2016). The Elasticity of Factor Substitution Between Capital and Labor in the U.S. Economy: A Meta-Regression Analysis. *CEPIE Working Paper Series 03/16*.

- Kumar, T. K. and Gapinski, J. H. (1974). Nonlinear Estimation of the CES Production Parameters: A Monte Carlo Study. *The Review of Economics and Statistics*, 56(4):563–567.
- Lagos, R. (2006). A Model of TFP. *The Review of Economic Studies*, 73(4):983–1007.
- Lerner, A. (1933). II.—The Diagrammatical Representation. *The Review of Economic Studies*, 1(1):68–71.
- Levhari, D. (1968). A Note on Houthakker’s Aggregate Production Function in a Multifirm Industry. *Econometrica*, 36(1):151–154.
- Léon-Ledesma, M. A., McAdam, P., and Willman, A. (2010). Identifying the Elasticity of Substitution with Biased Technical Change. *American Economic Review*, 100(4):1330–1357.
- Léon-Ledesma, M. A., McAdam, P., and Willman, A. (2015). Production Technology Estimates and Balanced Growth. *Oxford Bulletin of Economics and Statistics*, 77(1):40–65.
- Léon-Ledesma, M. A. and Satchi, M. (forthcoming). Appropriate Technology and Balanced Growth. *The Review of Economic Studies*.
- Lotka, A. J. (1926). The Frequency Distribution of Scientific Productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323.
- Luttmer, E. G. J. (2007). Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics*, 122(3):1103–1144.
- Maddala, G. S. and Kadane, J. B. (1966). Some Notes on the Estimation of the Constant Elasticity of Substitution Production Function. *The Review of Economics and Statistics*, 48(3):340–344.
- Maddala, G. S. and Kadane, J. B. (1967). Estimation of Returns to Scale and the Elasticity of Substitution. *Econometrica*, 35(3/4):419–423.
- Maki, D. R. and Meredith, L. N. (1987). A Note on Unionization and the Elasticity of Substitution. *The Canadian Journal of Economics / Revue canadienne d’Econometrie*, 20(4):792–801.
- Mallick, D. (2012). The Role of the Elasticity of Substitution in Economic Growth: A Cross-Country Investigation. *Labour Economics*, 19(5):682–694.
- Masanjala, W. H. and Papageorgiou, C. (2004). The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity. *Journal of Applied Econometrics*, 19(2):171–201.
- Matveenko, A. (2011). Stochastic Models of the Technological Ideas Flow as a Foundation of the Production Function. *International Days of Statistics and Economics*, pages 368–380.

- Matveenko, A. and Matveenko, V. (2015). Technological Choice Generating Normalized CES Production Functions. In *Conference on Dynamics, Economic Growth, and International Trade (DEGIT)*, volume 20.
- Matveenko, V. (2010). Anatomy of Production Functions: A Technological Menu and a Choice of the Best Technology. *Economics Bulletin*, 30(3):1906–1913.
- Matveenko, V. (2013). Differential Bargaining Games as Microfoundations for Production Function. *Contributions to Game Theory and Management*, 6:289–300.
- McAdam, P. (2016). de La Grandville, Olivier: Economic Growth: A Unified Approach. *Journal of Economics*, 119(1):91–96.
- McManus, W. S. (1988). Aggregation and "The" Elasticity of Substitution. *The American Economist*, 32(2):41–44.
- Miyagiwa, K. (2008). Comment on the Endogenous Aggregate Elasticity of Substitution for a Small Open Economy. *Journal of Macroeconomics*, 30(2):641–644.
- Miyagiwa, K. and Papageorgiou, C. (2007). Endogenous Aggregate Elasticity of Substitution. *Journal of Economic Dynamics and Control*, 31(9):2899–2919.
- Morawetz, D. (1976). Elasticities of Substitution in Industry: What do we learn from Econometric Estimates? *World Development*, 4(1):11–15.
- Nakamura, H. (2009). Micro-Foundation for a Constant Elasticity of Substitution Production Function through Mechanization. *Journal of Macroeconomics*, 31(3):464–472.
- Nakamura, H. (2010). Factor Substitution, Mechanization, and Economic Growth. *The Japanese Economic Review*, 61(2):266–281.
- Nakamura, H. and Nakamura, M. (2008). Constant-Elasticity-of-Substitution Production Function. *Macroeconomic Dynamics*, 12(5):694–701.
- Nakamura, Y. (2015). Productivity Versus Elasticity: A Normalized Constant Elasticity of Substitution Production Function Applied to Historical Soviet Data. *Applied Economics*, 47(53):5805–5823.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Science + Business Media, New York, 2nd edition.
- Newman, M. E. J. (2005). Power Laws, Pareto Distributions and Zipf's Law. *Contemporary Physics*, 46(5):323–351.
- Ofer, G. (1987). Soviet Economic Growth: 1928-1985. *Journal of Economic Literature*, 25(4):1767–1833.
- Osumi, Y. (2015). Aggregate Elasticity of Substitution and Structural Change. In *Studies in Medium-Run Macroeconomics Growth, Fluctuations, Unemployment, Inequality and Policies*, chapter 7, pages 169–186. World Scientific Publishing Co. Pte. Ltd.

- Palivos, T. (2008). Comment on “ σ : The Long and Short of it”. *Journal of Macroeconomics*, 30(2):687–690.
- Phelps, E. S. (1963). Substitution, Fixed Proportions, Growth and Distribution. *International Economic Review*, 4(3):265–288.
- Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, MA, London.
- Piketty, T. and Saez, E. (2003). Income Inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, 118(1):1–41.
- Pitchford, J. D. (1960). Growth and the Elasticity of Factor Substitution. *Economic Record*, 36(76):491–504.
- Reed, W. J. (2001). The Pareto, Zipf and other Power Laws. *Economics Letters*, 74(1):15–19.
- Robinson, J. (1933). *The Economics of Imperfect Competition*. London: Macmillan, 1st edition.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2):71–102.
- Rosen, S. (1978). Substitution and Division of Labor. *Economica*, 45(179):235–250.
- Saam, M. (2008). Openness to Trade as a Determinant of the Macroeconomic Elasticity of Substitution. *Journal of Macroeconomics*, 30(2):691–702.
- Schefold, B. (2008). C.E.S. Production Functions in the Light of the Cambridge Critique. *Journal of Macroeconomics*, 30(2):783–797.
- Shahiri, H. I. and Osman, Z. (2017). Substitutability of Capital-Labour in the Presence of Unions in the US Postal Services Industry. *Institutions and Economies*, 8(2):55–73.
- Silverberg, G. and Verspagen, B. (2007). The Size Distribution of Innovations Revisited: An Application of Extreme Value Statistics to Citation and Value Measures of Patent Significance. *Journal of Econometrics*, 139(2):318–339.
- Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1):65–94.
- Solow, R. M. (1962). Substitution and Fixed Proportions in the Theory of Capital. *The Review of Economic Studies*, 29(3):207–218.
- Solow, R. M. (1964). *The Behavior of Income Shares: Selected Theoretical and Empirical Issues*, chapter Capital, Labor, and Income in Manufacturing, pages 101–142. Princeton University Press.
- Solow, R. M. (1967). Some Recent Developments in the Theory of Production. In *The Theory and Empirical Analysis of Production*, pages 25–53. NBER.

- Stern, D. I. (2011). Elasticities of Substitution and Complementarity. *Journal of Productivity Analysis*, 36(1):79–89.
- Stewart, K. G. and Li, J. (2018). Are factor biases and substitution identifiable? The Canadian evidence. *Canadian Journal of Economics*, 51(2):528–548.
- Sue Wing, I. (2006). Representing Induced Technological Change in Models for Climate Policy Analysis. *Energy Economics*, 28(5-6):539–562.
- Sveikauskas, L. (1974). Bias in Cross-Section Estimates of the Elasticity of Substitution. *International Economic Review*, 15(2):522–528.
- Takayama, A. (1974). On Biased Technological Progress. *The American Economic Review*, 64(4):631–639.
- Thursby, J. G. (1980). Alternative CES Estimation Techniques. *The Review of Economics and Statistics*, 62(2):295–299.
- Thursby, J. G. and Lovell, C. A. K. (1978). An Investigation of the Kmenta Approximation to the CES Function. *International Economic Review*, 19(2):363–377.
- Toda, A. A. and Walsh, K. (2015). The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy*, 123(5):1177–1200.
- Turnovsky, S. J. (2002). Intertemporal and Intratemporal Substitution, and the Speed of Convergence in the Neoclassical Growth Model. *Journal of Economic Dynamics and Control*, 26(9-10):1765–1785.
- Turnovsky, S. J. (2008). The Role of Factor Substitution in the Theory of Economic Growth and Income Distribution: Two Examples. *Journal of Macroeconomics*, 30(2):604–629.
- Ventura, J. (1997). Growth and Interdependence. *The Quarterly Journal of Economics*, 112(1):57–84.
- Weitzman, M. L. (1970). Soviet Postwar Economic Growth and Capital-Labor Substitution. *The American Economic Review*, 60(4):676–692.
- Xue, J. and Yip, C. K. (2013). Aggregate Elasticity of Substitution and Economic Growth: A Synthesis. *Journal of Macroeconomics*, 38(Part A):60–75.
- Young, A. T. (2013). US Elasticities of Substitution and Factor-Augmentation at the Industry Level. *Macroeconomic Dynamics*, 17(4):861–897.
- Yuhn, K.-h. (1991). Economic Growth, Technical Change Biases, and the Elasticity of Substitution: A Test of the De La Grandville Hypothesis. *The Review of Economics and Statistics*, 73(2):340–346.
- Zeira, J. (1998). Workers, Machines, and Economic Growth. *The Quarterly Journal of Economics*, 113(4):1091–1117.
- Zeira, J. (2008). Machines as Engines of Growth. *Unpublished manuscript*.