Measuring Capital-Labor Substitution: The Importance of Method Choices and Publication Bias*

Sebastian Gechert[a], Tomas Havranek[b], Zuzana Irsova[b], Dominika Kolcunova[b,c]

[a]Macroeconomic Policy Institute, Düsseldorf
[b]Charles University, Prague
[c]Czech National Bank

December 17, 2020

Abstract

We show that the large elasticity of substitution between capital and labor estimated in the literature on average, 0.9, can be explained by three issues: publication bias, use of cross-country variation, and omission of the first-order condition for capital. The mean elasticity conditional on the absence of these issues is 0.3. To obtain this result, we collect 3,186 estimates of the elasticity reported in 121 studies, codify 71 variables that reflect the context in which researchers produce their estimates, and address model uncertainty by Bayesian and frequentist model averaging. We employ nonlinear techniques to correct for publication bias, which is responsible for at least half of the overall reduction in the mean elasticity from 0.9 to 0.3. Our findings also suggest that a failure to normalize the production function leads to a substantial upward bias in the estimated elasticity. The weight of evidence accumulated in the empirical literature emphatically rejects the Cobb-Douglas specification.

Keywords: Elasticity of substitution, capital, labor, publication bias, model uncertainty

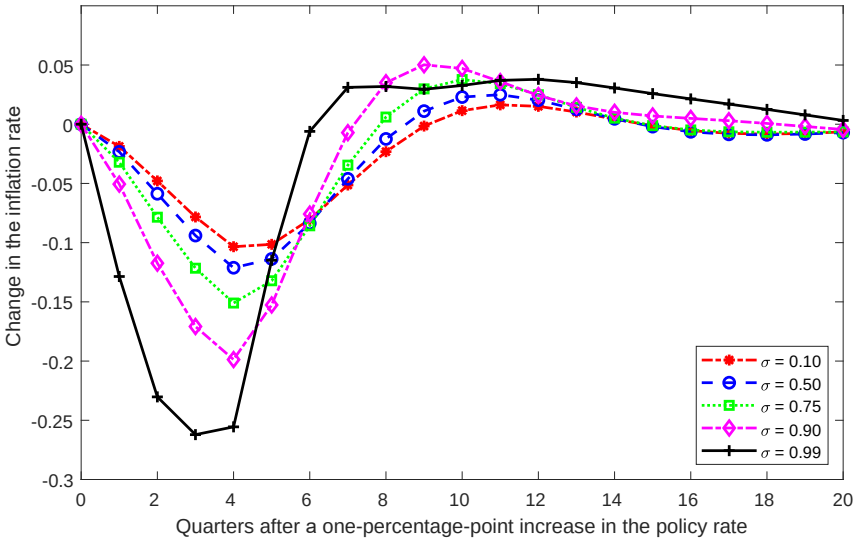
JEL Codes: D24, E23, O14

1 Introduction

A key parameter in economics is the elasticity of substitution between capital and labor. Among other things, the size of the elasticity has practical consequences for monetary policy, as Figure 1 illustrates. In the SIGMA model used by the Federal Reserve Board, the effectiveness of interest rate changes in steering inflation doubles when one assumes the elasticity to equal 0.9 instead of 0.5, yielding very different policy implications. We choose the SIGMA model for the illustration

*An online appendix with data and code is available at `meta-analysis.cz/sigma`. Corresponding author: Tomas Havranek, `tomas.havranek@ies-prague.org`. Kolcunova acknowledges support from the Czech Science Foundation (grant #21-09231S) and Charles University (project Primus/17/HUM/16). Havranek and Irsova acknowledge support from the Czech Science Foundation (grant #19-26812X).

Figure 1: The elasticity of substitution matters for monetary policy



Notes: The figure shows simulated impulse responses of inflation to a monetary policy shock. We use a calibrated version of the SIGMA model of Erceg *et al.* (2008) developed for the Federal Reserve Board and vary the value of the capital-labor substitution elasticity while leaving other parameters at their original values. The model does not have a stable solution for σ larger than one.

because, as one of very few models employed by central banks, it actually allows for different values of the elasticity of substitution. Almost all models use the convenient simplification of the Cobb-Douglas production function, which implicitly assumes that the elasticity equals one. If the true elasticity is smaller, these models overstate the strength of monetary policy and should imply a more aggressive campaign of interest rate cuts in response to a recession (Chirinko & Mallick, 2017, make a related argument).[1] In this paper we show that the Cobb-Douglas specification is at odds with the empirical evidence on the elasticity.

Aside from convenience, the other reason for the widespread use of the Cobb-Douglas production function is that, at first sight, empirical investigations into the value of the elasticity have produced many central estimates close to 1. When each study gets the same weight, the mean elasticity reported in the literature reaches 0.9—at least based on our attempt to collect all published estimates, in total 3,186 coefficients from 121 studies. But we show that the picture is seriously distorted by publication bias. After correcting for the bias, the mean reported elasticity shrinks to 0.5. This correction alone can imply halving the effectiveness of monetary policy in a structural model, as shown by Figure 1.

[1]The SIGMA model calibrates the elasticity at 0.5, which is consistent with our estimate of the mean elasticity reported in the entire literature and corrected for publication bias. While our preferred estimate (based on preferred methodology) is even lower at 0.3, such a decrease would make little difference for the SIGMA model, as illustrated in Figure 1. Further decreases below 0.5 have negligible impact on the modeled transmission of monetary policy. Note also that we use a calibrated version of the SIGMA model from the Macroeconomic Model Data Base. If some parameters of the model were estimated, the effect of changes in the elasticity would likely be attenuated. In general the modeled effect of an interest rate hike on inflation is small when the elasticity of substitution is 0.5 (a decrease of about 0.12 in the inflation rate after an increase in the policy rate of 100 basis points), but the relatively small effect of interest rate changes on prices is consistent with recent surveys of empirical evidence from VARs by Rusnak *et al.* (2013), Ramey (2016), and Cochrane (2018).

The finding of strong publication bias predominates in our results. The bias arises when different estimates have a different probability of being reported depending on sign and statistical significance. The identification builds on the fact that almost all econometric techniques used to estimate the elasticity assume that the ratio of the estimate to its standard error has a symmetrical distribution, typically a t-distribution. So the estimates and standard errors should represent independent quantities. But if statistically significant positive estimates are preferentially selected for publication, large standard errors (given by noise in data or imprecision in estimation) will become associated with large estimates. Because researchers command plenty of degrees of freedom in estimation design, a large estimate of the elasticity always emerges if the researcher looks for it long enough, and an upward bias in the literature arises. A useful analogy appears in McCloskey & Ziliak (2019), who liken publication bias to the Lombard effect in biology: speakers increase their effort in the presence of noise. Apart from linear techniques based on the Lombard effect, we employ recently developed methods by Ioannidis *et al.* (2017), Andrews & Kasy (2019), Bom & Rachinger (2019), and Furukawa (2019), which account for the potential nonlinearity between the standard error and selection effort. We also use methods by Gerber & Malhotra (2008) and van Aert & van Assen (2020) that address potential endogeneity of the standard error.[2]

The studies in our dataset do not estimate a single population parameter; rather, the precise interpretation of the elasticity differs depending on the context in which authors derive their results. We collect 71 variables that reflect the different contexts and find that our conclusions regarding publication bias hold when we control for context. Because of the richness of the literature on the elasticity of substitution, we face substantial model uncertainty with many controls and address it by using Bayesian (Eicher *et al.*, 2011; Steel, 2020) and frequentist (Hansen, 2007; Amini & Parmeter, 2012) model averaging. We investigate how the estimated elasticities depend on publication bias and the data and methods used in the analysis. Our results suggest that three factors drive the heterogeneity in the literature: publication bias (the size of the standard error), source of variation in input data (cross-country vs. industry-level variation), and identification approach (whether or not information from the first-order condition for capital is accounted for). Estimations using systems of equations tend to deliver results similar to those of single-equation approaches focused on the first-order condition for capital. In addition, the normalization of the production function used in recent studies typically brings much smaller reported elasticities, by 0.3 on average. We also find that different assumptions regarding technical change have little systematic effect on the reported elasticity.

As the bottom line of our analysis, we construct a synthetic study that uses all the estimates reported in the literature but assigns more weight to those that are arguably better specified. The result represents a mean estimate implied by the literature but conditional on the absence of publication bias, use of best-practice methodology, and other aspects related to quality (such

[2]Publication bias in economics has also been recently discussed, among others, by Havranek (2015), Brodeur *et al.* (2016), Bruns & Ioannidis (2016), Havranek & Irsova (2017), Havranek *et al.* (2017), Christensen & Miguel (2018), Astakhov *et al.* (2019), Bajzik *et al.* (2020), Blanco-Perez & Brodeur (2020), Brodeur *et al.* (2020), Cazachevici *et al.* (2020), Matousek *et al.* (2021), and Imai *et al.* (2021).

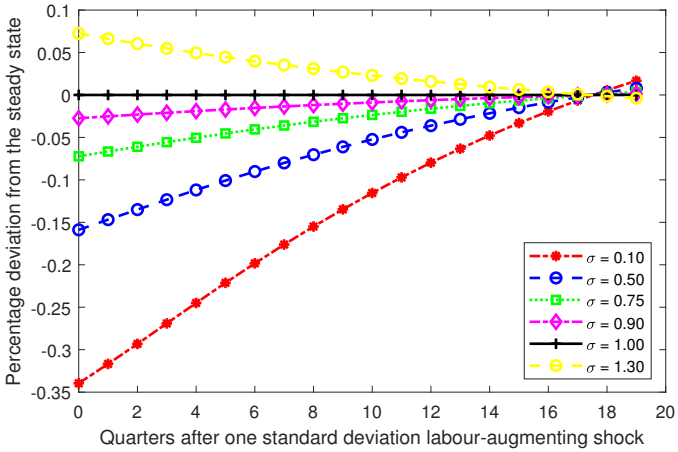
as publication in a leading journal or a large number of citations). In this way we obtain an elasticity of 0.3 with an upper bound of the 95% confidence interval at 0.6. Though certainly not the definitive point estimate for the elasticity, it is the best guess we can make when looking at half a century of accumulated empirical evidence.

Defining best-practice methodology is subjective, and different authors will have different preferences on study design. But to arrive at 0.3, it is enough to hold two preferences: i) using variation across industries is superior to using variation across countries (which is substantiated, e.g., by Nerlove, 1967; Chirinko, 2008) and ii) including information from the first-order condition for capital is superior to ignoring it (and, for example, focusing exclusively on the first-order condition for labor). To put these numbers into perspective, we once again turn to the Fed's SIGMA model, which employs a value of 0.5 for the elasticity of substitution (Erceg *et al.*, 2008). This calibration corresponds to the mean estimate in the literature corrected for publication bias, without discounting any estimates based on data and methodology. The model employed by the Bank of Finland (Kilponen *et al.*, 2016), on the other hand, uses the elasticity of 0.85, which is close to the mean estimate in the literature without correction for publication bias. The calibration closest to our final result is that of Cantore *et al.* (2015), who use a prior of 0.4. Their posterior estimate is even lower, though, at below 0.2.

The elasticity of substitution between capital and labor is central to a host of problems aside from monetary policy. Our understanding of long-run growth depends on the value of the elasticity (Solow, 1956). The sustainability of growth in the absence of technological change is contingent on whether or not the elasticity of substitution exceeds one (Antras, 2004). Klump & de La Grandville (2000) suggest that a larger elasticity in a country results in higher per capita income. Turnovsky (2002) argues that a smaller elasticity leads to faster convergence. Nekarda & Ramey (2013) argue that the countercyclicality of the price markup over marginal cost also depends on the elasticity of substitution. The elasticity represents an important parameter in analyzing the effects of fiscal policies, including the effect of corporate taxation on capital formation, and in determining optimal taxation of capital (Chirinko, 2002).

But perhaps most prominently, the elasticity of substitution is a key parameter in the literature on the labor share. The evidence of a declining labor share has in fact revived general interest in estimating the elasticity because some of the explanations depend critically on the value of the elasticity (σ). Oberfield & Raval (2014) categorize these explanations into two groups: (1) mechanisms decreasing the labor share via changing factor prices and (2) mechanisms decreasing the labor share via changing technology. Regarding group (1), the explanations put forward by Piketty (2014) and Karabarbounis & Neiman (2014) hold only when the elasticity surpasses one. Then the global decline in the labor share can be attributed to an increasing capital-labor ratio, either via capital deepening (Piketty, 2014) or as a response to falling investment prices (Karabarbounis & Neiman, 2014). With $\sigma < 1$, however, declining prices of capital and increased capital accumulation raise the labor share. Yet, as we show in this paper, $\sigma < 1$ is consistent with the bulk of the empirical estimates of the elasticity. In this context, Glover & Short (2020a) assert that capital deepening cannot explain the observed

Figure 2: The elasticity of substitution matters for the labor share



Notes: The figure shows simulated impulse responses of the labor share to a labor-augmenting technology shock. We use the model developed by Cantore *et al.* (2014) and Cantore *et al.* (2015).

decline; they point to biases that led to the high elasticity estimates of Karabarbounis & Neiman (2014). Regarding group (2), alternative explanations stress changes in automation, offshoring, directed technological change (as in Oberfield & Raval, 2014; Eden & Gaggl, 2015; Koh *et al.*, 2016), a slowdown in labor productivity (as in Grossman *et al.*, 2017), a rise in concentration (Autor *et al.*, 2017), and demographic changes (Glover & Short, 2020b); explanations that do not hinge on high values of σ.

The elasticity also has important effects on the short-run dynamics of the labor share. This channel can be illustrated by computing the response of the labor share to a labor-augmenting technology shock, as we do in Figure 2 based on the model developed by Cantore *et al.* (2014) and Cantore *et al.* (2015). In the case of the Cobb-Douglas production function the labor share remains constant, while with $\sigma < 1$ the share decreases after a labor-augmenting shock. As the figure illustrates, the response is highly sensitive to changes in σ. A model with a lower elasticity, consistent with our results, is able to match the actual dynamics of the data on the labor share better than the Cobb-Douglas case (Cantore *et al.*, 2015).

The remainder of the paper is structured as follows: Section 2 briefly discusses how the elasticity of substitution is estimated; Section 3 describes how we collect estimates of the elasticity from primary studies and provides a bird's-eye view of the data; Section 4 examines publication bias; Section 5 investigates the drivers of heterogeneity in the reported elasticities and calculates the mean elasticity implied by best practice in the literature; Section 6 concludes the paper. Appendix A illustrates the working of publication bias and basic meta-analysis tools via a Monte Carlo simulation. Appendix B and Appendix C describe the bias-correction techniques designed by Furukawa (2019) and Andrews & Kasy (2019). Appendix D shows summary statistics of the variables that reflect study context, Appendix E presents robustness checks, and Appendix F includes the list of studies from which we extract estimates. The data and code are available in an online appendix at `meta-analysis.cz/sigma`.

2 Estimating the Elasticity

To set the stage for data collection and identification of factors driving heterogeneity in results, we provide a short description of the most common approaches to estimating the elasticity of substitution between capital and labor. The concept was introduced by Hicks (1932) and almost simultaneously and independently by Robinson (1933), whose more popular definition treats the elasticity as a percentage change of the ratio of two production factors divided by the percentage change of the ratio of their marginal products. Under perfect competition, both inputs are paid their marginal products, so the elasticity of substitution can be written as

$$\sigma = \frac{d(K/L)/(K/L)}{d(w/r)/(w/r)} = -\frac{d\,\log(K/L)}{d\,\log(r/w)}, \tag{1}$$

where K and L denote capital and labor, r is the rental price of capital, and w is the wage rate. Under a quasiconcave production function the elasticity attains any number in the interval $(0, \infty)$. If $\sigma = 0$, capital and labor are perfect complements, always used in a fixed proportion in the Leontief production function. If the elasticity lies in the interval $(0, 1)$, capital and labor form gross complements. If $\sigma = 1$, the production function becomes Cobb-Douglas, and the relative change in quantity becomes exactly proportional to the relative change in prices. If the elasticity lies in the interval $(1, \infty)$, capital and labor form gross substitutes.

Although the concept of the elasticity of substitution was introduced in the 1930s, empirical estimates were only enabled by an innovation that came more than 20 years later: the introduction of the constant elasticity of substitution (CES) production function by Solow (1956), later popularized by Arrow *et al.* (1961). The CES production function can be written as

$$Y_t = C[\pi(A_t^K K_t)^{\frac{\sigma-1}{\sigma}} + (1-\pi)(A_t^L L_t)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where σ denotes the elasticity of substitution, K and L are capital and labor, C is an efficiency parameter, and π is a distributional parameter. The fraction $\frac{\sigma-1}{\sigma}$ is often labeled as ρ, a transformation of the elasticity called the substitution parameter. A_t^K and A_t^L denote the level of efficiency of the respective inputs, and variations in A_t^K and A_t^L over time reflect capital- and labor-augmenting technological change. When $A_t^K = A_t^L = A_t$, technological change becomes Hicks-neutral, which means that the marginal rate of substitution does not change when an innovation occurs.

The CES production function is nonlinear in parameters, and in contrast to the Cobb-Douglas case, a simple analytical linearization does not emerge. Thus the CES production function can be estimated (i) in its nonlinear form, (ii) in a linearized form as suggested by Kmenta (1967), or (iii) by using first-order conditions (FOCs). Kmenta (1967) introduced a logarithmized version of Equation 2 with Hicks-neutral technological change:

$$\log Y_t = \log C + \frac{\sigma}{\sigma-1} \log\left[\pi K_t^{\frac{\sigma-1}{\sigma}} + (1-\pi) L_t^{\frac{\sigma-1}{\sigma}}\right] \tag{3}$$

and then applied a second-order Taylor series expansion to the term $\log[\cdot]$ around the point $\sigma = 1$ to arrive at a function linear in σ:

$$\log Y_t = \log C + \pi \log K_t + (1-\pi)\log L_t - \frac{(\sigma-1)\pi(1-\pi)}{2\sigma}(\log K_t - \log L_t)^2. \tag{4}$$

Estimation of σ via first-order conditions was first suggested by Arrow *et al.* (1961). The underlying assumptions involve constant returns to scale and fully competitive factor and product markets. The FOC with respect to capital can be written as follows:

$$\log\left(\frac{Y_t}{K_t}\right) = \sigma \log\left(\frac{1}{\pi}\right) + (1-\sigma)\log(A_t^K C) + \sigma \log\left(\frac{r_t}{p_t}\right). \tag{5}$$

Consequently, the FOC with respect to labor implies

$$\log\left(\frac{Y_t}{L_t}\right) = \sigma \log\left(\frac{1}{1-\pi}\right) + (1-\sigma)\log(A_t^L C) + \sigma \log\left(\frac{w_t}{p_t}\right), \tag{6}$$

where p is the price of the output. Both conditions can be combined to yield

$$\log\left(\frac{K_t}{L_t}\right) = \sigma \log\left(\frac{\pi}{1-\pi}\right) + (\sigma-1)\log\left(\frac{A_t^K}{A_t^L}\right) + \sigma \log\left(\frac{w_t}{r_t}\right). \tag{7}$$

In a similar way, one can derive FOCs with respect to the labor share $(wL)/Y$, capital share $(rK)/Y$, or their reversed counterparts. The FOCs can be estimated separately as single equations, within a system of two or three FOCs, and as a system of FOCs coupled with a nonlinear or linearized CES production function. The latter approach (also called a supply-side system approach) has become especially popular in recent studies. León-Ledesma *et al.* (2010) assert that using the supply-side system approach dominates one-equation estimation, especially when coupled with cross-equation restrictions and normalization, which was suggested by de La Grandville (1989) and Klump & de La Grandville (2000). After scaling technological progress so that $A_0^K = A_0^L = 1$, the normalized production function can be written as

$$Y_t = Y_0 \left[\pi_0 \left(\frac{A_t^K K_t}{K_0}\right)^{\frac{\sigma-1}{\sigma}} + (1-\pi_0)\left(\frac{A_t^L L_t}{L_0}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{8}$$

where $\pi_0 = r_0 K_0/(r_0 K_0 + w_0 L_0)$ denotes the capital income share evaluated at the point of normalization. The point of normalization can be defined, for instance, in terms of sample means. In other words, normalization means rewriting the production function in an indexed number form (Klump *et al.*, 2012).

Though the aforementioned approaches to estimating the elasticity dominate the literature, we also consider other approaches, in particular the translog production function. The translog function is quadratic in the logarithms of inputs and outputs and provides the second-order

approximation to any production frontier (omitting now subscript t for ease of exposition):

$$\log Y = \log \alpha_0 + \sum_i \alpha_i \log X_i + \frac{1}{2} \sum_i \sum_j \alpha_{ij} \log X_i \log X_j, \tag{9}$$

where α_0 denotes the state of technological knowledge, and X_i and X_j are inputs, in our case capital and labor. The translog production frontier provides a wider set of options for substitution and transformation patterns than a frontier based on the CES production function. Due to the duality principle, researchers often employ the translog cost function instead:

$$\log C = \alpha_0 + \theta_1 \log Y + \frac{1}{2}\theta_2 (\log Y)^2 + \sum_i \beta_i \log P_i + \frac{1}{2} \sum_i \sum_j \epsilon_{ij} \log P_i \log P_j + \sum_i \delta_i \log P_i \log Y, \tag{10}$$

where C denotes total costs, $i = K, L$, and P_i is input factor price (that is, w and r). Using Sheppard's lemma, the following cost share functions can be derived:

$$S_i = \beta_i + \sum_i \epsilon_{ij} \log P_j + \delta_i \log Y, \tag{11}$$

where S_i denotes the share of the i-th factor in total costs. In this case, Allen partial elasticities of substitution are most often estimated and are defined as

$$\sigma_{ij} = \frac{\gamma_{ij} + S_i S_j}{S_i S_j}. \tag{12}$$

We include estimates from all of the aforementioned specifications, as each of them provides a measure of the elasticity of substitution between capital and labor, broadly defined. Then we control for the various aspects of the context in which researchers obtain their estimates. These aspects are presented and discussed in detail later in Section 5, while the following section describes the dataset of the estimated elasticities.

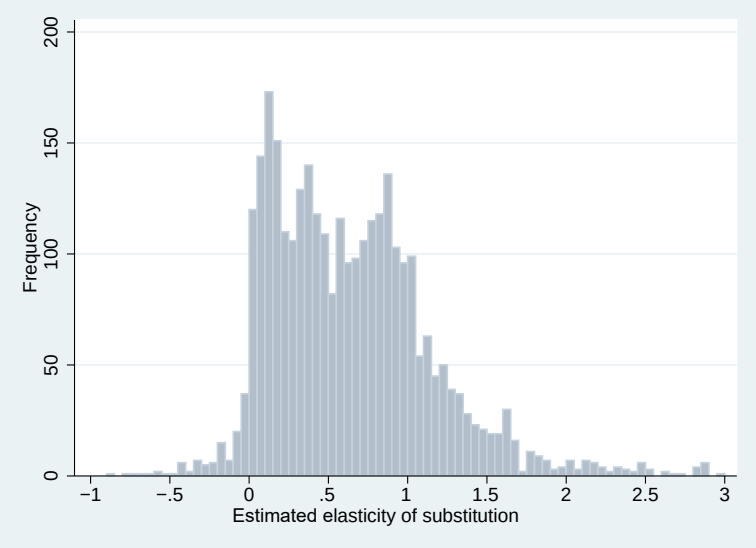
3 Data

We use Google Scholar to search for studies estimating the elasticity. Google's algorithm goes through the full text of studies, thus increasing the coverage of suitable published estimates, irrespective of the precise formulation of the study's title, abstract, and keywords. Our search query, available in the online appendix, is calibrated so that it yields the best-known relevant studies among the first hits. We examine the first 500 papers returned by the search. In addition, we inspect the lists of references in these studies and their Google Scholar citations to check whether we can find usable studies not captured by our baseline search—a method called "snowballing" in the literature on research synthesis. We terminate the search on August 1, 2018, and do not add any new studies beyond that date.

To be included in our dataset, a study must satisfy three criteria. First, at least one

estimate in the study must be directly comparable with the estimates described in Section 2. Second, the study must be published. This criterion is mostly due to feasibility since even after restricting our attention to published studies the dataset involves a manual collection of hundreds of thousands of data points. Moreover, we expect published studies to exhibit higher quality on average and to contain fewer typos and mistakes in reporting their results. Note that the inclusion of unpublished papers is unlikely to alleviate publication bias (Rusnak *et al.*, 2013): researchers write their papers with the intention to publish.[3] Third, the study must report standard errors or other statistics from which the standard error can be computed. If the elasticity is not reported directly, but can be derived from the presented results, we use the delta method to approximate the standard error. Omitting the estimates with approximated standard errors does not change our results up to a second decimal place.

Figure 3: Distribution of the estimated elasticities

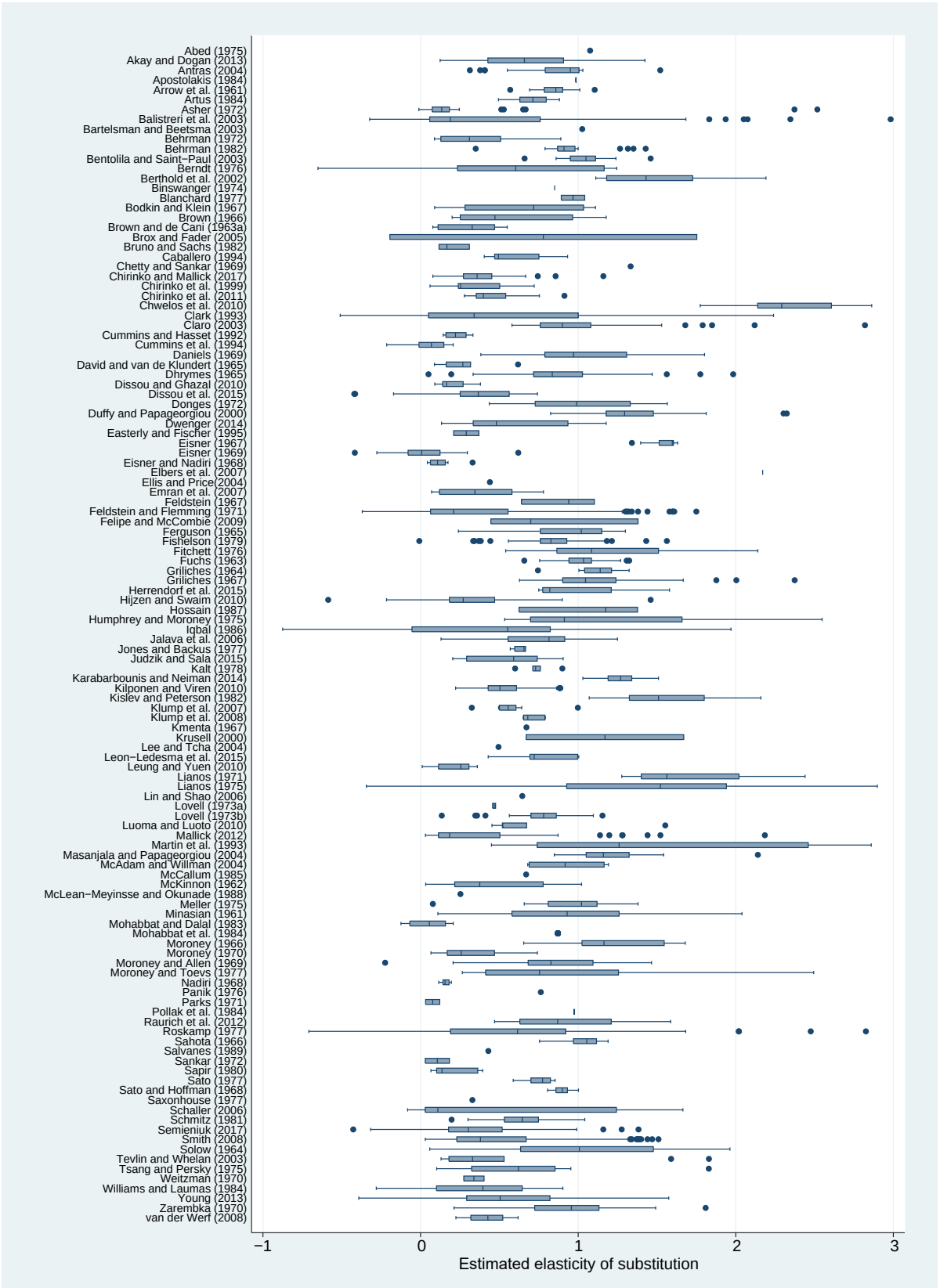


Notes: Estimates smaller than −1 and larger than 3 are excluded from the figure for ease of exposition but included in all statistical tests.

Using the search algorithm and inclusion criteria described above, we collect 3,186 estimates of the elasticity of substitution from 121 studies. To our knowledge, this makes our paper the largest meta-analysis conducted in economics so far: Doucouliagos & Stanley (2013), for example, survey dozens of meta-analyses and find that the largest one uses 1,460 estimates. Ioannidis *et al.* (2017) report that the mean number of estimates used in economics meta-analyses is 400. The literature on the elasticity of substitution is vast, with a long tradition spanning six decades and more than 100 countries. The list of the studies we include in the dataset (we call them "primary studies") is available in Appendix F. Out of the 121 studies, 19 are published in the five leading journals in economics. Altogether, they have received more than 20,000 citations in Google Scholar, highlighting the importance of the topic.

[3] A more precise label for publication bias is therefore "selective reporting," but we use the former, more common one to maintain consistency with previous studies on the topic, such as DeLong & Lang (1992), Card & Krueger (1995), and Ashenfelter & Greenstone (2004).

Figure 4: Estimates vary both across and within studies

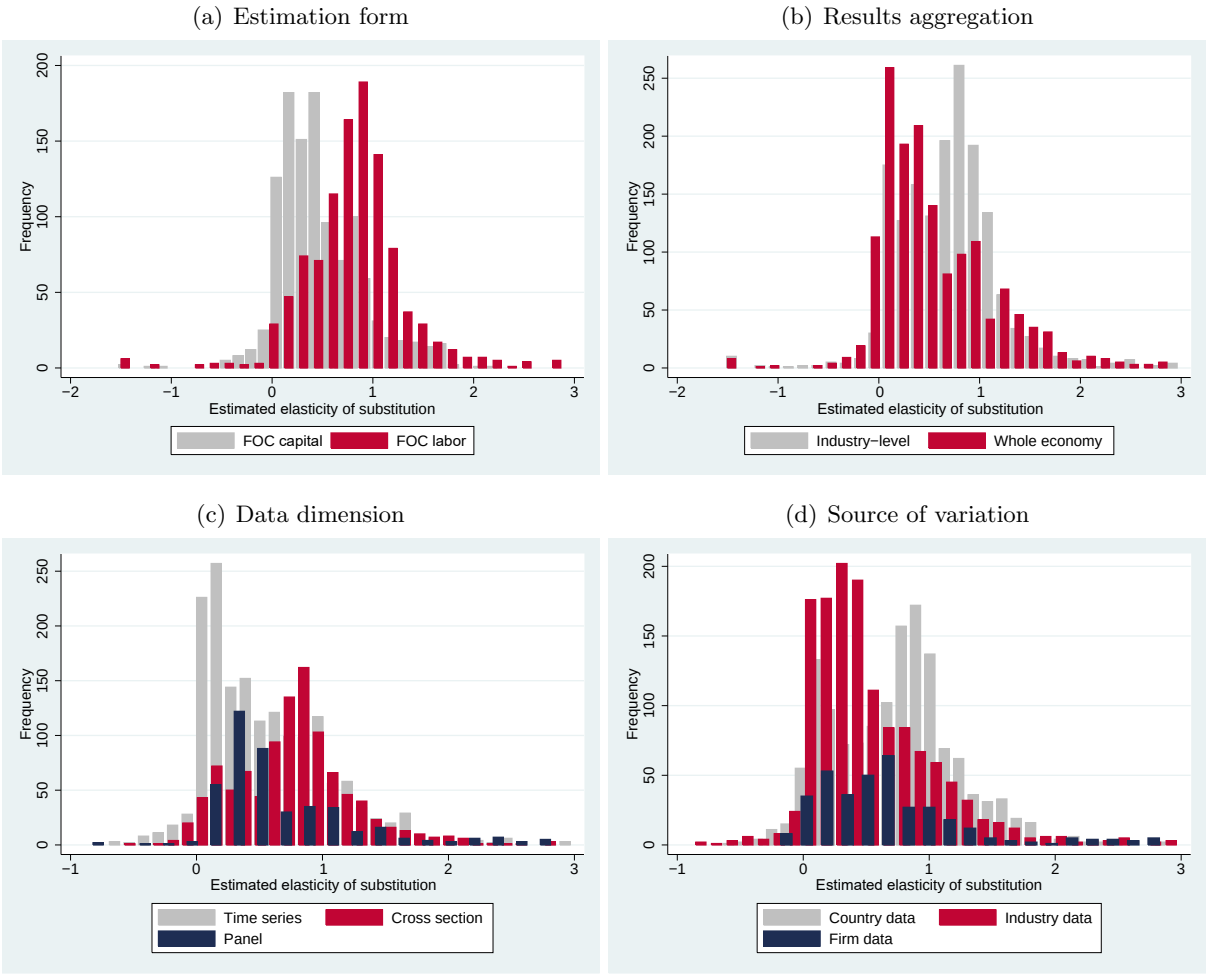


Notes: The figure shows a box plot of the estimates of the elasticity of substitution reported in individual studies. The box shows interquartile range (P25–P75) and the median highlighted. Whiskers cover (P25 − 1.5*interquartile range) to (P75 + 1.5*interquartile range). The dots are remaining (outlying) estimates. Estimates smaller than −1 and larger than 3 are excluded from the figure for ease of exposition but included in all statistical tests.

The mean reported estimate of the elasticity of substitution is 0.9 when we give the same weight to each study; that is, when we weight the estimates by the inverse of the number of observations reported per study. A simple mean of all estimates is 0.8. We consider the weighted mean to be more informative, because the simple mean is driven by studies that report many estimates, typically the results of robustness checks, and we see little reason to place more weight on such studies. For both such constructed means, in any case, the deviation from the Cobb-Douglas specification is not dramatic, and one could use the mean estimate from the literature as a justification of why the Cobb-Douglas production function presents a solid approximation of the data. We will argue that such an interpretation of the data misleads the reader because of publication bias and misspecifications in the literature.

Figure 3 shows the distribution of the estimates. Curiously, the distribution is bimodal, with peaks near 0 and slightly under 1, pointing to strong and systematic heterogeneity among the estimates. Three-quarters of the estimates lie between 0 and 1, 21% are greater than one, and only 4% attain a theoretically implausible negative value. At first sight it is apparent that

Figure 5: Prima facie patterns in the data



Notes: FOC = first-order condition. Estimates smaller than −1 and larger than 3 are excluded from the figure for ease of exposition but included in all statistical tests.

a researcher wishing to calibrate her structural model can find some empirical justification for any value of the elasticity between 0 and 1.5. There are a few extreme outliers in the data, thus we winsorize the estimates at the 5% level (our main results hold with different winsorization levels). In Figure 4 we show the box plot of the estimates. Not only do elasticities vary across studies, but also within studies. Most studies report at least some estimates close to 1, giving further (but superficial, as we will show later) credence to the Cobb-Douglas specification.

Apart from the estimates of σ and their standard errors, we collect 71 variables that capture the context in which different estimates are obtained. In consequence, we had to collect more than 220,000 data points from primary studies—a laborious but complex exercise. The data were collected by two of the coauthors of this paper, each of whom then double-checked random portions of the data collected by the other coauthor in order to minimize potential mistakes arising from manually coding so many entries. The entire process took seven months, and the final dataset is available in the online appendix. Out of the 71 variables that we collect, 50 are included in the baseline model, while the rest only appear in the subsamples of the data for which they apply.

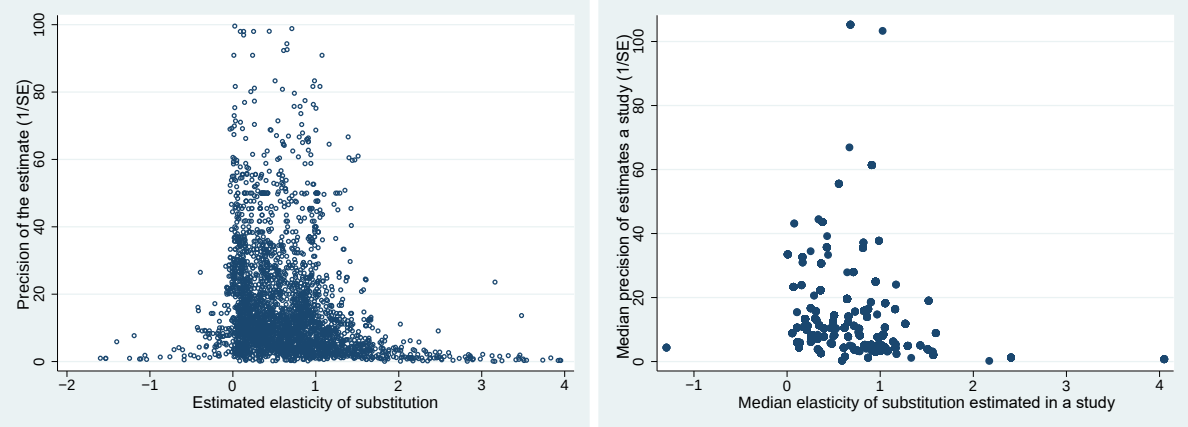
A casual look at the estimates reveals systematic differences among the reported elasticities derived from different data and identified using different methodologies. The most striking patterns are shown in Figure 5. For instance, while the mean of the estimates coming from the first-order condition for capital is 0.4, for the first-order condition for labor the mean is twice as much. The mean of the elasticities based on time series data is 0.5, while for cross-sectional data it reaches 0.8. Estimates based on industry-level data appear to be systematically smaller than those based on country-level data, and elasticities presented for individual industries are on average larger than aggregated estimates. These patterns may explain the bimodality of the overall histogram presented in Figure 3. Nevertheless, at this point we cannot be sure whether the differences are fundamental or whether they reflect correlations with other factors. A detailed analysis of heterogeneity is available in Section 5. Some of the differences among the estimates can also be attributable to publication bias, an issue to which we turn next.

4 Publication Bias

Theory and intuition provides little backing for a zero or negative elasticity of substitution between capital and labor, so it seems natural to discard such estimates. Previous researchers (most prominently, Ioannidis *et al.*, 2017) have shown that such a censoring distorts inference drawn from the literature, and here we document that publication bias is strong in the case of the elasticity of substitution. Even when the true elasticity is positive in every single estimation context, given sufficient noise in data and methods both negative and zero (statistically insignificant) estimates will appear. For each individual author who obtains such estimates, it makes little sense to focus on them; it will bring their study closer to the truth if they find and highlight a specification that yields a clearly positive elasticity. The problem is that noise in data and methods will also produce estimates that are much larger than the true effect, and such estimates are hard to identify: no upper threshold symmetrical to zero exists that would tell

the researcher the estimates are implausible. If many small imprecise estimates are discarded but many large imprecise estimates are reported, an upward bias arises on average. Thus a paradox arises: publication bias can be beneficial at the micro level of individual studies, but is detrimental at the macro level of the entire literature. Ioannidis *et al.* (2017) document that the typical exaggeration due to publication bias in economics is twofold. We find it remarkable that no study has addressed potential publication bias in the literature on the elasticity of substitution between capital and labor, one of the most important parameters in economics.

Figure 6: Negative estimates of the elasticity are underreported



Notes: In the absence of publication bias the scatter plot should resemble an inverted funnel symmetrical around the most precise estimates. The left panel shows all estimates, the right panel shows median estimates from each study. Estimates smaller than −2 and larger than 4 (together with precision values above 100 in the left panel) are excluded from the figure for ease of exposition but included in all statistical tests.

Figure 6 provides a graphical illustration of the mechanism outlined in the previous paragraph. In the scatter plot the horizontal axis measures the magnitude of the estimated elasticities, and the vertical axis measures their precision. In the absence of publication bias, the scatter plot will form an inverted funnel: the most precise estimates will lie close to the true mean elasticity, imprecise estimates will be more dispersed, and both small and large imprecise estimates will appear with the same frequency. (The scatter plot is thus typically called a funnel plot, Stanley & Doucouliagos 2010.) The figure shows the predicted funnel shape, still with plenty of heterogeneity at the top—but also shows asymmetry. For the funnel to be symmetrical, and hence consistent with the absence of publication bias, we should observe many more reported negative and zero estimates. In Appendix A we use a simple Monte Carlo simulation to further explain the mechanism of publication bias and the baseline meta-analysis estimators we use.

4.1 Baseline Methods

To identify publication bias numerically, we refer to the analogy with the Lombard effect mentioned in the Introduction: other things being equal, under publication bias authors will increase their effort (specification search) in response to noise (imprecision resulting from data or methodology). Thus publication bias is consistent with finding a correlation between estimates

of the elasticity and their standard errors. In contrast, if there is no bias, there should be no correlation, because the properties of the techniques used to obtain the elasticity ensure that the ratio of the estimate to its standard error has a t-distribution. It follows that estimates and standard errors should be statistically independent quantities. In any case, the intercept in the regression of the estimated elasticities on their standard errors can be interpreted as the mean elasticity corrected for potential publication bias (Stanley, 2005). It represents the mean elasticity conditional on the standard error approaching zero, and because in this specification publication bias forms a linearly increasing function of the standard error, the intercept measures the corrected estimate. The coefficient on the standard error measures publication bias and can be thought of as a test of the asymmetry of the funnel plot. So we have

$$\hat{\sigma}_{ij} = \sigma_0 + \gamma SE(\hat{\sigma}_{ij}) + u_{ij}, \tag{13}$$

where $\hat{\sigma}$ is the i-th estimated elasticity in study j, γ denotes the intensity of publication bias, and σ_0 represents the mean elasticity corrected for the bias.

In Table 1 we report the results of several specifications based on Equation 13. We cluster standard errors at both the study and the country level, as estimates are unlikely to be independent within these two dimensions; our implementation of two-way clustering follows Cameron *et al.* (2011). We also report wild bootstrap confidence intervals (Cameron *et al.*, 2008). In all specifications we find a statistically significant and positive coefficient on the standard error (publication bias) and a significant and positive intercept (the mean elasticity corrected for the bias). After correcting for publication bias, the mean elasticity drops from 0.9 to 0.5.

Table 1: Linear tests of funnel asymmetry suggest publication bias

	OLS	FE	BE	Precision	Study
SE	0.881***	0.656***	1.111***	0.755***	0.888***
Publication bias	(0.086)	(0.201)	(0.190)	(0.190)	(0.094)
	[0.49; 1.21]	–	–	[0.12; 1.40]	[0.62; 1.22]
Constant	0.492***	0.529***	0.499***	0.484***	0.544***
Mean beyond bias	(0.028)	(0.033)	(0.048)	(0.028)	(0.039)
	[0.38; 0.61]	–	–	[0.39; 0.66]	[0.44; 0.64]
Studies	121	121	121	121	121
Observations	3,186	3,186	3,186	3,186	3,186

Notes: The table presents the results of regression $\hat{\sigma}_{ij} = \sigma_0 + \gamma SE(\hat{\sigma}_{ij}) + u_{ij}$. $\hat{\sigma}_{ij}$ and $SE(\hat{\sigma}_{ij})$ are the i-th estimates of elasticity of substitution and their standard errors reported in the j-th study. The standard errors of the regression parameters are clustered at both the study and country level and shown in parentheses (the implementation of two-way clustering follows Cameron *et al.*, 2011). OLS = ordinary least squares. FE = study-level fixed effects. BE = study-level between effects. Precision = the inverse of the reported estimate's standard error is used as the weight. Study = the inverse of the number of estimates reported per study is used as the weight. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level. Standard errors in parentheses. Whenever possible, in square brackets we also report 95% confidence intervals from wild bootstrap clustering; implementation follows Roodman (2019), and we use Rademacher weights with 9999 replications.

The first column of Table 1 reports a simple OLS regression. The second column adds study-level fixed effects in order to account for unobserved study-specific characteristics, but little changes. (Adding country dummies would also produce similar results.) The third column

Table 2: Nonlinear techniques corroborate publication bias

	Bom & Rachinger (2019)	Furukawa (2019)	Andrews & Kasy (2019)	Ioaninidis et al. (2017)
Mean beyond bias	0.52	0.55	0.43	0.50
	(0.09)	(0.21)	(0.02)	(0.06)

Notes: Standard errors in parentheses. The method developed by Bom & Rachinger (2019) searches for a precision threshold above which publication bias is unlikely. Methods developed by Furukawa (2019) and Andrews & Kasy (2019) are described in detail in Appendix B and Appendix C. The method developed by Ioannidis *et al.* (2017) focuses on estimates with adequate power.

uses between-study variance instead of within-study variance, and the estimate of the corrected mean remains not much affected. Next, we apply two weighting schemes. First, precision becomes the weight, as suggested by Stanley & Doucouliagos (2017), which adjusts for the heteroskedasticity in the regression. Similar weights are also used in physics for meta-analyses of particle mass estimates (Baker & Jackson, 2013). The corrected mean elasticity becomes a bit smaller, but not far from 0.5. Second, we weight the data by the inverse of the number of observations reported in a study, so that each study has the same impact on the results. Again, the difference is small in comparison to other specifications.

The simple tests based on the Lombard effect and presented in Table 1 are intuitive but can themselves be biased if publication selection does not form a linear function of the standard error. For example, it might be the case that estimates are automatically reported if they cross a particular precision threshold. This is the intuition behind the estimator due to Bom & Rachinger (2019) presented in Table 2. Bom & Rachinger (2019) show how to estimate this threshold for each literature and introduce an "endogenous kink" technique that extends the linear test based on the Lombard effect. Next, Furukawa (2019) provides a nonparametric method that is robust to various assumptions regarding the functional form of publication bias and the underlying distribution of true effects. Furukawa (2019) suggests using only a portion of the most precise estimates, the stem of the funnel plot, and determines this portion by minimizing the trade-off between variance (decreasing in the number of estimates included) and bias (increasing in the number of imprecise estimates included). The stem-based method is generally more conservative than those commonly used, producing wide confidence intervals; the details are available in Appendix B.

Another nonlinear method to correct for publication bias is advocated by Andrews & Kasy (2019). They show how the conditional publication probability (the probability of publication as a function of a study's results) can be nonparametrically identified and then describe how publication bias can be corrected if the conditional publication probability is known. The underlying intuition involves jumps in publication probability at conventional p-value cut-offs. Using their method, we estimate that positive elasticities are six times more likely to be published than negative ones. We include more details on the approach and estimation in Appendix C. Finally, the remaining estimate in Table 2 arises using the approach championed by Ioannidis *et al.* (2017), who focus only on estimates with adequate statistical power. We conclude that both linear and nonlinear techniques agree that 0.5 represents a robust estimate of the mean

elasticity of substitution after correcting the literature for publication bias. Since the uncorrected mean equals 0.9, the exaggeration due to publication bias is almost twofold, consistent with the rule of thumb suggested by Ioannidis *et al.* (2017).

4.2 Extensions

Our results presented so far regarding publication bias can be criticized along three main lines. First, the distribution of elasticity estimates in some studies does not have to be symmetrical if the elasticity is not estimated directly but as a function of regression parameters from reduced-form estimations like (4). Such asymmetry in the distribution could give rise to the asymmetry of the funnel plot even in the absence of publication bias. Second, both the estimate and standard error of the elasticity can be jointly influenced by characteristics of data and methods, which would violate the exogeneity assumption and again yield an asymmetrical funnel plot even when no publication bias is present. Third, our tests of publication bias assume that researchers compare their estimates with zero. But other publication hurdles can potentially be more important: departure from the Cobb-Douglas case or other important benchmarks in the literature, such as the estimate of 1.3 by Karabarbounis & Neiman (2014) in the context of the labor share. We thank two referees of this Journal for bringing these important problems to our attention. In the remainder of this section we focus on the linear models of publication bias because they are simpler and we have shown earlier that they bring results similar to the more complex non-linear models.

Table 3: Direct estimates of the elasticity

	OLS	FE	BE	Precision	Study
SE	0.976^{***}	0.868^{***}	1.358^{***}	0.752^{*}	1.019^{***}
Publication bias	(0.167)	(0.317)	(0.271)	(0.396)	(0.132)
	[-0.23; 1.46]	–	–	[-0.61; 2.13]	[0.59; 1.35]
Constant	0.459^{***}	0.472^{***}	0.429^{***}	0.455^{***}	0.494^{***}
Mean beyond bias	(0.0226)	(0.0408)	(0.0575)	(0.0319)	(0.0354)
	[0.35; 0.57]	–	–	[0.31; 0.74]	[0.40; 0.60]
Studies	67	67	67	67	67
Observations	2,316	2,316	2,316	2,316	2,316

Notes: The table presents the results of regression $\hat{\sigma}_{ij} = \sigma_0 + \gamma SE(\hat{\sigma}_{ij}) + u_{ij}$. $\hat{\sigma}_{ij}$ and $SE(\hat{\sigma}_{ij})$ are the i-th estimates of elasticity of substitution and their standard errors reported in the j-th study. In this specification we only include direct estimates of the elasticity, i.e. the cases in which the regression parameter reported in a paper directly corresponds to the elasticity and no re-computation is needed. The standard errors of the regression parameters are clustered at both the study and country level and shown in parentheses (the implementation of two-way clustering follows Cameron *et al.*, 2011). OLS = ordinary least squares. FE = study-level fixed effects. BE = study-level between effects. Precision = the inverse of the reported estimate's standard error is used as the weight. Study = the inverse of the number of estimates reported per study is used as the weight. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level. Whenever possible, in square brackets we also report 95% confidence intervals from wild bootstrap clustering; implementation follows Roodman (2019), and we use Rademacher weights with 9999 replications.

First, we address the natural asymmetry in the estimates from some studies. Table 3 shows the results of publication bias tests when we exclude all estimates that can potentially be asymmetrically distributed. In other words, we retain only estimates for which the reported

regression coefficient can be directly interpreted as the elasticity of substitution (so that no re-computation is needed, neither by us nor by the authors of the primary studies) and at the same time the coefficient features a symmetrical distribution given by the properties of the estimation technique. Doing so restricts our sample to 2,316 estimates from 67 studies, but the results remain remarkably consistent: we find strong upward publication bias and a corrected mean elasticity of about 0.5 or slightly less. Even the most conservative technique in this case, precision weighting with wild bootstrap, gives us an upper bound of the 95% confidence interval at 0.74, safely below the Cobb-Douglas case.

Second, we address the likely endogeneity of the standard error in some studies. Table 4 presents the results of an instrumental variable (IV) regression and a new technique called p-uniform*. IV presents a crucial robustness check because in primary studies estimates and standard errors are jointly determined by the estimation technique. If some techniques produce systematically larger standard errors and point estimates, our finding of publication bias could be spurious. An intuitive instrument for the standard error is the inverse of the square root of the number of observations used in the primary study: the root is correlated with the standard error by definition but is unlikely to be much correlated with the use of a particular estimation technique. Employing IV in the first column of Table 4 we obtain a larger estimate of publication bias and a smaller estimate of the mean elasticity corrected for publication bias, 0.3, compared to our baseline estimation presented earlier.

Table 4: Relaxing the exogeneity assumption

	IV	p-uniform*
Publication bias	2.186^{***}	YES^{***}
	(0.413)	(*0.005*)
	[1.20; 3.68]	
Mean beyond bias	0.279^{***}	0.416^{**}
	(0.0702)	(*0.042*)
	[0.04; 0.47]	[0.01; 0.74]
Studies	121	121
Observations	3,186	3,186

Notes: IV = the inverse of the square root of the number of observations employed by researchers is used as an instrument for the standard error. P-uniform* = a technique developed by van Aert & van Assen (2020) and based on the distribution of p-values. For IV, standard errors are clustered at both the study and country level and reported in parentheses. For p-uniform*, p-values are reported in parentheses. For both techniques, the corresponding 95% confidence intervals are reported in square brackets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

The second column of Table 4 presents the results of p-uniform*. The technique was developed by van Aert & van Assen (2020) for standardized coefficients used in psychology, but it can also be applied to regression coefficients. At the heart of p-uniform* lies the statistical principle that p-values should be uniformly distributed *at the mean underlying effect size*: when testing the hypothesis that the estimated coefficient equals the underlying value of the effect. Publi-

Table 5: Potential sources of endogeneity

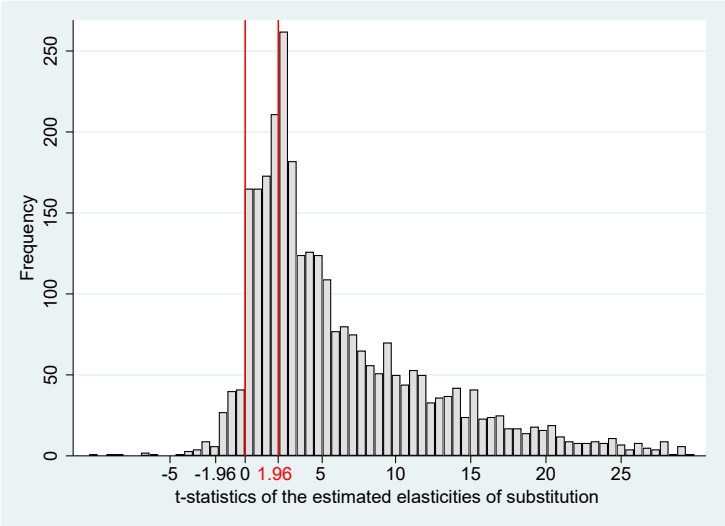
	Identif.	Data aggr.	Results aggr.	K: perpetual	Translog	Short run	All
SE	0.649***	0.803**	0.624***	0.754***	0.664***	0.473***	0.647**
(*publication bias*)	(0.219)	(0.318)	(0.146)	(0.259)	(0.212)	(0.0903)	(0.247)
Constant	0.512***	0.553***	0.569***	0.551***	0.529***	0.587***	0.613***
(*mean beyond bias*)	(0.0357)	(0.0420)	(0.0449)	(0.0337)	(0.0321)	(0.0155)	(0.0421)
SE * Identification	-0.0323						0.0874
	(0.332)						(0.263)
SE * Data aggr.		-0.299					-0.00928
		(0.334)					(0.202)
SE * Results aggr.			0.0616				-0.169
			(0.249)				(0.237)
SE * K: perpetual				-0.334			-0.285
				(0.289)			(0.285)
SE * Translog					-0.127		-0.0127
					(0.344)		(0.312)
SE * Short run						1.741*	1.707**
						(0.885)	(0.846)
Studies	121	121	121	121	121	121	121
Observations	3,186	3,186	3,186	3,186	3,186	3,186	3,186

Notes: Study-level fixed effects and non-interacted variables are included but not reported. Standard errors are reported in parentheses and clustered at the study and country level. Identification = 1 if instrumental variables are used for identification. Data aggregation = 1 if state or country aggregation is used for input data. Results aggregation = 1 if the reported elasticity corresponds to an aggregate one (in contrast to elasticities corresponding to industries disaggregated at least at the 2-digit level). K: perpetual = 1 if input data for capital are measured via the perpetual inventory method. Translog = 1 if the elasticity is estimated using the translog functional form. Short run = 1 if the coefficient is taken from an explicitly short-run specification. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

cation bias affects some segments of the distribution of p-values (under-representation of large p-values, over-representation of p-values just below 0.05), but not the entire distribution. The idea of p-uniform* is to find a coefficient at which the distribution of p-values is approximately uniform; this is achieved by recomputing the reported p-values for various possible values of the underlying effect and then comparing the resulting distribution to the uniform one. In a similar vein, the technique's test for publication bias evaluates whether p-values are uniformly distributed at the precision-weighted mean reported in the literature. (The technique yields a binary result for the test of publication bias and a corresponding p-value.) Once again we obtain evidence for publication bias; the corrected mean elasticity is 0.4.

Another way to approach the endogeneity problem is to explicitly control for the most likely causes of endogeneity. We do so in Table 5, where we include interactions of the standard error with dummy variables for six study characteristics along with study fixed effects. We focus on the following characteristics: the use of IV, data aggregation, results aggregation, the use of the perpetual inventory method to approximate capital, the use of the translog function, and short-run estimation. For example, studies using IV techniques can be expected to deliver less precision, but at the same time systematically different results if endogeneity is an important issue in the primary literature. If a characteristic is associated with publication bias, or simply with systematically different standard errors that might give a false impression of publication bias, the interaction should prove strong. But we see no such pattern. Of the 12 coefficients for interactions estimated in Table 5, one is significant at the 10% level and one at the 5% level,

Figure 7: The distribution of t-statistics shows jumps at 0 and 1.96



Notes: Outliers are excluded from the figure but included in all tests.

which could easily arise by chance. Moreover, the coefficient on the non-interacted standard error remains statistically significant in all cases, and the mean beyond bias remains close to our baseline estimates. We thus fail to model the violations of exogeneity (or, alternatively, the sources of publication bias) explicitly.

The exogeneity assumption can also be relaxed by using the caliper test (Gerber & Malhotra, 2008), which moreover allows us to address the third main issue of our baseline approach, the focus on the zero threshold. The caliper test uses the simple idea that publication bias is the best explanation for sudden jumps in the distribution of the t-statistic. In a narrow caliper around 1.96, for example, the number of t-statistics reported above the threshold should equal the number of t-statistics below the threshold. If the former significantly outweigh the latter, we conclude publication bias likely plagues the literature. The distribution of t-statistics (Figure 7) does indeed show conspicuous jumps: at 0 and 1.96. The jump at 0 is so large that no statistical tests are necessary to conclude that negative estimates are discriminated against, either due to bias or a rational tendency not to report nonsensical results. In Table 6 we test the threshold of 1.96, which is associated with statistical significance at the 5% level. In a narrow caliper of 0.05 (corresponding to t-statistics between 1.935 and 1.985), estimates above the threshold outnumber those below the threshold 30 to 9. The difference remains statistically significant with wider calipers.

In the second and third column of Table 6 we adapt the caliper test to examine publication hurdles other than zero and 5% statistical significance with respect to zero. We focus on two values: 1.3, which is an important benchmark result by Karabarbounis & Neiman (2014), and 1, which corresponds to the Cobb-Douglas case and also the baseline noisy estimate for many regression equations like (4) or those that test the FOC of labor shares. Some studies do explicitly compare their estimates to these benchmarks; for the rest we recompute the t-statistics so that they correspond to this new hypothesis. We ask whether statistical (in)significance of the

Table 6: Caliper tests for t-statistics corresponding to 5% significance thresholds

	Full sample $H_0 : \sigma = 0$ upper threshold	Labor share $H_0 : \sigma = 1.3$ lower threshold	Base at 1 $H_0 : \sigma = 1$ lower threshold
Caliper width = 0.05	0.277^{***}		
	(0.067)		
	N = 39		
Caliper width = 0.1	0.165^{***}	-0.248	
	(0.056)	(0.251)	
	N = 71	N = 4	
Caliper width = 0.15	0.139^{***}	-0.236	
	(0.049)	(0.197)	
	N = 96	N = 6	
Caliper width = 0.2	0.098^{***}	-0.236	
	(0.041)	(0.197)	
	N = 142	N = 6	
Caliper width = 0.25	0.071^{**}	-0.317^{**}	0.322^{**}
	(0.037)	(0.137)	(0.156)
	N = 177	N = 9	N = 7
Caliper width = 0.3	0.088^{***}	-0.338^{**}	0.244^{*}
	(0.033)	(0.123)	(0.165)
	N = 221	N = 10	N = 8
Caliper width = 0.35	0.107^{***}	-0.185	0.266^{*}
	(0.030)	(0.140)	(0.150)
	N = 258	N = 12	N = 9
Caliper width = 0.4	0.106^{***}	-0.128	0.266^{*}
	(0.029)	(0.140)	(0.150)
	N = 292	N = 13	N = 9
Caliper width = 0.45	0.071^{***}	-0.080	0.331^{**}
	(0.027)	(0.137)	(0.125)
	N = 326	N = 14	N = 10
Caliper width = 0.5	0.061^{**}	-0.080	0.315^{**}
	(0.026)	(0.137)	(0.117)
	N = 353	N = 14	N = 12

Notes: The table reports the results of the caliper test by Gerber & Malhotra (2008). The test compares the relative frequency of estimates above and below an important threshold for the t-statistic; with a sufficiently narrow caliper, there should be no difference. We use calipers of different sizes depending on the number of observations available. A test statistic of 0.139, for example, means that 63.9% estimates are above the threshold and 36.1% estimates are below the threshold. Standard errors are reported in parentheses and clustered at the study level. In the first column (full sample) the original reported t-statistics are evaluated. In the second column (labor share) only estimates from papers about the labor share are used, and t-statistics are recomputed to reflect the hypothesis $H_0 : \sigma = 1.3$. In the third column (base at 1) we include only reduced-form estimates for which an estimated regression parameter of zero translates to an elasticity of 1; the t-statistics of the elasticity are recomputed to reflect the hypothesis $H_0 : \sigma = 1$. N = number of estimates. The missing values for some calipers indicate no estimates available for the caliper. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

differences from the benchmarks influences the probability of reporting the estimate. Regarding the value 1.3, we restrict our attention to estimates derived in papers on the labor share because the estimate by Karabarbounis & Neiman (2014) is relevant especially in this context (though the result would hold if all estimates were used). We see little effect of the threshold. Next, when examining the Cobb-Douglas case we include only reduced-form estimates for which a zero regression coefficient translates to an elasticity of 1. The fact that a noisy and small regression coefficient implies a unitary elasticity may affect the mechanism of publication bias, but the caliper test result would hold if we included all the estimates. Here we obtain significant results: estimates that are just consistent with the Cobb-Douglas case are reported more often than those that are significantly smaller than unity at the 5% level. Thus we find evidence of publication bias against three thresholds: positive sign, statistical significance with respect to zero at the 5% level, and consistency with the Cobb-Douglas production function.

5 Heterogeneity

In the previous section we have shown that when we give the same weight to all approaches used in primary studies, the empirical literature as a whole provides no support for the Cobb-Douglas production function. But perhaps poor data and misspecifications bias the mean estimate downwards. We investigate this issue here. In Section 2 and Section 3 we discussed several prominent aspects of study design that might systematically influence the reported estimates of the elasticity. But many additional study characteristics can certainly play a role, and we need to control for them. To assign a pattern to the apparent heterogeneity in the literature, we collect 71 variables that reflect the context in which researchers obtain their estimates. The variables capture the characteristics of the data, specification choice, econometric approach, definition of the production function, and publication characteristics. (Moreover, the effects of different ways of measuring capital and labor are examined in subsamples of the main dataset and presented in Appendix E.) The variables, grouped in these categories, are discussed below and listed in Table D1 in Appendix D together with their definitions and summary statistics.

5.1 Variables

Data characteristics A central distinguishing feature of the studies concerns the source of variation. Almost half (45%) of the studies exploit variation across country or state-level, which forms our reference category. We include dummy variables equal to one if the study exploits variation across industries (43% of the estimates) or firms (12% of the estimates). Nerlove (1967) suggests that exploiting cross-country variation, where there may be systematic correlation between efficiency levels, product prices and wages, can lead to an upward bias in the estimated elasticity. Moreover, Chirinko (2008) discusses several drawbacks of cross-country variation in comparison to firm or industry-level variation, including limited variation available for identification and unaccounted heterogeneity.

We also include a dummy equal to one when the resulting estimate is reported at a very

disaggregated level for various industries. Moreover, we add controls for potential cross-country differences: a dummy for the US, developed European countries, and developing countries, as the substitutability between capital and labor may differ with the level of economic development and across institutional settings. For instance, Duffy & Papageorgiou (2000) suggest that capital and labor become less substitutable in poorer countries.

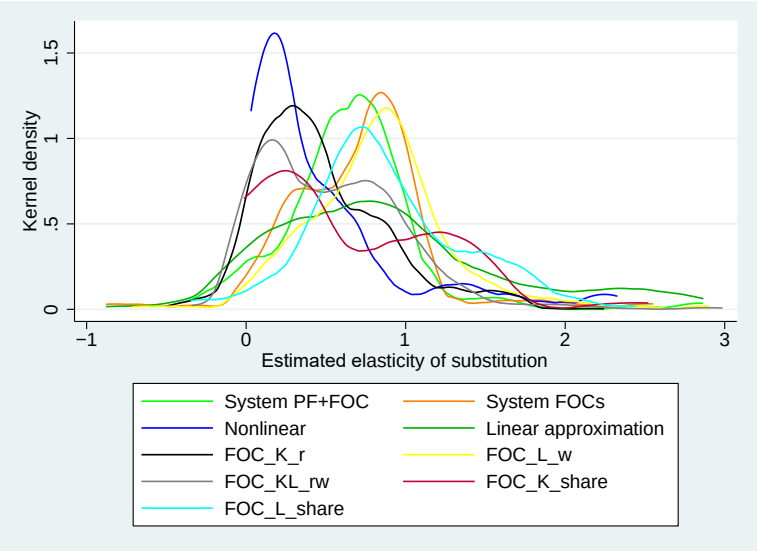
To account for potential small-sample bias, we control for the number of observations used in each study. We also include the midpoint of the data period to capture a potential positive trend in the elasticity over time, which could be due to economic development within a country, a changing composition of the inputs, or changes in their relative efficiency (Cantore *et al.*, 2017). Regarding data frequency, 89% of the estimates employ annual data; we thus use annual data as the baseline category and include a dummy variable for the use of quarterly data. Moreover, we control for data dimension—whether time series, cross-sectional, or panel data are used. Most of the studies employ time series data (around 53%), which we take as the reference category.

The final subset of variables covering data characteristics describes the source of data. Many estimates are based on data from the same databases—the largest number of studies employ data from the US Annual Survey of Manufactures and Census of Manufacturers. The second largest group is the KLEM database by Jorgenson (2007), followed by the OECD's International Sectoral Database and Structural Analysis Database. We do not have a prior on how data sources should affect estimates, yet still prefer not to ignore this potential source of differences in results and include the corresponding dummies as control variables.

Specification Concerning the specification of the various studies described in Section 2, we distinguish between estimation via single first-order conditions (FOCs); systems of more than one FOC; systems of the production function plus FOCs; linear approximations of the production function; and nonlinear estimation of the production function. We also discriminate between the FOC for labor based on the wage rate, FOC for capital based on the rental rate of capital, FOC for the capital-labor ratio based on the ratio between the wage rate and the rental rate of capital, FOC for capital share, and FOC for labor share in income. In total, this gives us nine distinct categories for estimation specification. We choose the FOC for capital based on the rental rate as the reference category because it represents the most frequently used specification (35%), though closely followed by the FOC for labor based on the wage rate (33% of estimates). A special case of the FOC for capital is its inverse estimation, in which the resulting estimates are labeled user-cost elasticities; examples include Smith (2008) and Chirinko *et al.* (2011).

The differences in estimates derived from the various specifications are clearly visible in the data (Figure 8). While the mean of the estimates derived from the FOC for labor based on the wage rate reaches 1.1, estimates derived from the FOC for capital based on the rental rate of capital are on average only 0.5. Estimates obtained from the linear approximation of the production function also stand out, reaching a mean value of 1.1. Some of these patterns were noted early in the history of the estimation of the elasticity, for example, by Berndt (1976), and later discussed by Antras (2004) and Young (2013). We attempt to quantify the patterns, while simultaneously controlling for other influences.

Figure 8: Estimation form matters for the reported elasticities



Notes: A detailed description of the variables is available in Table D1.

Regarding system estimations, two other important specification aspects can influence the reported elasticities: normalization and cross-equation restrictions. Normalization, suggested by de La Grandville (1989), further explored by Klump & de La Grandville (2000), and first implemented empirically by Klump *et al.* (2007), has been used by only a small fraction of the studies in our database. Normalization starts from the observation that a family of CES functions whose members are distinguished only by different elasticities of substitution needs a common benchmark point. Since the elasticity of substitution is defined as a point elasticity, one needs to fix benchmark values for the level of production, factor inputs, and the marginal rate of substitution, or equivalently for per capita production, capital deepening, and factor income shares. Normalization essentially implies representing the production function in a consistent indexed number form. A proper choice of the point of normalization facilitates the identification of deep technical parameters. According to León-Ledesma *et al.* (2010), the superiority of the system estimation compared to the single FOC approach is further enhanced when complemented with normalization. In their Monte Carlo experiment they show that without normalization, estimates tend towards one.

Some estimations of systems employ cross-equation restrictions that restrict parameters across two or more equations to be equal, as in Zarembka (1970), Krusell *et al.* (2000), and Klump *et al.* (2007). To account for possible differences, we additionally include a dummy for cross-equation restrictions.

While the vast majority of estimates come from single-level production functions, estimates of the elasticity of substitution between capital and labor can also be found in studies using two-level production functions, including additional inputs such as energy and material, (e.g., Van der Werf, 2008; Dissou *et al.*, 2015). We control for two-level production functions as a special case. Moreover, when estimates of the elasticity rely on such two-level production func-

tions, linear approximations of the production function, or a system of a linear approximation in conjunction with share factors, researchers commonly report partial elasticities of substitution, for which we control as well. Our results are robust to excluding partial elasticities.

Econometric approach Our reference category for the choice of the econometric technique is OLS. We include a dummy for the case when the model is dynamic, which holds for approximately one-quarter of all observations. The second dummy we include equals one if seemingly unrelated regression (SUR) is used—often employed for the estimation of systems of equations (11% of all estimates). An important aspect of estimating the elasticity, as pointed out by Chirinko (2008), is whether the estimate refers to a long-run or a short-run elasticity. Our reference category consists of explicit long-run specifications, that is, models in which coefficients are meant to be long-run and the specification is adjusted accordingly. We opt for long-run elasticities as a reference point as they are regarded as more informative for economic decisions. Explicit long-run specifications include estimations of cointegration relations or interval-difference models, where data are averaged over longer intervals to mimic lower frequencies; distributed lag models can also give a long-run estimate. Conversely, the short-run approach modifies the estimating equation to account for temporal dynamics. Examples include estimation of implicit investment equations, as in Eisner & Nadiri (1968) or Eisner (1969), differenced models, and estimation of short-run elements from error correction models or distributed lag models. The vast majority of estimates (70%) are meant to be long-run but the specification is unadjusted.

Production function components The fourth category of control variables comprises the ingredients of the production function. We include a dummy variable for the case when other inputs (energy, materials, human capital) are considered as additional factors of production, for instance by Humphrey & Moroney (1975), Bruno & Sachs (1982), and Chirinko & Mallick (2017). We include a dummy that equals one when a study differentiates between skilled and unskilled labor. We also subject the estimates to the following questions. Does the production function assume Hicks-neutral technological change (our reference category), Harrod-neutral technological change (i.e. labor-augmenting, LATC), or Solow-neutral technological change (i.e. capital-augmenting, CATC)? Are the dynamics of technological change important in explaining the heterogeneity? The growth rate of technological change can be either zero (our reference), constant or—with flexible Box & Cox (1964) transformation—exponential, hyperbolic, or logarithmic. According to the impossibility theorem suggested by Diamond *et al.* (1978), it is infeasible to identify both the elasticity of substitution and the parameters of technological change at the same time, so researchers tend to impose one of the three specific forms of technological change and implicit or explicit assumptions on its growth rate. We include the corresponding dummy variables.

We distinguish between estimates of gross and net elasticity, based on whether gross or net data for output and the capital stock are used. As pointed out in Semieniuk (2017), the distinction between net and gross elasticity is important with respect to the inequality argument of Piketty (2014): for his explanation of the decline in the labor share to hold, σ needs to

exceed one in net terms. Elasticities based on net quantities should naturally yield smaller results (Rognlie, 2014). Finally, we include two additional dummies—first, for the case when researchers abandon the assumption of constant returns to scale; second, for the case when researchers relax the assumption of perfectly competitive markets.

Publication characteristics We include four study-level variables: the year of the appearance of the first draft of the paper in Google Scholar, a dummy for the paper being published in a top five journal, the recursive discounted RePEc impact factor of the outlet, and the number of citations per year since the first appearance of the paper in Google Scholar. We include these variables in order to capture aspects of study quality not reflected by observable differences in data and methods.

Moreover, we include two additional dummies. The first variable measures whether the study's central focus is the elasticity of substitution between capital and labor or whether the estimate is a byproduct of a different exercise, such as in Cummins & Hassett (1992) and Chwelos *et al.* (2010). The second variable equals one if the author explicitly prefers the estimate in question, and equals minus one if the estimate is explicitly discounted. Nevertheless, researchers typically do not reveal their exact preferences regarding the individual estimates they produce, so the variable equals zero for most estimates.

5.2 Estimation

An obvious thing to do at this point is to regress the reported elasticities on the variables reflecting the context in which researchers obtain their estimates:

$$\hat{\sigma}_{ij} = \alpha_0 + \sum_{l=1}^{49} \beta_l X_{l,ij} + \gamma SE(\hat{\sigma}_{ij}) + \mu_{ij}, \tag{14}$$

where $\hat{\sigma}_{ij}$ again denotes estimate i of the elasticity of substitution reported in study j, $X_{l,ij}$ represents control variables described in Subsection 5.1, γ again denotes the intensity of publication bias, and α_0 represents the mean elasticity corrected for publication bias but *conditional* on the definition of the variables included in X—that is, the intercept means nothing on its own, and μ_{ij} stands for the error term.

But using one regression is inadequate because of model uncertainty. With so many variables reflecting study design, including all of them would substantially attenuate the precision of our estimation. (We use 50 variables in the baseline estimation; the remaining 21 variables related to measurement of capital and labor and industry-level characteristics are included in the three subsamples presented in Appendix E.) One solution is to reduce the number of variables to about 10, which could allow for simple estimation—but doing so would ignore many aspects in which estimates and studies differ. Another commonly applied solution to model uncertainty is stepwise regression, but sequential t-tests are statistically problematic as individual variables can be excluded by accident. The solution that we choose here is Bayesian model averaging

(BMA; see, for example, Eicher *et al.*, 2011; Steel, 2020), which arises naturally as a response to model uncertainty in the Bayesian setting.[4]

BMA runs many regression models with different subsets of variables; in our case there are 2^{50} possible subsets. Assigned to each model is a posterior model probability (PMP), an analog to information criteria in frequentist econometrics, measuring how well the model performs compared to other models. The resulting statistics are based on a weighted average of the results from all the regressions, the weights being the posterior model probabilities. For each variable we thus obtain a posterior inclusion probability (PIP), which denotes the sum of the posterior model probabilities of all the models in which the variable is included. Using the laptop on which we wrote this paper, it would take us decades to estimate all the possible models. So we opt for a model composition Markov Chain Monte Carlo algorithm (Madigan & York, 1995) that walks through the models with the highest posterior model probabilities. In the baseline specification we use a uniform model prior (each model has the same prior probability) and unit information g-prior (the prior that all regression coefficients equal zero has the same weight as one observation in the data), but we also use alternative priors in Appendix E. BMA has been used in meta-analysis, for example, by Havranek *et al.* (2015); Zigraiova & Havranek (2016); Havranek *et al.* (2018a,b,c).

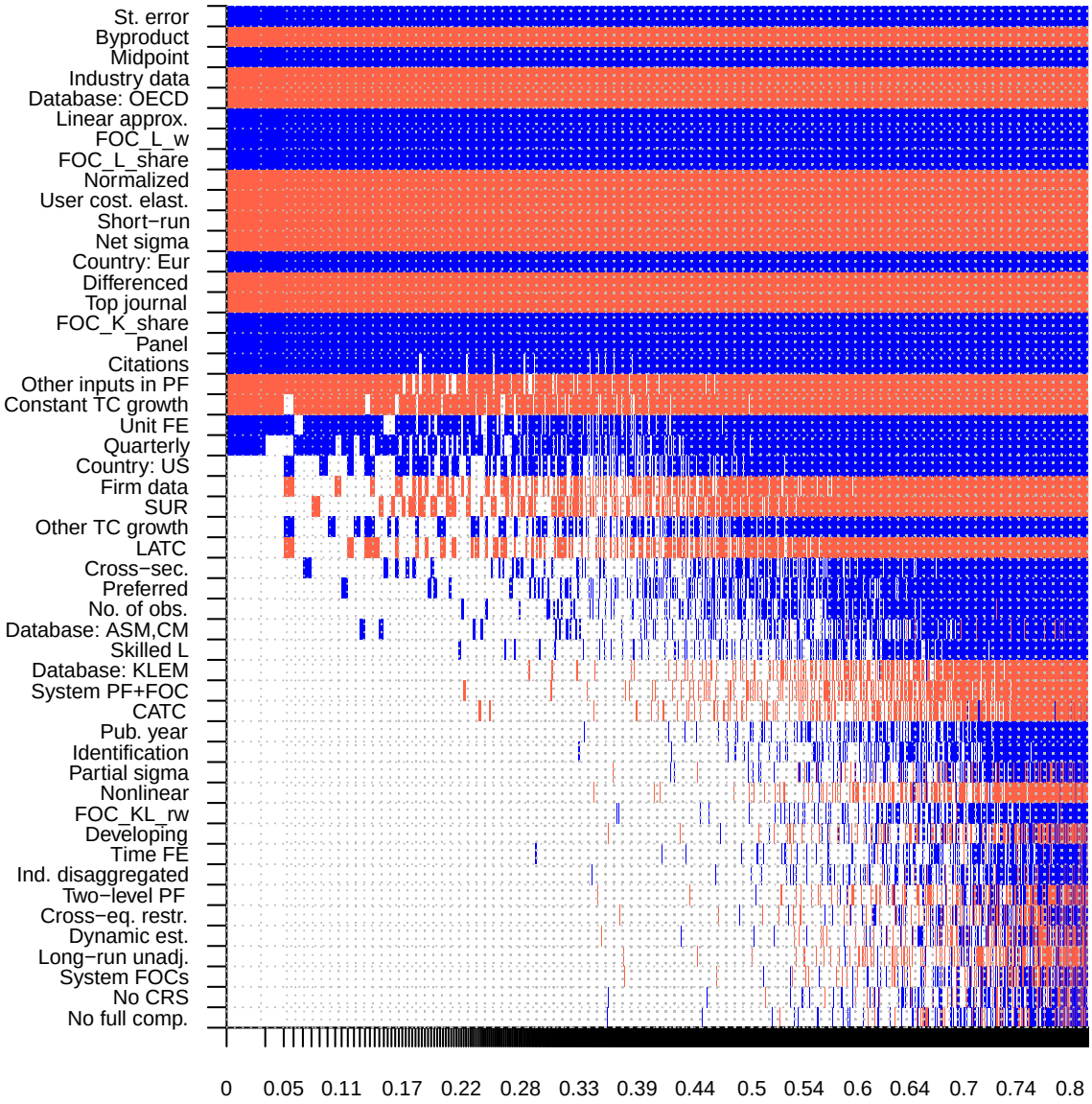
Second, as a simple robustness check of our baseline BMA specification, we run a hybrid frequentist-Bayesian model. We employ variable selection based on BMA (specifically, we only include the variables with PIPs above 80%) and estimate the resulting model using OLS with clustered standard errors. We label this specification a "frequentist check" of the baseline BMA exercise. Third, we employ frequentist model averaging (FMA). Our implementation of FMA uses Mallows's criteria as weights since they prove asymptotically optimal (Hansen, 2007). The problem is that, using a frequentist approach, we have no straightforward alternative to the model composition Markov Chain Monte Carlo algorithm, and it appears infeasible to estimate all 2^{50} potential models. We therefore follow the approach suggested by Amini & Parmeter (2012) and resort to orthogonalization of the covariate space.

5.3 Results

Figure 9 illustrates our results. The vertical axis depicts explanatory variables sorted by their posterior inclusion probabilities; the horizontal axis shows individual regression models sorted by their posterior model probabilities. The blue color indicates that the corresponding variable appears in the model and the estimated parameter has a positive sign, while the red color indicates that the estimated parameter is negative. In total, 21 variables appear to drive heterogeneity in the estimates, as their posterior inclusion probabilities surpass 80%. Table 7

[4]If a simple OLS brought results similar to model averaging, we could simplify the analysis and just present OLS. But in our case a simple OLS regression including all variables would yield results quite different from Bayesian model averaging: 29% of the variables would lose their statistical significance (or importance in the Bayesian setting), while 17% of the variables would now be wrongly significant. The median change in the magnitude of the estimated coefficients for these variables would reach 133% in absolute value. (But note that our key results concerning publication bias and the best-guess elasticity would continue to hold.) We thus opt for the more complex but statistically more appropriate approach.

Figure 9: Model inclusion in Bayesian model averaging



Notes: The response variable is the estimate of the elasticity of capital-labor substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. FOC = first-order condition. CATC = capital-augmenting technical change. LATC = labor-augmenting technical change. CRS = constant returns to scale. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. To ensure convergence we employ 100 million iterations and 50 million burn-ins. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model. Numerical results of the BMA exercise are reported in Table 7. A detailed description of all variables is available in Table D1.

Table 7: Why do estimates of the elasticity of substitution differ?

Response variable:	Bayesian model averaging			Frequentist check		
Estimate of σ	Post. mean	Post. SD	PIP	Coef.	Std. er.	p-value
SE (publication bias)	0.614	0.038	1.000	0.633	0.042	0.000
Data characteristics						
No. of obs.	0.003	0.009	0.107			
Midpoint	0.118	0.022	1.000	0.123	0.036	0.001
Cross-sec.	0.009	0.023	0.160			
Panel	0.161	0.041	0.985	0.177	0.048	0.000
Quarterly	0.070	0.060	0.642			
Firm data	-0.033	0.049	0.363			
Industry data	-0.191	0.026	1.000	-0.191	0.064	0.003
Country: US	0.030	0.036	0.468			
Country: Eur	0.119	0.029	1.000	0.103	0.051	0.043
Developing country	0.000	0.003	0.014			
Database: ASM/CM	0.004	0.016	0.071			
Database: OECD	-0.277	0.039	1.000	-0.276	0.099	0.005
Database: KLEM	-0.003	0.014	0.042			
Disaggregated σ	0.000	0.003	0.012			
Specification						
System PF+FOC	-0.002	0.014	0.039			
System FOCs	0.000	0.003	0.008			
Nonlinear	-0.001	0.011	0.016			
Linear approx.	0.235	0.039	1.000	0.227	0.108	0.037
FOC_L_w	0.278	0.023	1.000	0.261	0.023	0.000
FOC_KL_rw	0.000	0.005	0.015			
FOC_K_share	0.230	0.064	0.993	0.212	0.253	0.402
FOC_L_share	0.209	0.038	1.000	0.204	0.064	0.001
Cross-equation restr.	0.000	0.004	0.010			
Normalized	-0.277	0.038	1.000	-0.289	0.066	0.000
Two-level PF	0.000	0.007	0.011			
Partial σ	0.001	0.012	0.017			
User cost elast.	-0.385	0.044	1.000	-0.368	0.061	0.000
Econometric approach						
Dynamic est.	0.000	0.003	0.009			
SUR	-0.027	0.041	0.348			
Identification	0.000	0.005	0.018			
Differenced	-0.111	0.025	1.000	-0.109	0.025	0.000
Time FE	0.000	0.006	0.013			
Unit FE	0.093	0.065	0.735			
Short-run σ	-0.380	0.034	1.000	-0.381	0.053	0.000
Long-run σ unadj.	0.000	0.002	0.009			
Production function components						
Other inputs in PF	-0.103	0.054	0.852	-0.128	0.070	0.068
CATC	-0.001	0.007	0.038			
LATC	-0.018	0.028	0.327			
Skilled L	0.006	0.029	0.061			
Constant TC growth	-0.078	0.040	0.844	-0.101	0.038	0.009
Other TC growth	0.029	0.045	0.332			
No CRS	0.000	0.002	0.008			
No full comp.	0.000	0.004	0.008			
Net σ	-0.376	0.048	1.000	-0.260	0.054	0.000
Publication characteristics						
Top journal	-0.092	0.023	0.998	-0.074	0.032	0.021
Pub. year	0.000	0.004	0.024			
Citations	0.033	0.014	0.916	0.037	0.018	0.040
Preferred est.	0.005	0.014	0.154			
Byproduct	-0.152	0.028	1.000	-0.143	0.075	0.059
Constant	0.059		1.000	0.071	0.143	0.619
Observations	3,186			3,186		

Notes: σ = elasticity of capital-labor substitution, PIP = posterior inclusion probability. SD = standard deviation. FOC = first-order condition. CATC = capital-augmenting technical change. LATC = labor-augmenting technical change. CRS = constant returns to scale. The table shows unconditional moments for BMA. In the frequentist check we include only explanatory variables with PIP > 0.8. The standard errors in the frequentist check are clustered at the study level. A detailed description of all variables is available in Table D1.

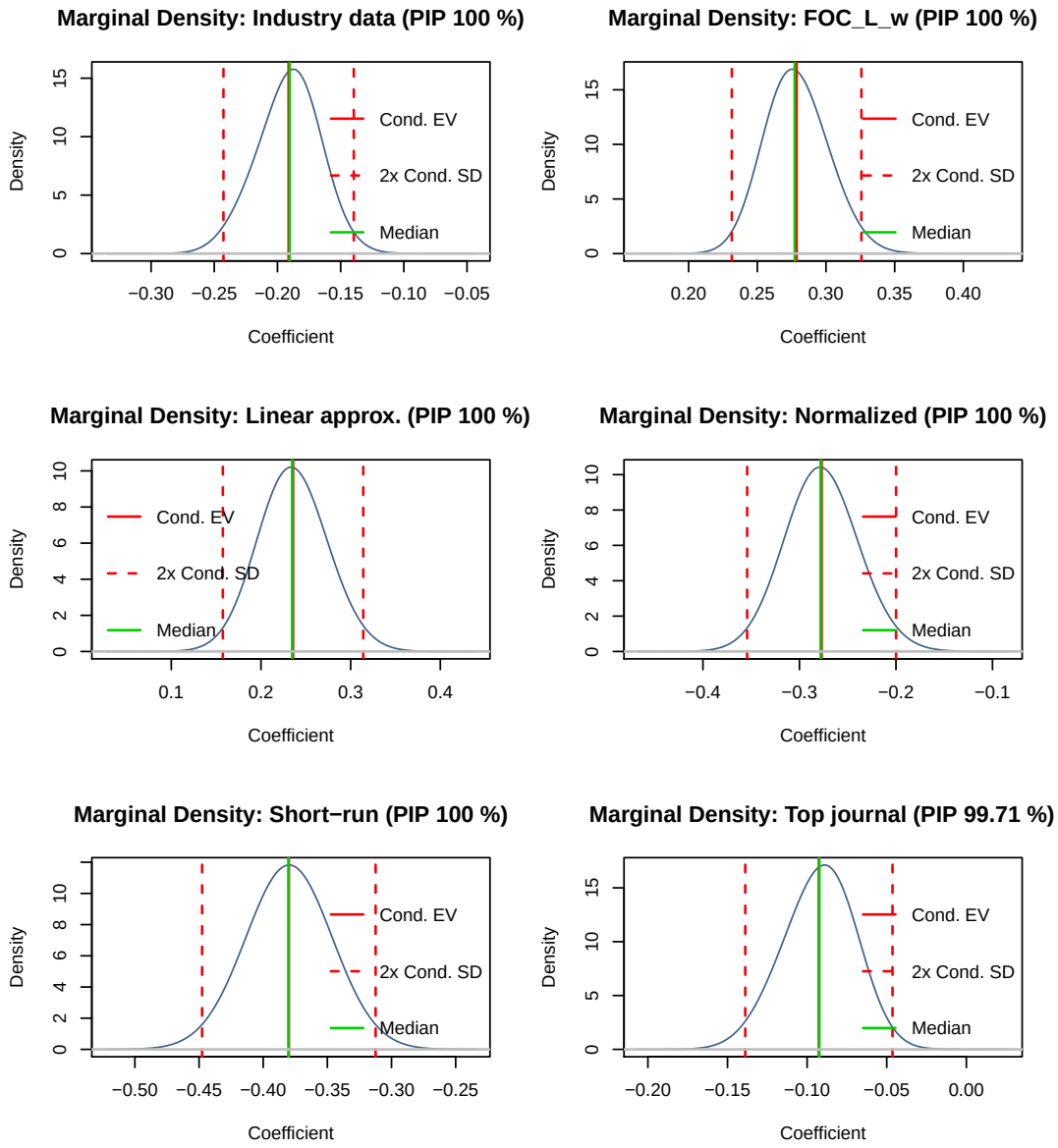
provides numerical results for BMA and the frequentist check. In the frequentist check we only include the 21 variables with PIPs above 80%. Choosing a 50% threshold, for example, would result in including merely two more variables with virtually unchanged results for the remaining ones. Figure 10 plots posterior coefficient distributions of selected variables. The results of the FMA exercise are reported in Table E1 in Appendix E.

The first conclusion that we make based on these results is that our findings of publication bias presented in the previous section remain robust when we control for the context in which the elasticity is estimated. Indeed, the variable corresponding to publication bias, the standard error of the estimate, represents the single most effective variable in explaining the heterogeneity in the reported estimates of the elasticities of substitution (though several other variables also have posterior inclusion probabilities very close to 100% and are rounded to that number in Table 7). We observe that the publication bias detected by the correlation between estimates and standard errors is not driven by aspects of data and methods omitted from the univariate regression in Equation 13.

Data characteristics Several characteristics related to the data used in primary studies systematically affect the estimates of the elasticity. Our results suggest a mild upward trend in the reported elasticities, which increase on average by 0.004 each year. (The yearly change does not equal the regression coefficient because the variable is in logs; the precise definition is available in Table D1.) The finding resonates with Cantore *et al.* (2017), who point to a similar time trend. But the upward trend constitutes a poor reason to resurrect the Cobb-Douglas specification, because at this pace the specification will become consistent with the literature in about 175 years. Next, estimates of the elasticity that exploit variation across industries tend to be significantly smaller than those using variation across countries and states, a result corroborating the prima facie pattern in the literature shown in Figure 5(d) in Section 3. This is consistent with Nerlove (1967) and Chirinko (2008), who argue that exploiting variation across countries can lead to an upward bias due to disregarded heterogeneity.

Concerning data dimension, our results suggest that panel data tend to yield larger estimates of the elasticity than time series data. The other prima facie pattern in the literature, the systematic and large difference between the results of time series and cross-section studies shown in Figure 5(c), breaks apart when controlling for other variables in BMA (the variable is statistically significant in FMA, but the estimated coefficient is small). Similarly, our results do not suggest that much of the differences between estimates can be explained by differences in data frequency. Another prima facie data pattern, the importance of results aggregation presented in Figure 5(b), disappears in the BMA analysis. Elasticities computed for individual industries do not differ systematically from elasticities computed for the entire economy. Nevertheless, that is not to say that the elasticity does not vary across industries; we will return to this issue in Appendix E. Concerning cross-country differences, the reported elasticities tend to be larger in Europe than in other regions, but only by 0.1. Finally, our results suggest that datasets coming from the OECD database are associated with substantially smaller elasticities compared to all other data sources.

Figure 10: Posterior coefficient distributions for selected variables



Notes: FOC_L_w = 1 if the elasticity is estimated within the FOC for labor based on the wage rate. The figure depicts the densities of the regression parameters encountered in different regressions in which the corresponding variable is included (that is, the depicted mean and standard deviation are conditional moments, in contrast to those shown in Table 7). For example, the regression coefficient for Linear approximation is positive in all models, irrespective of specification. The most common value of the coefficient is 0.23.

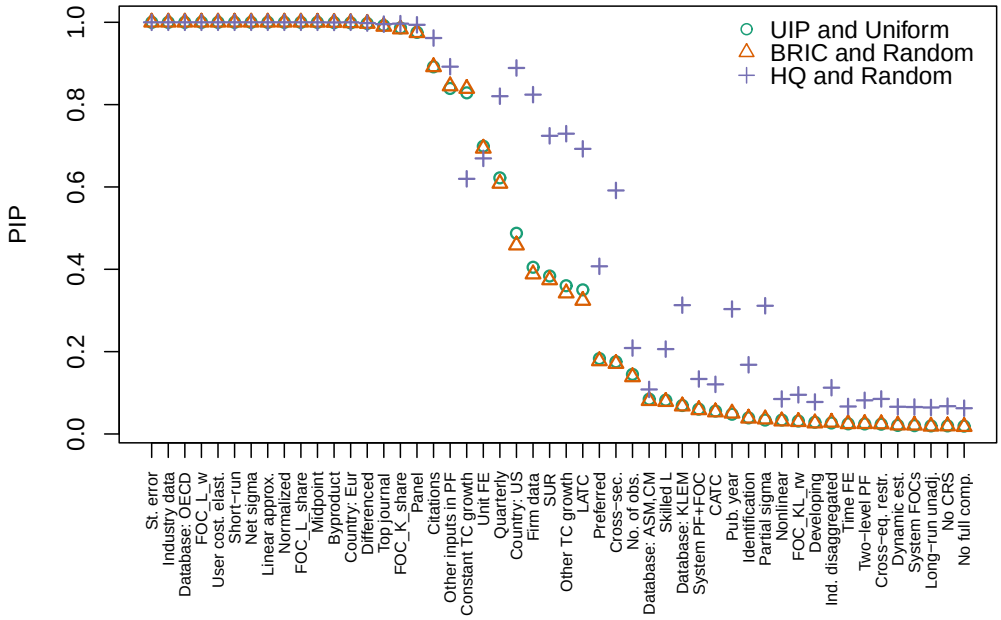
Specification A stylized fact in the literature on capital-labor substitution has it that estimations based on the first-order condition for labor deliver larger elasticities than estimations based on the first-order condition for capital; see Figure 5(a) in Section 3. The BMA analysis corroborates this stylized fact and elaborates on it: when a system of FOCs is used, the results tend to be close to those derived from the FOC for capital. Omitting information from the FOC for capital, in contrast, exaggerates the reported elasticity by 0.2 or more. The FOC for capital thus seems to be more important for proper identification of the elasticity than the

FOC for labor. The elasticity also becomes inflated by 0.2 when a linear approximation of the production function (using either the Kmenta or translog approach) is employed. As pointed out by Thursby & Lovell (1978) and León-Ledesma *et al.* (2010), linear approximations of the production function tend to be biased towards $\hat{\sigma} = 1$, as an elasticity of one usually serves as the initial point of expansion.

On the other hand, normalization of the production function systematically reduces the estimated elasticity by allowing for the identification of technological change parameters. Finally, if the FOC for capital is estimated in an inverse form (user cost elasticity of capital), the estimates tend to be on average much smaller. These results are robust across all the estimations we run: BMA, FMA, and the frequentist check. A similarly robust result is that the mean implied elasticity is 0.3 when made conditional on three aspects: (i) no publication bias, (ii) no use of cross-country variation in input data, and (iii) not ignoring information from the FOC for capital. We will expand and provide more details on the computation of the implied elasticity at the end of this section.

Econometric approach We find little evidence that the econometric approach used in primary studies is responsible for systematic differences in the reported elasticities. Naturally, short-run elasticities are smaller than long-run ones: estimations in differences tend to deliver elasticities that are smaller by 0.1; explicitly short-run estimations tend to deliver elasticities smaller by 0.4. Adjusted and unadjusted long-run estimates do not differ much from each other.

Figure 11: Posterior inclusion probabilities across different prior settings



Notes: UIP (unit information prior) and Uniform model prior = priors according to Eicher *et al.* (2011). BRIC and Random = the benchmark g-prior for parameters with the beta-binomial model prior for the model space, which means that each model size has equal prior probability. HQ prior asymptotically mimics the Hannan-Quinn criterion. PIP = posterior inclusion probability.

Production function components The results suggest that assumptions regarding technical change have little systematic effect on the resulting elasticities of substitution. Allowing for capital- or labor-augmenting technological change brings, on average, elasticities similar to the case when Hicks-neutral technological change is assumed. Allowing for constant growth in technological change (in comparison to no growth) decreases the estimate, but only by a small margin. The apparent irrelevance of assumptions on technological change for the estimation of the elasticity of substitution contrasts with Antras (2004), who argues that Hicks-neutral technological change biases the results towards the Cobb-Douglas specification. The irrelevance finding holds for both BMA and FMA and regardless of whether we include labor- and capital-augmenting technological change as separate dummies or jointly in one dummy.

Including other inputs in the production function aside from labor and capital has a negative effect on the resulting size of the elasticity. When the elasticity is estimated in the net form, it tends to be smaller by 0.4 on average, but very few studies pursue this approach.

Publication characteristics Out of the five variables grouped together as publication characteristics, three are systematically associated with the magnitude of the reported elasticity. First, compared to other outlets, the top five journals in economics tend to publish slightly smaller elasticities. Second, studies that provide larger elasticities tend to receive more citations—potentially, such studies are more useful to researchers trying to justify the use of the Cobb-Douglas production function in their model, but it could also mean that studies reporting larger estimates are of higher quality. Third, the reported elasticity tends to be smaller if it does not represent the central focus of the study but merely a byproduct of a different exercise. One can interpret the finding as further indirect evidence of publication bias against small estimates, or, alternatively, as evidence that more thorough examinations yield larger estimates.

Aside from our baseline BMA, FMA, and frequentist check, we run several sensitivity analyses with respect to different subsamples of data, control variables, priors, and weighting schemes. Regarding priors, Figure 11 shows that the implied relative importance of the variables changes little when different priors are used for BMA. In Appendix E, we also run BMA on weighted data: first, data are weighted by the inverse of the number of estimates reported by each study so that each study has the same weight (Figure E1); second, data are weighted by the inverse of the standard error (Figure E2). Our key results continue to hold in these specifications.

Economic significance and implied elasticity

We close the analysis with a discussion of (i) the economic significance of the variables identified as important by BMA and FMA and (ii) the mean elasticity of substitution implied by the literature after taking into account the pattern that some data and method choices create in the reported estimates. Economic significance is explored in Table 8, which shows the effect on the reported elasticity when we increase the value of the corresponding variable by one standard deviation (the left-hand panel) and from minimum to maximum (the right-hand panel). Increasing from minimum to maximum perhaps makes more sense for dummy variables, while

Table 8: Economic significance of key variables

	One-std.-dev. change		Maximum change	
	Effect on σ	% of best practice	Effect on σ	% of best practice
Standard error	0.117	39%	0.461	154%
Byproduct	-0.047	-16%	-0.152	-51%
Midpoint	0.056	19%	0.588	196%
Industry data	-0.095	-32%	-0.191	-64%
Database: OECD	-0.069	-23%	-0.277	-92%
Linear approx.	0.062	21%	0.235	78%
FOC_L_w	0.132	44%	0.278	93%
Normalized	-0.061	-20%	-0.277	-92%
Short-run σ	-0.083	-28%	-0.380	-127%
Net σ	-0.059	-20%	-0.376	-125%

Notes: The table shows *ceteris paribus* changes in the reported elasticities implied by changes in the variables that reflect the context in which researchers obtain their estimates. For example, increasing the estimate's standard error by one standard deviation is associated with an increase in the estimated elasticity by 0.117, more than a third of the size of the best practice estimate (one conditional on ideal data, method, and publication characteristics, as described in Table 9). Increasing the standard error from the sample minimum to the sample maximum is associated with an increase in the estimated elasticity by 0.461, more than one and a half of the best practice estimate. A detailed description of the variables is available in Table D1 in Appendix D.

for continuous variables, such as the midpoint of data, the one-standard-deviation change is typically more informative. In the second and fourth column, the table also casts the effects as percentages of the "best-practice" estimate implied by the literature, which we discuss below. It is apparent from the table that the variables with the largest effect on the elasticity are the standard error (publication bias), use of variation at the industry level, FOC for labor (ignoring FOC for capital), normalization of the production function, and the assumption of short-run or net elasticity. Changes in these variables can alter the resulting elasticity by 50% or more.

The mean implied elasticity is explored in Table 9. In essence, we create a synthetic study in which we use all the reported estimates but give different weights to certain aspects of data, methodology, and publication. We have already noted that the implied elasticity is 0.3, when we hold three preferences: the estimate should be conditional on the absence of publication bias, use of variation across industries instead of countries, and use of information from the first-order condition for capital. Next, we augment the list of preferences to construct a best-practice estimate. For the computation we use the results of FMA because, unlike BMA, it allows us to construct confidence intervals around the implied elasticities (linear combinations of FMA coefficients and the chosen values for each variable). We compute fitted values of the elasticity by plugging in sample maxima for variables reflecting best practice in the literature, sample minima for variables reflecting departures from best practice, and sample means for variables where we cannot determine best practice.

We prefer large studies using newer data, so we plug in sample maxima for the number of observations and midpoint of data. We prefer a system of production function together with FOCs for both capital and labor, tied with normalization and cross-equation restrictions. We also prefer the use of factor-augmenting technological change and joint estimation of equations by Zellner's method instead of OLS. As for the publication characteristics, we prefer studies

Table 9: Results from a synthetic study

	Implied elasticity	95% confidence interval
Best practice	0.30	(-0.01, 0.60)
Short-run	-0.11	(-0.38, 0.15)
Net σ	-0.02	(-0.30, 0.25)
Country-level data	0.50	(0.18, 0.81)
Quarterly data	0.42	(0.08, 0.76)
Time series	0.25	(-0.10, 0.60)
Cross-sections	0.32	(0.07, 0.56)
System of FOCs	0.35	(0.07, 0.64)

Notes: The table shows mean estimates of the elasticity of substitution conditional on data, method, and publication characteristics. The exercise is akin to a synthetic study that uses all information reported in the literature but puts more weight on selected aspects of study design. The result in the first column is conditional on our definition of best practice (see the main text for details). The remaining rows change one aspect in the definition of best practice: for example the second row shows the result for short-run instead of long-run estimates.

that are highly cited and published in top journals. In contrast, we do not prefer linear approximation, byproduct estimates, elasticities that are supposed to be long-run but are not properly adjusted, and partial elasticities: we plug in zero for these variables. We do not have any strong opinion on the various sources of data or data dimension (whether time series or cross-sectional studies should be used, what data frequency should be employed). Thus, next to the central "best practice" estimate we generate multiple estimates for these data and method choices. We also show implied elasticities for exploiting variation across countries, often used in the literature, and for short-run elasticity, net elasticity, and the use of a system of FOCs without a production function.

The results, shown in Table 9, illustrate the high degree of uncertainty that such an exercise entails: the 95% confidence intervals for all estimates are approximately 0.6 wide. Our central estimate is still 0.3, which means that other aspects of best practice (on top of the three preferences made in the beginning) cancel each other out—even though now the estimate becomes barely statistically significantly different from zero at the 5% level. But even such a conservative estimation rejects the Cobb-Douglas specification in all cases. The implied short-run and net elasticities are close to zero. When one prefers quarterly data instead of showing equal treatment to estimates derived from data of different frequencies, the implied estimate increases to 0.4. A preference for time series data, cross-sectional data, or a system of FOCs without a production function would result in a smaller change in the elasticity. Even a preference for exploiting variation across countries would only take the implied estimate to 0.5, with the upper bound of the 95% confidence interval at 0.8, making the result safely inconsistent with the Cobb-Douglas specification.

6 Concluding Remarks

The Cobb-Douglas production function contradicts the data. This is the result we obtain after analyzing 3,186 estimates of the capital-labor substitution elasticity reported in 121 published

studies. When we give the same weight to all the different approaches used to identify the elasticity, we find that the value most representative of the literature is 0.5, tightly estimated with the upper bound of the 95% confidence interval at 0.6. The representative value corresponds to the mean reported elasticity corrected for publication bias, a phenomenon that has not been previously addressed in the vast literature on the elasticity of substitution. The representative estimate further shrinks to 0.3 when one imposes the restrictions that identification must come from industry-level instead of aggregated, country-level data and that information from the first-order condition for capital must be considered instead of ignored. The representative estimate stays at 0.3 when we control for 71 aspects of study design and select a best-practice value for each aspect (plugging in mean values where no reasonable choice can be made). Such best-practice elasticity is imprecisely estimated, with the upper bound of the 95% confidence interval still at 0.6. Other researchers will have different opinions on what constitutes best practice and might arrive at a point estimate different from 0.3. But no matter the preferences, after acknowledging publication bias, the Cobb-Douglas production function with the elasticity at 1 becomes indefensible in the light of empirical evidence.

We are not the first to highlight the disconnect between the Cobb-Douglas specification commonly used in macroeconomic models and the empirical literature estimating the elasticity of substitution. Chirinko (2008) and Knoblach *et al.* (2020) provide useful surveys of portions of the literature, and both studies suggest that the Cobb-Douglas production function is not backed by the available evidence. We argue that after controlling for publication bias and model uncertainty the case against Cobb-Douglas strengthens to the point where one has to warn against the continued use of this convenient simplification. As we show in the Introduction, a structural model built to aid monetary policy is biased from the beginning if it uses an elasticity of one for capital-labor substitution. Computational convenience should yield to the stylized fact established by half a century of meticulous research: capital and labor are gross complements.

Three caveats to the value of our central estimate, 0.3, are in order. First, the elasticities that we collect are unlikely to be independent because they are frequently derived from the same or similar datasets. We partially address this problem by clustering standard errors at both the study and country level when controlling for publication bias and additionally compute wild bootstrap confidence intervals. Second, the value of 0.3 is a mean estimate and certainly does not fit all situations and calibrations. While we are able to address several issues that we see as problems in the literature, in meta-analysis one can only solve methodological problems that have already been addressed by at least one previous study. The value of 0.3 is our best guess conditional on the available literature published prior to 2019, not the definitive point estimate for the elasticity. Third, we do our best to include all published studies estimating the elasticity of substitution, but still we might have missed some. Such an omission will not affect our results much as long as it remains random. We experimented with randomly omitting 50% of our data set, and the main findings continue to hold in such simulations.

References

van Aert, R. C. & M. van Assen (2020): "Correcting for publication bias in a meta-analysis with the p-uniform* method." *Working paper*, Tilburg University.

Amini, S. M. & C. F. Parmeter (2012): "Comparison of Model Averaging Techniques: Assessing Growth Determinants." *Journal of Applied Econometrics* **27(5)**: pp. 870–876.

Andrews, I. & M. Kasy (2019): "Identification of and Correction for Publication Bias." *American Economic Review* **109(8)**: pp. 2766–2794.

Antras, P. (2004): "Is the US Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution." *Contributions in Macroeconomics* **4(1)**: pp. 1–36.

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

Ashenfelter, O. & M. Greenstone (2004): "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review* **94(2)**: pp. 454–460.

Astakhov, A., T. Havranek, & J. Novak (2019): "Firm Size And Stock Returns: A Quantitative Survey." *Journal of Economic Surveys* **33(5)**: pp. 1463–1492.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, & J. Van Reenen (2017): "Concentrating on the Fall of the Labor Share." *American Economic Review* **107(5)**: pp. 180–85.

Bajzik, J., T. Havranek, Z. Irsova, & J. Schwarz (2020): "Estimating the Armington elasticity: The importance of study design and publication bias." *Journal of International Economics* **127(C)**: p. 103383.

Baker, R. D. & D. Jackson (2013): "Meta-Analysis Inside and Outside Particle Physics: Two Traditions that Should Converge?" *Research Synthesis Methods* **4(2)**: pp. 109–124.

Behrman, J. R. (1972): "Sectoral Elasticities of Substitution between Capital and Labor in a Developing Economy: Times Series Analysis in the Case of Postwar Chile." *Econometrica* **40(2)**: pp. 311–326.

Berndt, E. R. (1976): "Reconciling Alternative Estimates of the Elasticity of Substitution." *Review of Economics and Statistics* **58(1)**: pp. 59–68.

Blanco-Perez, C. & A. Brodeur (2020): "Publication Bias and Editorial Statement on Negative Findings." *Economic Journal* **130(629)**: pp. 1226–1247.

Bom, P. R. D. & H. Rachinger (2019): "A Kinked Meta-Regression Model for Publication Bias Correction." *Research Synthesis Methods* **10(4)**: pp. 497–514.

Box, G. E. & D. R. Cox (1964): "An Analysis of Transformations." *Journal of the Royal Statistical Society: Series B (Methodological)* **26(2)**: pp. 211–243.

Brodeur, A., N. Cook, & A. Heyes (2020): "Methods Matter: P-Hacking and Causal Inference in Economics." *American Economic Review* **110(11)**: pp. 3634–60.

Brodeur, A., M. Lé, M. Sangnier, & Y. Zylberberg (2016): "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

Brown, M. (1966): "A Measure of the Change in Relative Exploitation of Capital and Labor." *Review of Economics and Statistics* **48(2)**: pp. 182–192.

Bruno, M. & J. Sachs (1982): "Input Price Shocks and the Slowdown in Economic Growth: The Case of UK Manufacturing." *Review of Economic Studies* **49(5)**: pp. 679–705.

Bruns, S. B. & J. P. A. Ioannidis (2016): "p-Curve and p-Hacking in Observational Research." *PloS ONE* **11(2)**: p. e0149144.

Cameron, A. C., J. B. Gelbach, & D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* **90(3)**: pp. 414–427.

Cameron, A. C., J. B. Gelbach, & D. L. Miller (2011): "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics* **29(2)**: pp. 238–249.

Cantore, C., F. Ferroni, & M. A. Leon-Ledesma (2017): "The Dynamics of Hours Worked and Technology." *Journal of Economic Dynamics and Control* **82**: pp. 67–82.

Cantore, C., M. León-Ledesma, P. McAdam, & A. Willman (2014): "Shocking Stuff: Technology, Hours, and Factor Substitution." *Journal of the European Economic Association* **12(1)**: pp. 108–128.

Cantore, C., P. Levine, J. Pearlman, & B. Yang (2015): "CES Technology and Business Cycle Fluctuations." *Journal of Economic Dynamics and Control* **61(C)**: pp. 133–151.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-Analysis." *American Economic Review* **85(2)**: pp. 238–243.

Cazachevici, A., T. Havranek, & R. Horvath (2020): "Remittances and Economic Growth: A Meta-Analysis." *World Development* **134**: p. 105021.

Chirinko, R. S. (2002): "Corporate Taxation, Capital Formation, and the Substitution Elasticity between Labor and Capital." *National Tax Journal* **55(2)**: pp. 339–355.

Chirinko, R. S. (2008): "σ: The Long and Short of it." *Journal of Macroeconomics* **30(2)**: pp. 671–686.

Chirinko, R. S., S. M. Fazzari, & A. P. Meyer (2011): "A New Approach to Estimating Production Func-

tion Parameters: The Elusive Capital–Labor Substitution Elasticity." *Journal of Business & Economic Statistics* **29(4)**: pp. 587–594.

Chirinko, R. S. & D. Mallick (2017): "The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model." *American Economic Journal: Macroeconomics* **9(4)**: pp. 225–53.

Christensen, G. & E. Miguel (2018): "Transparency, Reproducibility, and the Credibility of Economics Research." *Journal of Economic Literature* **56(3)**: pp. 920–980.

Chwelos, P., R. Ramirez, K. L. Kraemer, & N. P. Melville (2010): "Does Technological Progress Alter the Nature of Information Technology as a Production Input? New Evidence and New Results." *Information Systems Research* **21(2)**: pp. 392–408.

Cochrane, J. H. (2018): "Michelson-Morley, Fisher, and Occam: The Radical Implications of Stable Quiet Inflation at the Zero Bound." *NBER Macroeconomics Annual* **32(1)**: pp. 113–226.

Cummins, J. G. & K. A. Hassett (1992): "The Effects of Taxation on Investment: New Evidence from Firm Level Panel Data." *National Tax Journal* **45(3)**: pp. 243–251.

DeLong, J. B. & K. Lang (1992): "Are All Economic Hypotheses False?" *Journal of Political Economy* **100(6)**: pp. 1257–72.

Dhrymes, P. J. (1965): "Some Extensions and Tests for the CES Class of Production Functions." *Review of Economics and Statistics* **47(4)**: pp. 357–366.

Diamond, P., D. McFadden, & M. Rodriguez (1978): "Measurement of the Elasticity of Factor Substitution and Bias of Technical Change." In M. Fuss & D. McFadden (editors), "Contributions to Economic Analysis," volume 2, pp. 125–147. Elsevier.

Dissou, Y., L. Karnizova, & Q. Sun (2015): "Industry-Level Econometric Estimates of Energy-Capital-Labor Substitution with a Nested CES Production Function." *Atlantic Economic Journal* **43(1)**: pp. 107–121.

Doucouliagos, C. & T. D. Stanley (2013): "Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

Duffy, J. & C. Papageorgiou (2000): "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification." *Journal of Economic Growth* **5(1)**: pp. 87–120.

Eden, M. & P. Gaggl (2015): "On the welfare implications of automation." *Policy Research Working Paper 7487*, The World Bank.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Eisner, R. (1969): "Tax Policy and Investment Behavior: Comment." *American Economic Review* **59(3)**: pp. 379–388.

Eisner, R. & M. I. Nadiri (1968): "Investment Behavior and Neo-Classical Theory." *Review of Economics and Statistics* **50(3)**: pp. 369–382.

Erceg, C. J., L. Guerrieri, & C. Gust (2008): "Trade Adjustment and the Composition of Trade." *Journal of Economic Dynamics and Control* **32(8)**: pp. 2622–2650.

Ferguson, C. E. (1965): "Time-Series Production Functions and Technological Progress in American Manufacturing Industry." *Journal of Political Economy* **73(2)**: pp. 135–147.

Furukawa, C. (2019): "Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method." *Unpublished paper*, MIT.

Gerber, A. & N. Malhotra (2008): "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* **3(3)**: pp. 313–326.

Glover, A. & J. Short (2020a): "Can capital deepening explain the global decline in labor's share?" *Review of Economic Dynamics* **35**: pp. 35–53.

Glover, A. & J. Short (2020b): "Demographic Origins of the Decline in Labor's Share." *BIS working papers no. 874*, Bank for International Settlements.

Grossman, G. M., E. Helpman, E. Oberfield, & T. Sampson (2017): "The productivity slowdown and the declining labor share: A neoclassical exploration." *NBER Working Papers 23853*, National Bureau of Economic Research.

Hall, R. E. & D. W. Jorgenson (1967): "Tax Policy and Investment Behavior." *American Economic Review* **57(3)**: pp. 391–414.

Hansen, B. E. (2007): "Least Squares Model Averaging." *Econometrica* **75(4)**: pp. 1175–1189.

Havranek, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

Havranek, T., D. Herman, , & Z. Irsova (2018a): "Does Daylight Saving Save Electricity? A Meta-Analysis." *The Energy Journal* **39(2)**: pp. 35–61.

Havranek, T., R. Horvath, Z. Irsova, & M. Rusnak (2015): "Cross-Country Heterogeneity in Intertemporal Substitution." *Journal of International Economics* **96(1)**: pp. 100–118.

Havranek, T. & Z. Irsova (2017): "Do Borders Really Slash Trade? A Meta-Analysis." *IMF Economic Review* **65(2)**: pp. 365–396.

Havranek, T., Z. Irsova, & T. Vlach (2018b): "Measuring the Income Elasticity of Water Demand: The Importance of Publication and Endogeneity Biases." *Land Economics* **94(2)**: pp. 259–283.

Havranek, T., Z. Irsova, & O. Zeynalova (2018c): "Tuition Fees and University Enrolment: A Meta-

Regression Analysis." *Oxford Bulletin of Economics and Statistics* **80(6)**: pp. 1145–1184.

Havranek, T., M. Rusnak, & A. Sokolova (2017): "Habit formation in consumption: A meta-analysis." *European Economic Review* **95**: pp. 142–167.

Hedges, L. V. (1992): "Modeling Publication Selection Effects in Meta-Analysis." *Statistical Science* **7(2)**: pp. 246–255.

Hicks, J. R. (1932): "Marginal Productivity and the Principle of Variation." *Economica* **(35)**: pp. 79–88.

Humphrey, D. B. & J. R. Moroney (1975): "Substitution among Capital, Labor, and Natural Resource Products in American Manufacturing." *Journal of Political Economy* **83(1)**: pp. 57–82.

Imai, T., T. A. Rutter, & C. F. Camerer (2021): "Meta-Analysis of Present-Bias Estimation using Convex Time Budgets." *Economic Journal* **(forthcoming)**.

Ioannidis, J. P., T. D. Stanley, & H. Doucouliagos (2017): "The Power of Bias in Economics Research." *Economic Journal* **127(605)**: pp. 236–265.

Jorgenson, D. W. (1963): "Capital Theory and Investment Behavior." *American Economic Review* **53(2)**: pp. 247–259.

Jorgenson, D. W. (2007): "35 Sector KLEM." Harvard Dataverse.

Karabarbounis, L. & B. Neiman (2014): "The global decline of the labor share." *The Quarterly Journal of Economics* **129(1)**: pp. 61–103.

Kilponen, J., S. Orjasniemi, A. Ripatti, & F. Verona (2016): "The Aino 2.0 Model." *Bank of Finland Research Discussion Paper 16*, Bank of Finland.

Klump, R. & O. de La Grandville (2000): "Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions." *American Economic Review* **90(1)**: pp. 282–291.

Klump, R., P. McAdam, & A. Willman (2007): "Factor Substitution and Factor-Augmenting Technical Progress in the United States: A Normalized Supply-Side System Approach." *Review of Economics and Statistics* **89(1)**: pp. 183–192.

Klump, R., P. McAdam, & A. Willman (2012): "The Normalized CES Production Function: Theory and Empirics." *Journal of Economic Surveys* **26(5)**: pp. 769–799.

Kmenta, J. (1967): "On Estimation of the CES Production Function." *International Economic Review* **8(2)**: pp. 180–189.

Knoblach, M., M. Rossler, & P. Zwerschke (2020): "The Elasticity of Substitution Between Capital and Labour in the US Economy: A Meta-Regression Analysis." *Oxford Bulletin of Economic and Statistics* **82(1)**: pp. 62–82.

Koh, D., R. Santaeulàlia-Llopis, & Y. Zheng (2016): "Labor Share Decline and Intellectual Property Products Capital." *Working Papers 927*, Barcelona Graduate School of Economics.

Krusell, P., L. E. Ohanian, J.-V. Roos-Rull, & G. L. Violante (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* **68(5)**: pp. 1029–1053.

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

León-Ledesma, M. A., P. McAdam, & A. Willman (2010): "Identifying the Elasticity of Substitution with Biased Technical Change." *American Economic Review* **100(4)**: pp. 1330–57.

Lovell, C. K. (1973): "CES and VES Production Functions in a Cross-Section Context." *Journal of Political Economy* **81(3)**: pp. 705–720.

Madigan, D. & J. York (1995): "Graphical Models for Discrete Data." *International Statistical Review* **63(2)**: pp. 215–232.

Matousek, J., T. Havranek, & Z. Irsova (2021): "Individual Discount Rates: A Meta-Analysis of Experimental Evidence." *Experimental Economics* **(forthcoming)**.

McCloskey, D. N. & S. T. Ziliak (2019): "What Quantitative Methods Should We Teach to Graduate Students? A Comment on Swann's Is Precise Econometrics an Illusion?" *Journal of Economic Eduation* **50(4)**: pp. 356–361.

Nekarda, C. J. & V. A. Ramey (2013): "The Cyclical Behavior of the Price-Cost Markup." *NBER Working Paper 19099*, National Bureau of Economic Research.

Nerlove, M. (1967): "Recent Empirical Studies of the CES and Related Production Functions." In M. Brown (editor), "The Theory and Empirical Analysis of Production," pp. 55–136. NBER.

Oberfield, E. & D. Raval (2014): "Micro data and macro technology." *NBER Working Paper 20452*, National Bureau of Economic Research.

Piketty, T. (2014): *Capital in the 21st Century*. Cambridge, MA: Harvard University Press.

Ramey, V. (2016): "Macroeconomic Shocks and Their Propagation." In J. B. Taylor & H. Uhlig (editors), "Handbook of Macroeconomics," volume 2 of *Handbook of Macroeconomics*, chapter 1, pp. 71–162. Elsevier.

Robinson, J. (1933): *Economics of Imperfect Competition*. New York: The Macmillan Company.

Rognlie, M. (2014): "A Note on Piketty and Diminishing Returns to Capital." *Unpublished paper*, MIT.

Roodman, D. (2019): "BOOTTEST: Stata Module to Provide Fast Execution of the Wild Bootstrap with Null Imposed." Statistical Software Components from Boston College Department of Economics.

Rusnak, M., T. Havranek, & R. Horvath (2013): "How to Solve the Price Puzzle? A Meta-Analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

Semieniuk, G. (2017): "Piketty's Elasticity of Substitution: A Critique." *Review of Political Economy* **29(1)**: pp. 64–79.

Smith, J. (2008): "That Elusive Elasticity and the Ubiquitous Bias: Is Panel Data a Panacea?" *Journal of Macroeconomics* **30(2)**: pp. 760–779.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* **70(1)**: pp. 65–94.

Stanley, T., S. B. Jarrell, & H. Doucouliagos (2010): "Could it Be Better to Discard 90% of the Data? A Statistical Paradox." *American Statistician* **64(1)**: pp. 70–77.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Stanley, T. D. & H. Doucouliagos (2010): "Picture This: A Simple Graph that Reveals Much Ado about Research." *Journal of Economic Surveys* **24(1)**: pp. 170–191.

Stanley, T. D. & H. Doucouliagos (2017): "Neither Fixed nor Random: Weighted Least Squares Meta-Regression." *Research Synthesis Methods* **8(1)**: pp. 19–42.

Steel, M. F. J. (2020): "Model Averaging and its Use in Economics." *Journal of Economic Literature* **58(3)**: pp. 644–719.

Thursby, J. G. & C. K. Lovell (1978): "An Investigation of the Kmenta Approximation to the CES Function." *International Economic Review* **19(2)**: pp. 363–377.

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)**: pp. 1765–1785.

Van der Werf, E. (2008): "Production Functions for Climate Policy Modeling: An Empirical Analysis." *Energy Economics* **30(6)**: pp. 2964–2979.

Young, A. T. (2013): "US Elasticities of Substitution and Factor Augmentation at the Industry Level." *Macroeconomic Dynamics* **17(4)**: pp. 861–897.

Zarembka, P. (1970): "On the Empirical Relevance of the CES Production Function." *Review of Economics and Statistics* **52(1)**: pp. 47–53.

Zigraiova, D. & T. Havranek (2016): "Bank Competition And Financial Stability: Much Ado About Nothing?" *Journal of Economic Surveys* **30(5)**: pp. 944–981.

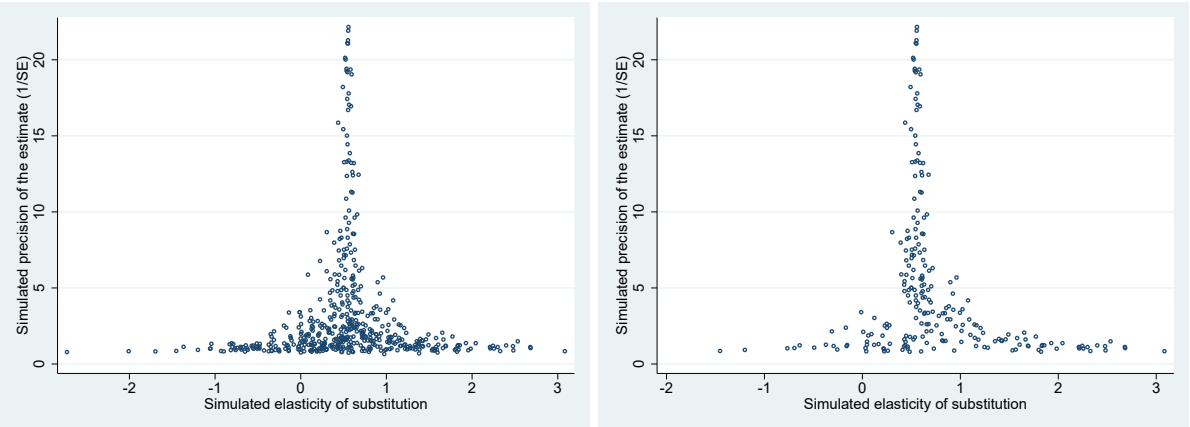
Appendix A Illustrating the Effects of Publication Bias in a Monte Carlo Simulation

The impact of publication bias on a naive literature summary and on the working of benchmark meta-analysis tools can be shown in a simple Monte Carlo simulation. We illustrate what happens to the mean of the reported elasticities if some estimates are systematically underreported. To this end we employ the central estimate of the elasticity from Antras (2004), a representative and well-cited study with a point estimate of $\hat{\sigma} = 0.551$ stemming from a specification with FOC on capital, allowing for biased technological change, and relying on US macroeconomic time series data. The estimation equation reads:

$$log(Y_t/K_t) = \alpha + \sigma log(R_t/P_t^Y) + (1-\sigma)\lambda_K \cdot t + \varepsilon_t. \tag{15}$$

For the Monte Carlo simulation we assume this estimate to be close to the unbiased true underlying value of σ. (Our results would be qualitatively the same if we chose a different study for the simulation.) We set up our data generating process by re-estimating $\hat{\sigma}$ from 500 draws of the Antras (2004) data by adding noise to the dependent variable $log(Y_t/K_t)$ (with a sample mean of 4.16) via a random error from a Gaussian distribution with $X \sim N(0, Var)$. In order to generate the familiar funnel shape for the scatter plot of estimates and standard errors, the variance Var of the noise term X is chosen not to be constant across draws but to vary from 0.0016 to 0.8. Note that the qualitative results of the simulation are independent of this specific parametrization. The funnel would still display a range comparable to our actual dataset shown in Figure 6, though it would look less pretty. The funnel plot from our 500 simulated estimations (noisy versions of Antras's model) is displayed on the left-hand side of Figure A1. It has an average $\hat{\sigma} = 0.534$ with a standard deviation equal to 0.685.

Figure A1: Simulated funnel plots without and with publication bias



Notes: In the absence of publication bias the scatter plot should resemble an inverted funnel symmetrical around the most precise estimates. The left panel shows estimates from all 500 Monte Carlo draws obtained from the replication of the estimate in Antras (2004) (Table 5, Column I, Row 1) and by adding random noise to the dependent variable, thereby producing a symmetric funnel around Antras's estimate. The right panel shows what happens to the funnel plot if 80% of estimates that are negative or insignificantly different from zero (at a 5% level) are discarded, which results in retaining only 227 observations.

The right-hand panel of Figure A1 shows how the funnel would change if we filtered out 80% of the simulated estimates that are either negative or insignificantly different from zero. This setup reflects a typical publication bias scenario in which significant and theory-compliant estimates are more likely to be reported. In this scenario, only 227 observations are left, and the funnel becomes asymmetric. In fact, however, it is less asymmetric that the actual funnel plot we observe in the literature (Figure 6), indicating that publication bias may be even more severe in practice than with the aforementioned filter. The filtered simulated dataset represents what a reviewer of the literature observes. Publication bias drives the observed average elasticity upwards from 0.534 to 0.743 and produces a correlation between point estimates and their standard errors, a correlation that was not present before (column 1 in Table A1).

Table A1 shows a funnel asymmetry test, a regression of estimated elasticities on the corresponding standard errors (as explained in Section 4) for different scenarios of bias. Column 1 refers to the unbiased symmetric funnel in Figure A1. The test indicates no bias, and the estimated mean beyond bias is close to the true mean. If all negative estimates are dropped (column 2), the naive mean increases to 0.726. The test detects publication bias and uncovers a mean of 0.521, close to the true one. Column 3 refers to the asymmetric funnel in the right-hand panel of Figure A1. Again the test detects publication bias and estimates the true mean fairly precisely. Columns 4 and 5 show that the working of the test does not hinge on the selection threshold of zero. If for example the Cobb-Douglas specification with $\sigma = 1$ serves as a benchmark for researchers, in the way that they discard 80% of all estimates that are significantly different from 1 at a 5% level, the mean of the reported estimates would also be biased upwards and meta-analysis tests again do a good job in detecting the bias. Even for the extreme example of column 5, where we drop 80% of estimates with $\sigma < 1.3$ and the uncorrected mean increases to 0.913, the funnel asymmetry test estimates the underlying true σ well.

Table A1: Monte Carlo simulation of publication bias

	(1) no filter	(2) drop < 0	(3) drop 80% of < 0 or insignif $\neq 0$	(4) drop 80% of signif $\neq 1$	(5) drop 80% of < 1.3
$\bar{\sigma}$ (mean)	0.534	0.726	0.743	0.616	0.913
SE (publication bias)	-0.0160	0.313^{***}	0.499^{***}	0.135^{**}	0.578^{***}
	(0.0534)	(0.0492)	(0.0956)	(0.0567)	(0.102)
Const (mean beyond bias)	0.548^{***}	0.521^{***}	0.515^{***}	0.550^{***}	0.488^{***}
	(0.00727)	(0.00716)	(0.00922)	(0.0145)	(0.0133)
Observations	500	423	227	391	151

Notes: The table shows detection of and correction for publication bias in five different scenarios. (1) Reporting all estimates. (2) Dropping all negative estimates of σ. (3) Dropping 80% of negative or insignificant (at the 5% threshold) estimates. (4) Dropping 80% of estimates that are significantly different from $\sigma = 1$ at the 5% level. (5) Dropping 80% of estimates that are smaller than $\sigma = 1.3$. The original data were obtained from Antras (2004), the specification FOC_K with trend from Table 5.1, Col I, Row 1. The Monte Carlo simulation adds noise to the dependent variable and estimates Antras's model 500 times. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level. Standard errors in parentheses.

Appendix B Furukawa's Method for Addressing Selective Reporting (for Online Publication)

Appendices B, C, D, E, and F are only presented here for the convenience of reviewers. If the manuscript is accepted for publication, this material will be relegated to an online appendix.

Furukawa (2019) proposes the so-called stem-based correction method, which relies on the most precise studies, corresponding to the stem of the funnel plot. The method is nonparametric, fully data-dependent and requires weaker assumptions for the underlying distribution of true effects and the publication selection process than other methods. Publication selection can be a function of the size of the estimates, their significance, or both at the same time, as imprecise null results are less likely to be published. By focusing on the n most precise estimates, Furukawa (2019) is able to account for various publication selection processes. The method extends the approach by Stanley *et al.* (2010), who suggest using 10% of the most precise estimates. Instead of selecting an arbitrary number of the most precise estimates, Furukawa (2019) suggests a formal method to calculate the optimal number n of the most precise studies to include by minimizing the mean squared error:

$$\min_n MSE(n) = Bias^2(n) + Var(n). \tag{16}$$

With more studies used, the squared bias term increases as less precise studies suffer from more bias, but the variance term decreases as more information increases efficiency. An empirical analog of the bias term is estimated nonparametrically using two algorithms. The inner algorithm computes the bias-corrected mean given an assumed value of squared precision, and the outer algorithm computes the implied variance and ensures that it is consistent with its assumed value. The inner algorithm ranks and indexes studies in an ascending order according to their standard error, se, and for each $n = 2, ..., N$ calculates the relevant bias squared and variance, given the assumed value of se_0:

$$\tilde{Bias}^2(n) = \frac{\sum_{i=2}^{n}\sum_{j\neq i}^{n} w_i w_j \beta_i \beta_j}{\sum_{i=2}^{n}\sum_{j\neq i}^{n} w_i w_j} - 2\beta_1 \frac{\sum_{i=2}^{n} w_i \beta_i}{\sum_{i=2}^{n} w_i}, \tag{17}$$

$$Var(n) = \sum_{i=1}^{n} w_i, \tag{18}$$

where $w_i = \frac{1}{se_i^2 + se_0^2}$. The optimal number of included studies is given by Equation 16. The stem-based corrected estimate follows:

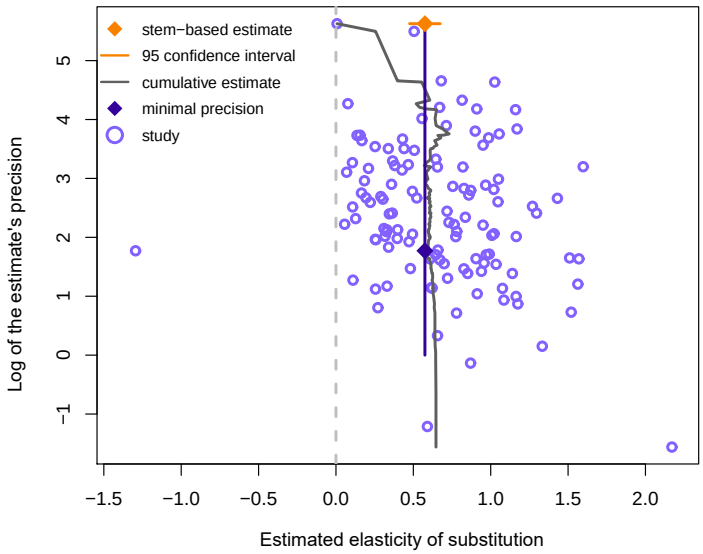
$$\hat{b}_{stem} = \frac{\sum_{i=1}^{n_{stem}} w_i \beta_i}{\sum_{i=1}^{n_{stem}} w_i}. \tag{19}$$

The outer algorithm then searches over se_0^2 so that the implied variance is consistent.

The stem-based method applied to the elasticity of substitution yields the following results: the mean underlying elasticity corrected for publication bias is 0.57 with a standard error of

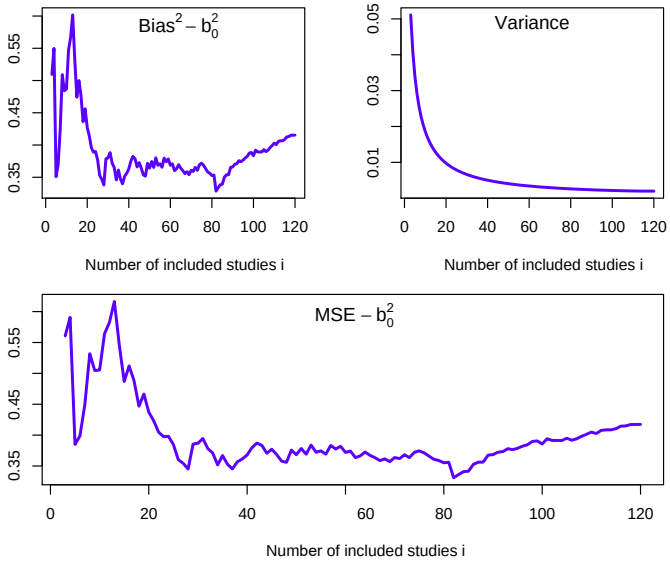
0.05. Overall, 77% of the total information in the data is utilized, and the 83 most precise studies (out of 121) are included. Because the stem-based method uses study-level estimates (as preferred by Furukawa), we select median values from each study. Figure B1 visualizes the stem-based bias correction method. Figure B2 visualizes the bias-variance trade-off in order to minimize the mean squared error. When all estimates instead of median estimates are used, the mean corrected elasticity is similar, 0.55, but the standard error increases to 0.21.

Figure B1: A graphical illustration of Furukawa's technique



Note: The orange (lighter in grayscale) diamond at the top corresponds to the stem-based estimate of the mean elasticity corrected for publication bias, with the orange line indicating the corresponding 95% confidence interval. The gray (lighter in grayscale) line denotes the estimate under various $n_{stem} \in 1, ..., N$. The blue (darker in grayscale) diamond indicates the minimum precision level that defines the "stem" of the funnel.

Figure B2: The trade-off between bias and variance



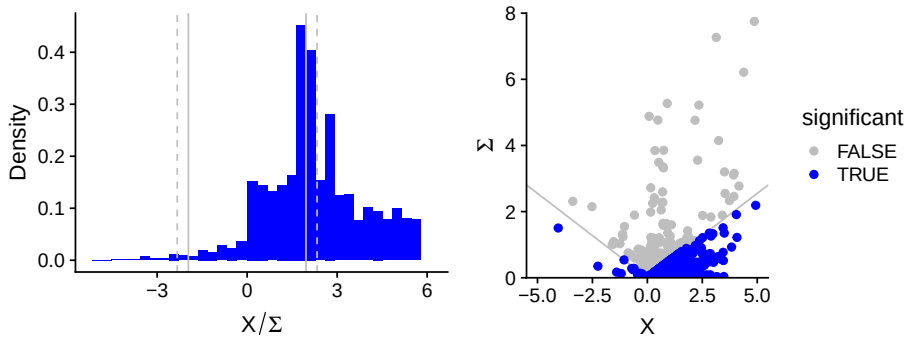
Note: The mean squared error (MSE) is the criterion for choosing the n_{stem}, the optimal number of studies to include in the stem-based estimator. The relevant components of MSE—bias and variance—are plotted.

Appendix C Andrews and Kasy's Method for Addressing Selective Reporting (for Online Publication)

Andrews & Kasy (2019) introduce two approaches for the identification of publication selection: the first one based on data from replication studies and the second one tailored for meta-analysis. They show that the meta-analysis approach delivers results similar to the approach based on replications. In the absence of publication bias, the distribution of the estimates from imprecise studies can be written as the distribution for precise studies plus noise; deviations from this form identify conditional publication probabilities. Andrews & Kasy (2019) identify publication probability similarly to Hedges (1992) using maximum likelihood: conditional publication probability, $p(\cdot)$, is a step function with jumps at conventional critical values of the p-value.

When applied to our data, the method by Andrews & Kasy (2019) yields the following results. The bias-corrected estimate is 0.43 with a standard error of 0.017. We impose a cutoff at zero, that is, we compare the publication probability of negative vs. positive estimates regardless of their significance. (Allowing for other jumps in publication probability would yield even smaller estimates of the mean elasticity corrected for publication bias.) Our results also suggest that positive estimates are six times more likely to be selected for publication than negative estimates (Table C1). In the case of the elasticity of substitution, publication selection based on statistical significance is apparently less pronounced than selection based on the sign of the estimate, as suggested by the right panel of Figure C1.

Figure C1: A graphical illustration of Andrews and Kasy's (2019) estimator



Note: The solid gray lines mark t-statistic equal to 1.96 in absolute value; the dashed gray line marks t-statistics equal to 2.33 in absolute value. We observe a jump at t-statistic equal to zero and then also jumps at conventional significance levels. The right-hand figure plots estimates X and their standard errors Σ; the gray line marks 1.96 in absolute value. Even though we observe discontinuity at the t-statistic corresponding to the 5% significance level, the right panels shows publication selection based on significance is not absolute, as some insignificant estimates (gray points) are reported.

Table C1: Results of Andrews and Kasy's (2019) estimator

	$\bar{\theta}$	$\bar{\tau}$	DF	β_p
Estimate	0.430	0.489	12.809	0.158
Standard error	0.017	0.012	0.707	0.019

Notes: $\bar{\theta}$ denotes the bias-corrected mean effect, $\bar{\tau}$ is a scale parameter, DF are degrees of freedom. β_p is a publication probability measured relative to the omitted category, in our case positive estimates. An estimate of 0.158 therefore implies that negative results are 15.8% as likely to be published as positive ones.

Appendix D Description of Variables (for Online Publication)

Table D1: Definitions and summary statistics of explanatory variables

Variable	Description	Mean	Std. dev.
Data characteristics			
No. of obs.	The logarithm of the number of observations used in the regression.	4.28	1.51
Midpoint	The logarithm of the mean year of the data used minus the earliest mean year in our data.	4.71	0.48
Cross-sec.	= 1 if cross-sectional data are used (reference category: time series).	0.33	0.47
Panel	= 1 if panel data are used (reference category: time series).	0.14	0.35
Quarterly	= 1 if the data frequency is quarterly (reference category: annual).	0.11	0.31
Industry data	= 1 if variation at the industry-/sector-level is exploited in input data (reference category: cross-country-/state-level variation).	0.43	0.50
Firm data	= 1 if variation at the firm-level is exploited in input data (reference category: cross-country-/state-level variation).	0.12	0.32
Country: US	= 1 if the estimate is for the US.	0.58	0.49
Country: Eur	= 1 if the estimate is for a developed European country.	0.17	0.37
Developing	= 2 if the estimate is for a developing country; = 1 if the estimate is a common estimate for a collection of developed and developing countries (reference category: developed countries).	0.22	0.54
Database: OECD	= 1 if the data come from the OECD database.	0.07	0.25
Database: KLEM	= 1 if the data come from the Jorgenson KLEM dataset.	0.15	0.36
Database: ASMCM	= 1 if the data come from the Annual Survey of Manufacturers and/or Census of Manufacturers.	0.14	0.35
Disaggregated σ	= 1 if the elasticity is estimated on a disaggregated level (industry-specific elasticity).	0.52	0.50
Specification			
System PF-FOC	= 1 if the elasticity is estimated within a system of CES with FOC(s) or with cost share functions.	0.06	0.23
System FOCs	= 1 if the elasticity is estimated within a system of FOCs.	0.05	0.23
Nonlinear	= 1 if the elasticity is estimated within the CES directly via nonlinear methods.	0.04	0.20
Linear approx.	= 1 if the elasticity is estimated via Taylor series expansion (Kmenta approach or translog approach).	0.07	0.26
FOC_L_w	= 1 if the elasticity is estimated within the FOC for labor based on the wage rate (reference category: FOC for capital based on the rental rate of capital).	0.33	0.47
FOC_KL_rw	= 1 if the elasticity is estimated within the FOC of K/L based on w/r (reference category: FOC for capital based on the rental rate of capital).	0.18	0.39
FOC_K_share	= 1 if the elasticity is estimated within the FOC for capital based on the capital share (reference category: FOC for capital based on the rental rate of capital).	0.03	0.16
FOC_L_share	= 1 if the elasticity is estimated within the FOC for labor based on the labor share (reference category: FOC for capital based on the rental rate of capital).	0.04	0.19
User cost elast.	= 1 if the user cost of capital elasticity is estimated.	0.17	0.38
Cross-equation rest.	= 1 if cross-equation restrictions are employed when using system estimation.	0.08	0.28
Normalized	= 1 if normalization is applied to the CES.	0.05	0.22

Continued on next page

Table D1: Definitions and summary statistics of explanatory variables (continued)

Variable	Description	Mean	Std. dev.
Two-level PF	= 1 if a two-level CES function is estimated (due to more than two factors of production).	0.03	0.18
Partial σ	= 1 if some form of partial elasticity is used (Allen-Uzawa, Hicks-Allen, Morishima).	0.06	0.24
Econometric approach			
Dynamic est.	= 1 if dynamic methods are used for estimation (VAR, a distributed lag model or error correction model; reference category: OLS).	0.24	0.42
SUR	= 1 if a system of seemingly unrelated regressions is used (Zellner's estimation; reference category: OLS).	0.11	0.31
Differenced	= 1 if the coefficient is taken from a regression in first differences or log differences.	0.23	0.42
Time FE	= 1 if time-fixed effects are used for estimation.	0.06	0.24
Unit FE	= 1 if unit-fixed effects are used for estimation.	0.04	0.20
Identification	= 1 if instrumental variables are used for identification.	0.13	0.34
Short-run σ	= 1 if the coefficient is taken from an explicit short-run specification (reference category: explicit long-run specification—cointegration, low-pass filter, interval-difference model).	0.05	0.22
Long-run σ unadj.	= 1 if the coefficient is meant to be long-run but the specification is not adjusted accordingly (reference category: explicit long-run specification).	0.68	0.47
Production function components			
Other inputs in PF	= 1 if the production function includes other inputs such as energy, materials, and human capital.	0.13	0.34
LATC	= 1 if the production function includes labor-augmenting technological change, i.e. Harrod-neutral technological change (reference category: Hicks-neutral technological change).	0.29	0.63
CATC	= 1 if the production function includes capital-augmenting technological change, i.e. Solow-neutral technological change (reference category: Hicks-neutral technological change).	0.26	0.57
Skilled L	= 1 if the production function distinguishes between skilled and unskilled labor.	0.02	0.13
Constant TC growth	= 1 if the technological change is modeled with constant growth rates (reference category: no growth of technology).	0.30	0.46
Other TC growth	= 1 if the technological change is modeled with nonconstant growth rates, e.g., logarithmic, linear (reference category: no growth of technology).	0.10	0.31
No CRS	= 1 if the authors assume nonconstant returns to scale.	0.09	0.36
No full comp.	= 1 if the authors do not assume factor markets to be perfectly competitive.	0.04	0.19
Net σ	= 1 if net elasticity is estimated (reference category: gross elasticity).	0.02	0.16
External info			
Top journal	= 1 if the study is published in a top five journal in economics.	0.31	0.46
Pub. year	The logarithm of the year when the first draft of the study appeared in Google Scholar minus the year when the first study on elasticity of substitution was written.	3.25	0.88
Impact	The recursive discounted RePEc impact factor of the outlet.	0.96	1.07
Citations	The logarithm of the number of per-year citations of the study since its first appearance on Google Scholar.	1.47	0.96
Preferred est.	= 1 if the estimate is preferred by authors or is explicitly considered to be better; -1 if it is considered inferior.	-0.04	0.47
Byproduct	= 1 if estimation of the elasticity is not the central focus of the paper but only a byproduct; = 0 if it is the central focus; = 0.5 if it is one of multiple main aims.	0.20	0.31

Continued on next page

Table D1: Definitions and summary statistics of explanatory variables (continued)

Variable	Description	Mean	Std. dev.
Measurement of variables			
y: index	= 1 if the input data for total output is in an index form.	0.03	0.18
y: other	= 1 if the input data for total output is measured differently than in gross domestic product or total value added (reference category: GDP, value added).	0.07	0.26
Labor-related			
Quality adj.	= 1 if the input data for labor incomes data are quality-adjusted.	0.22	0.41
Self empl.	= 1 if the input data for labor incomes data are adjusted for the income of self-employed people.	0.18	0.39
w: nominal	= 1 if the input data for the wage rate are nominal (reference category: the wage rate is in real terms).	0.09	0.29
w: direct	= 1 if the input data for the wage rate are measured directly (the wage rate calculated as total wages divided by the total number of employees).	0.14	0.36
L: hours	= 1 if the input data for the labor are measured in hours.	0.25	0.44
L: years	= 1 if the input data for the labor are measured in years.	0.07	0.25
L: FTE workers	= 1 if the input data for labor are measured by the full-time equivalent number of workers.	0.07	0.25
L: force	= 1 if number of workers labor is measured as the total number of people in the labor force.	0.04	0.20
Capital-related			
Capacity adj.	= 1 if the authors control for the capacity utilization in the regression.	0.09	0.28
r: quasi	= 1 if the input data for the rental rate of capital are measured as the quasi-rent, i.e., total output minus total wages divided by total capital stock (reference category: it is measured as the user cost of capital, Hall-Jorgenson formula).	0.24	0.43
r: nominal	= 1 if the input data for the rental rate of capital are expressed in nominal terms.	0.01	0.09
K: IT	= 1 if IT capital is used only.	0.02	0.13
K: equipment	= 1 if the measure of equipment capital is used only.	0.07	0.26
K: structures	= 1 if the measure of structures, land or plant is used only.	0.04	0.17
K: residential	= 1 if the measure of capital includes residential capital stock.	0.07	0.25
K: services	= 1 if capital is measured as service flow.	0.13	0.33
K: perpetual	= 1 if the input data for capital is measured via perpetual inventory method.	0.36	0.48
K: index	= 1 if the input data for capital are expressed in an index form.	0.17	0.37
Industry-related			
Primary ind.	= 1 if the elasticity is estimated for the primary sector.	0.02	0.14
Secondary ind.	= 1 if the elasticity is estimated for the secondary sector.	0.62	0.49
Tertiary ind.	= 1 if the elasticity is estimated for the tertiary sector.	0.03	0.18
Materials	= 1 if the elasticity is estimated for the 2-digit industry in the category "Materials" of the GICS industry classification.	0.25	0.43
Industrials	= 1 if the elasticity is estimated for the 2-digit industry in the category "Industrials" of the GICS industry classification.	0.09	0.29
Consumer	= 1 if the elasticity is estimated for the 2-digit industry in the category "Consumer goods" of the GICS industry classification.	0.14	0.34

Note: Collected from published studies estimating the elasticity of substitution between capital and labor. When dummy variables form groups, we mention the reference category.

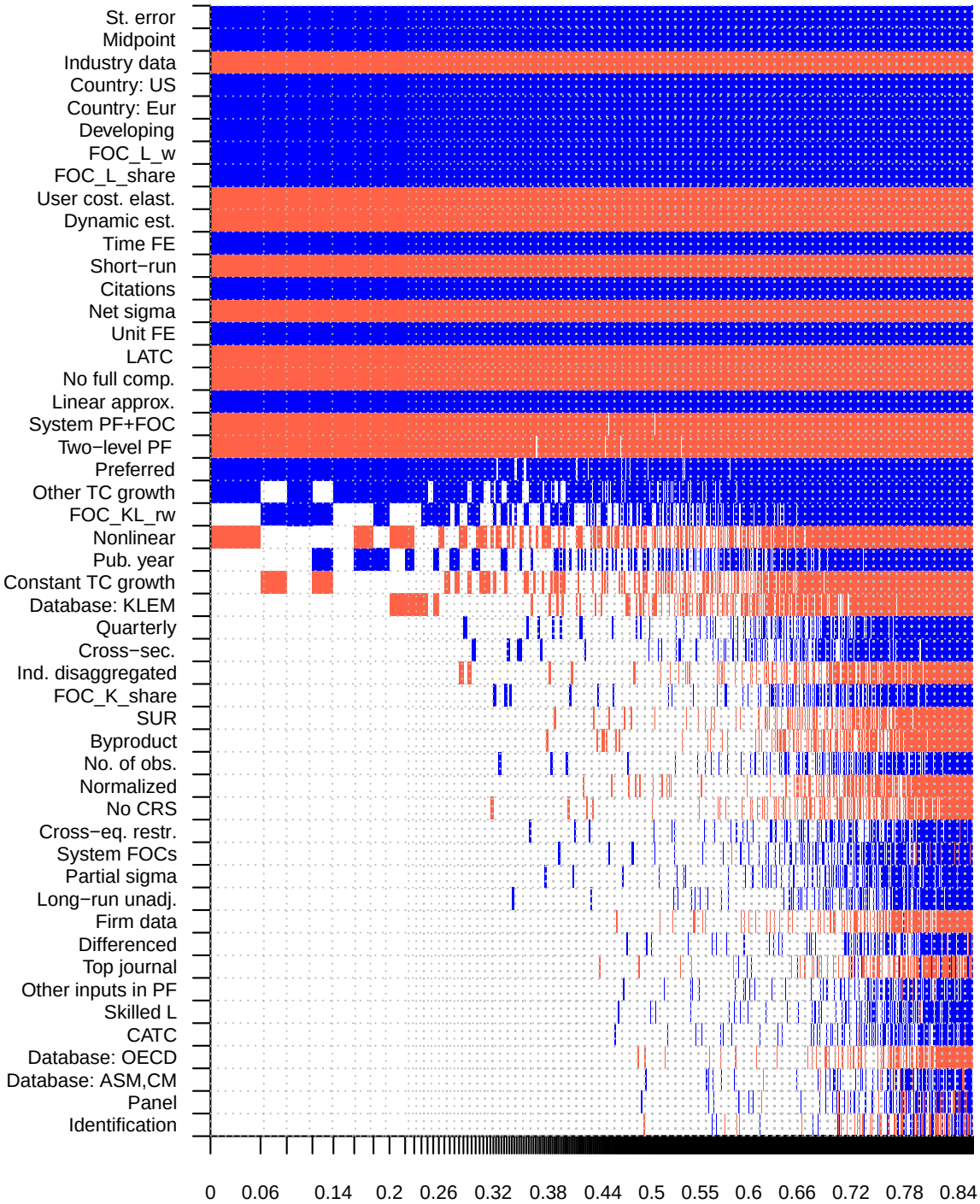
Appendix E Robustness Checks (for Online Publication)

Table E1: Results of frequentist model averaging

	Coef.	Std. er.	p-value
Standard error	0.557	0.042	0.000
Data characteristics			
No. of obs.	0.011	0.012	0.326
Midpoint	0.103	0.022	0.000
Cross-sec.	0.069	0.029	0.016
Panel	0.193	0.042	0.000
Quarterly	0.135	0.042	0.001
Firm data	-0.160	0.040	0.000
Industry data	-0.198	0.026	0.000
Country: US	0.121	0.031	0.000
Country: Eur	0.180	0.030	0.000
Developing	0.019	0.019	0.333
Database: ASM,CM	-0.031	0.037	0.402
Database: OECD	-0.301	0.044	0.000
Database: KLEM	-0.092	0.046	0.047
Disaggregated σ	0.043	0.024	0.077
Specification			
System PF+FOC	-0.111	0.059	0.061
System FOCs	-0.057	0.050	0.258
Nonlinear	-0.016	0.061	0.796
Linear approx.	0.268	0.050	0.000
FOC_L_w	0.324	0.032	0.000
FOC_KL_rw	0.007	0.032	0.832
FOC_K_share	0.226	0.063	0.000
FOC_L_share	0.251	0.048	0.000
Cross-eq. restr.	0.071	0.048	0.140
Normalized	-0.248	0.051	0.000
Two-level PF	-0.023	0.070	0.743
Partial sigma	0.130	0.055	0.018
User cost. elast.	-0.373	0.042	0.000
Econometric approach			
Dynamic est.	-0.005	0.029	0.854
SUR	-0.105	0.032	0.001
Identification	0.046	0.026	0.077
Differenced	-0.096	0.027	0.000
Time FE	-0.009	0.040	0.830
Unit FE	0.067	0.043	0.116
Short-run	-0.410	0.040	0.000
Long-run unadj.	-0.011	0.026	0.681
Production function components			
Other inputs in PF	-0.137	0.044	0.002
CATC	-0.003	0.026	0.904
LATC	-0.041	0.024	0.088
Skilled L	0.076	0.059	0.199
Constant TC growth	-0.032	0.025	0.191
Other TC growth	0.108	0.035	0.002
No CRS	-0.003	0.022	0.905
No full comp.	-0.022	0.042	0.598
Net sigma	-0.320	0.056	0.000
Publication characteristics			
Top journal	-0.085	0.025	0.001
Pub. year	0.032	0.015	0.038
Citations	0.037	0.011	0.001
Preferred	0.027	0.016	0.093
Byproduct	-0.130	0.032	0.000
(Intercept)	-0.123	0.130	0.342
Observations	3,186		

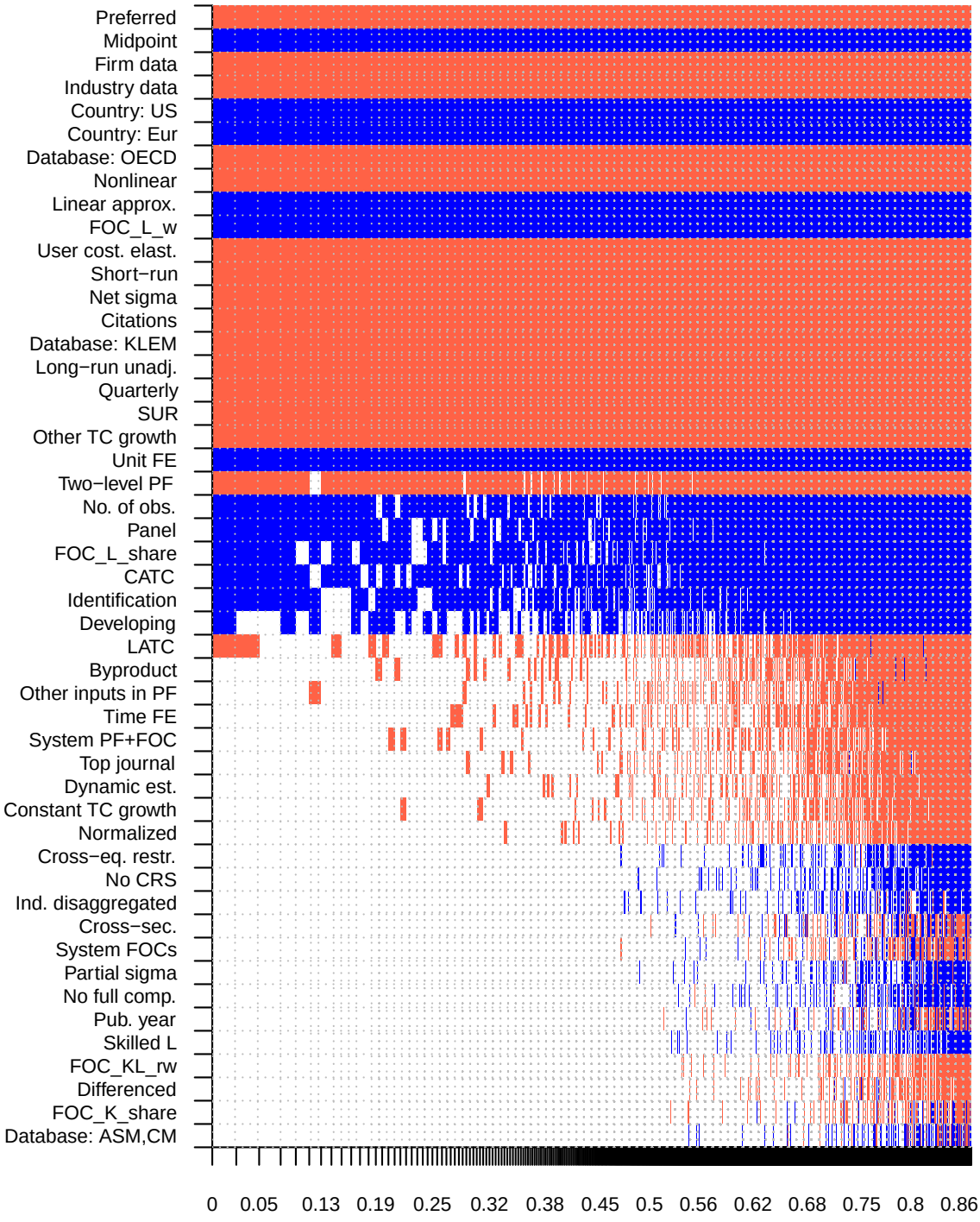
Notes: Our frequentist model averaging (FMA) exercise employs Mallow's weights (Hansen, 2007) and the orthogonalization of the covariate space suggested by Amini & Parmeter (2012). Dark gray color denotes variables that are deemed important also in the BMA exercise. Light gray color denote variables that are deemed important in the FMA but not BMA exercise.

Figure E1: Model inclusion in Bayesian model averaging, weighted by the inverse of the number of estimates per study



Notes: The response variable is the estimate of the elasticity of substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model.

Figure E2: Model inclusion in Bayesian model averaging, weighted by the inverse of the standard error



Notes: The response variable is the estimate of the elasticity of substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model.

Subsamples with measurement variables

As a complementary exercise to our baseline specification, we also run BMA analyses for subsamples of data in order to control for variables that are relevant only for a given subsample. We call these variables measurement variables. We need to create subsamples of the main dataset, because the variables relevant for the FOC for labor are not relevant for the FOC for capital, and vice versa. Regarding the estimates that utilize the FOC for labor, we include additional variables on how labor and the wage rate are measured. Regarding the estimates that utilize the FOC for capital, we include variables on how capital and the rental rate of capital are measured. Regarding industry-level estimates, we include the sector for which the elasticity was estimated, that is, primary, secondary and tertiary sectors; and, within the secondary sector, groups for industrial goods production, material goods production, and consumer goods production.

Concerning the measurement of labor, our reference category is measurement via the number of workers. We include a dummy equal to one if labor is measured using the number of hours worked. We also include a dummy variable that equals one if labor income is adjusted for self-employed labor income. As for the wage rate, we include dummy variables for the case when the rate is measured directly (in contrast to the situation when the wage rate is measured as the total amount paid to employees divided by the labor variable) and when the wage rate is used in nominal terms. In addition, we examine the effect of adjusting for changes in skill over time, for example, adjusting for the share of white- versus blue-collar workers.

Concerning the measurement of capital, our reference category is unspecified capital. We include dummies for specific measurements, including measurement as service flow, measurement via the perpetual inventory method, and capital stock in an index form. We code for special categories of capital stock: equipment, structures, IT, and residential capital stock. We include a separate dummy equal to one if the study controls for capacity utilization, either by adjusting the measurement variables or by adding it as a control. Underutilized capital would bias the results since it biases the effect of input on output (Brown, 1966); nevertheless, only a small portion of studies (Brown, 1966; Behrman, 1972; Dissou *et al.*, 2015, among others) explicitly use this approach, for example by including capacity utilization indices.

Regarding the rental rate of capital, the baseline category comprises the user cost of capital, or, in other words, the standard Hall-Jorgenson formula (Jorgenson, 1963; Hall & Jorgenson, 1967), which appears in two-thirds of all the estimations. The Hall-Jorgenson formula calculates the user cost of capital as a function of the relative price of capital, rate of return, and depreciation. We include a dummy for the case when the tax rate is an additional variable in the Hall-Jorgenson formula. The second most frequently used measurement is the quasi-rent approach, which calculates the rental rate of capital as a difference between total value added and total wages divided by the capital stock; this approach is used in 17% of the cases, for example in Dhrymes (1965), Ferguson (1965), and Lovell (1973). Further, the rental rate of capital can be measured either in gross terms or in net terms and in real or nominal terms; nevertheless, the variability in nominal user cost is almost zero, and thus we do not include the corresponding variable.

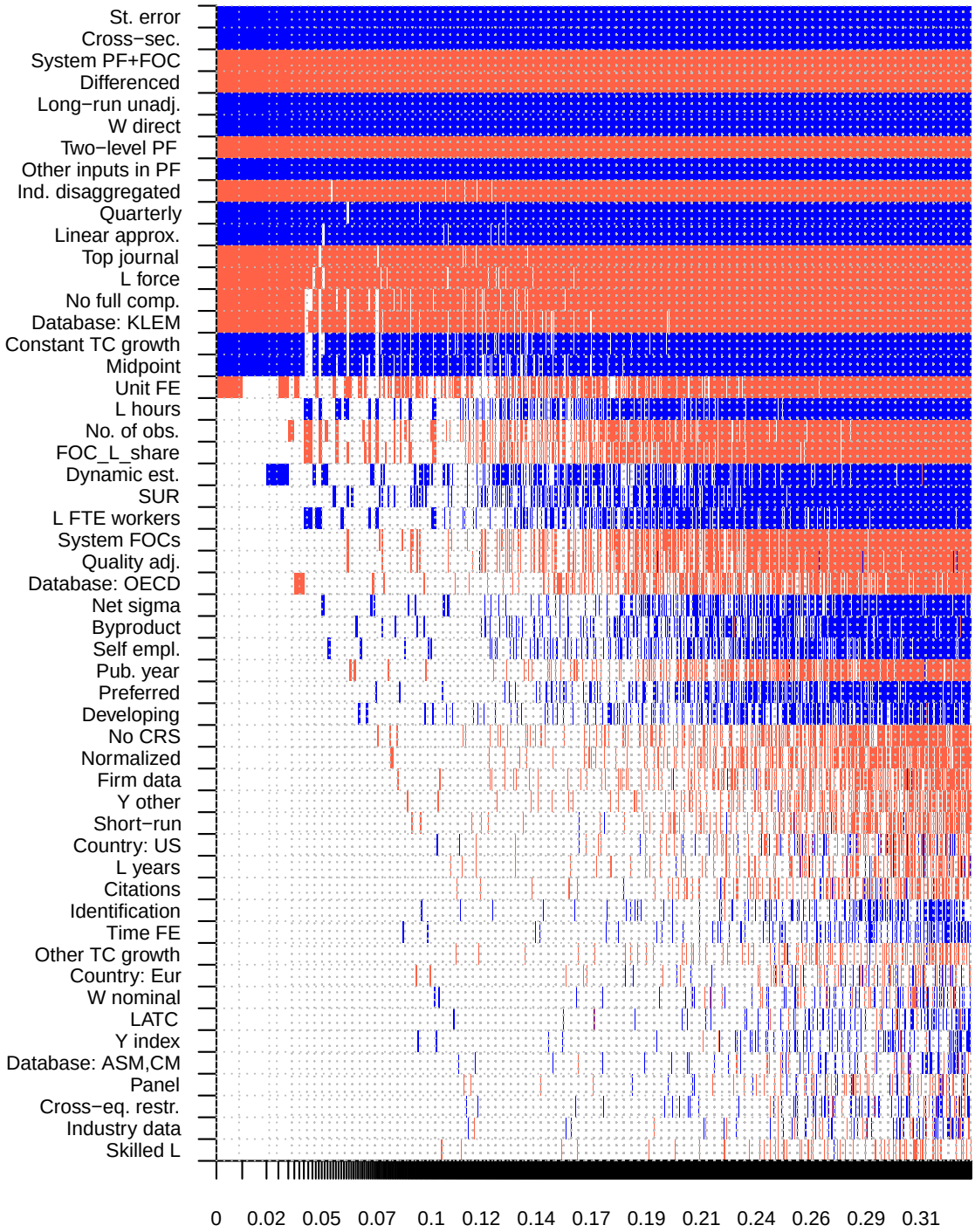
In all subsamples we control for the measurement of output: first, we include a dummy variable that equals one if output is not measured as gross product or in value added terms, but in another way—for example, as the amount of sales. Second, we include a dummy for the case when output is used in an index form.

How does the addition of these variables affect our results? First, we include labor-specific variables, which capture how labor and wage rate are measured, and run BMA on the subsample of data estimating the FOC for labor. The subsample covers less than half of the original dataset; the results are displayed in Figure E3. Only two of the newly included measurement variables are important for the explanation of the heterogeneity in the reported elasticities: direct measurement of the wage rate and measurement of labor as total labor force. The main drivers of heterogeneity remain the same while the total explanatory power of the analysis increases only marginally.

Concerning capital-related variables, we find that the type of capital under examination represents an important driver of the differences in results (Figure E4). IT capital and equipment capital are more substitutable with labor than other types of capital, such as buildings. When capital is measured as service flow, the estimates typically yield a larger elasticity of substitution. It also matters how the rental rate of capital, r, is computed, specifically whether the Hall-Jorgenson formula is used—we find that it yields smaller elasticities than do other approaches. The best-practice estimate derived from both subsamples and conditional on plugging in mean values for measurement variables would again equal 0.3, far from the Cobb-Douglas assumption.

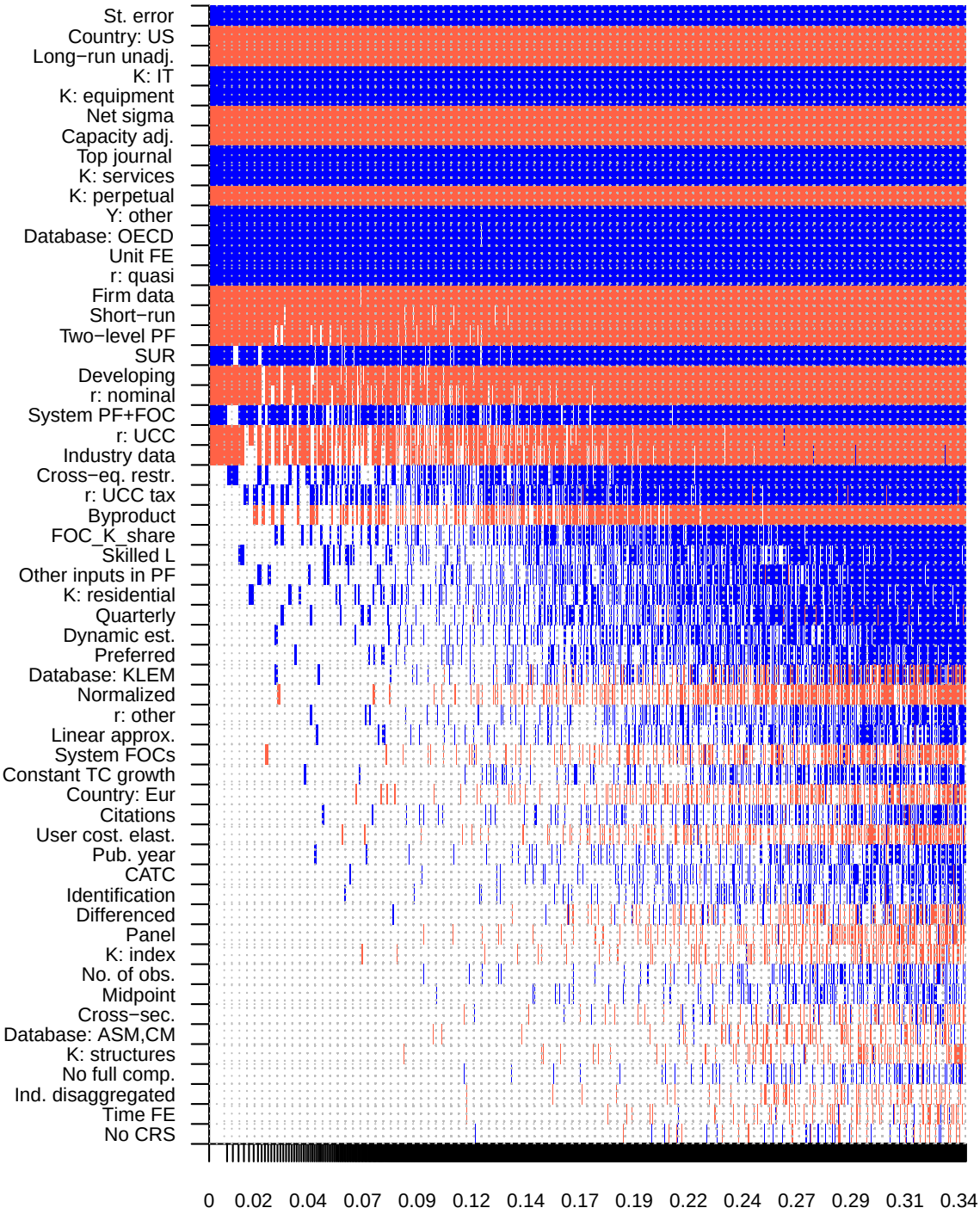
Finally, for the subsample of disaggregated elasticities we run the baseline BMA enriched with industry-relevant variables in Figure E5. We do not find any significant determinants that would suggest that the elasticity of capital-labor substitution differs systematically across sectors or industry groups (production of materials, production of industrial goods, production of consumer goods, and production of services). Given the number of variables in our analysis, it is infeasible to add more industry-specific variables since that would create troubles with collinearity.

Figure E3: Model inclusion in Bayesian model averaging, labor-specific variables



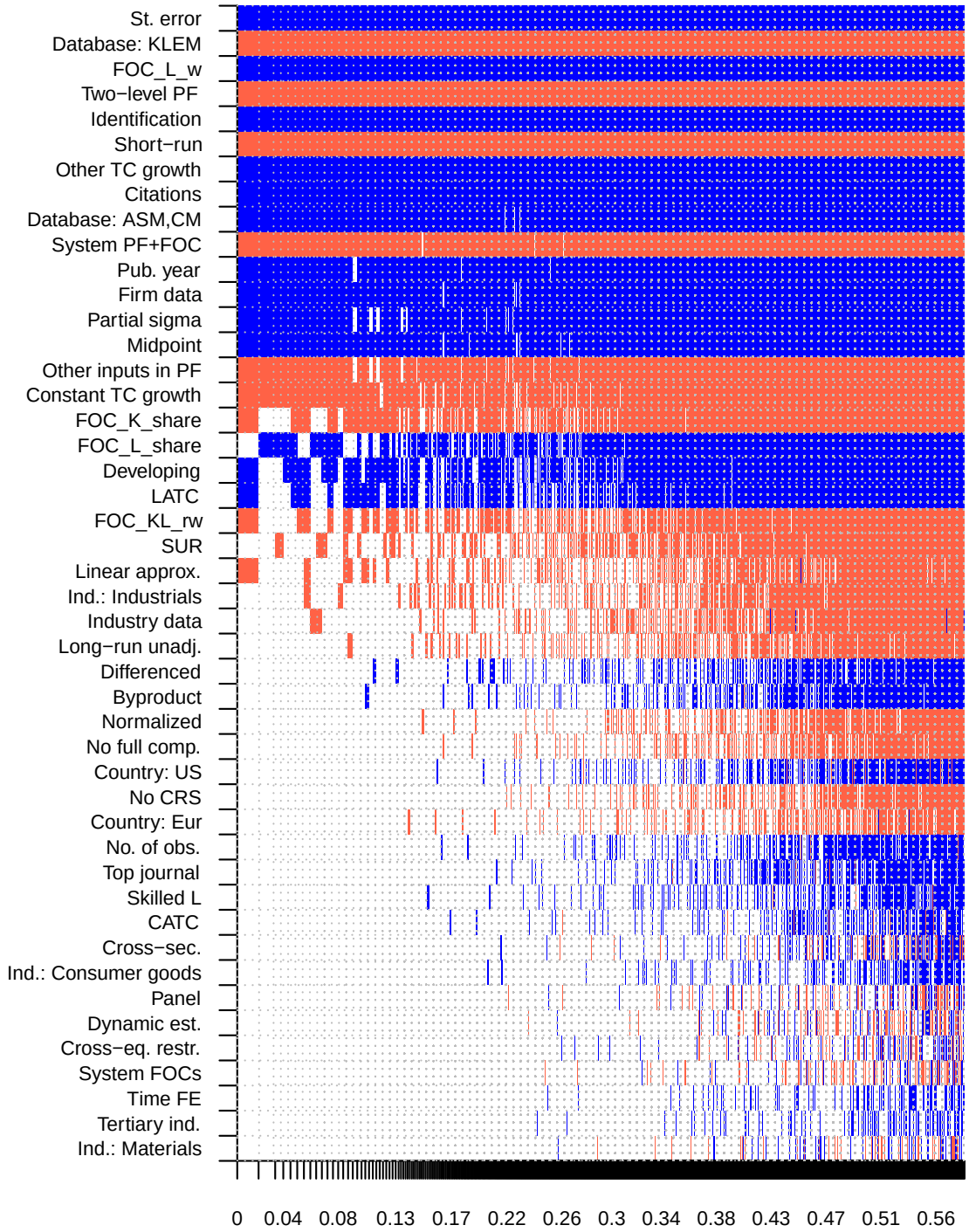
Notes: The response variable is the estimate of the elasticity of substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model.

Figure E4: Model inclusion in Bayesian model averaging, capital-specific variables



Notes: The response variable is the estimate of the elasticity of substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model.

Figure E5: Model inclusion in Bayesian model averaging, industry-specific variables



Notes: The response variable is the estimate of the elasticity of substitution. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes cumulative posterior model probabilities; only the 5,000 best models are shown. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model.

Appendix F Studies Included in the Dataset (for Online Publication)

Abed, G. T. (1975): "Labour Absorption in Industry: An Analysis with Reference to Egypt." *Oxford Economic Papers* **27(3)**: pp. 400–425.

Akay, G. H. & C. Dogan (2013): "The Effect of Labor Supply Changes on Output: Empirical Evidence from US Industries." *Journal of Productivity Analysis* **39(2)**: pp. 123–130.

Antras, P. (2004): "Is the US Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution." *Contributions in Macroeconomics* **4(1)**: pp. 1–36.

Apostolakis, B. E. (1984): "A Translogarithmic Cost Function Approach: Greece, 1953–1977." *Empirical Economics* **9(4)**: pp. 247–262.

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

Artus, J. R. (1984): "The Disequilibrium Real Wage Rate Hypothesis: An Empirical Evaluation." *IMF Staff Papers* **31(2)**: pp. 249–302.

Asher, E. (1972): "Industrial Efficiency and Biased Technical Change in American and British Manufacturing: The Case of Textiles in the Nineteenth Century." *Journal of Economic History* **32(2)**: pp. 431–442.

Balistreri, E. J., C. A. McDaniel, & E. V. Wong (2003): "An Estimation of US Industry-Level Capital–Labor Substitution Elasticities: Support for Cobb–Douglas." *North American Journal of Economics and Finance* **14(3)**: pp. 343–356.

Bartelsman, E. J. & R. M. Beetsma (2003): "Why Pay More? Corporate Tax Avoidance through Transfer Pricing in OECD Countries." *Journal of Public Economics* **87(9-10)**: pp. 2225–2252.

Behrman, J. (1982): "Country and Sectoral Variations in Manufacturing Elasticities of Substitution between Capital and Labor." In "Trade and Employment in Developing Countries, Volume 2: Factor Supply and Substitution," pp. 159–192. University of Chicago.

Behrman, J. R. (1972): "Sectoral Elasticities of Substitution between Capital and Labor in a Developing Economy: Times Series Analysis in the Case of Postwar Chile." *Econometrica* **40(2)**: pp. 311–326.

Bentolila, S. & G. Saint-Paul (2003): "Explaining Movements in the Labor Share." *Contributions in Macroeconomics* **3(1)**: pp. 1–33.

Berndt, E. R. (1976): "Reconciling Alternative Estimates of the Elasticity of Substitution." *Review of Economics and Statistics* **58(1)**: pp. 59–68.

Berthold, N., R. Fehn, & E. Thode (2002): "Falling Labor Share and Rising Unemployment: Long–Run Consequences of Institutional Shocks?" *German Economic Review* **3(4)**: pp. 431–459.

Binswanger, H. P. (1974): "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution." *American Journal of Agricultural Economics* **56(2)**: pp. 377–386.

Blanchard, O. J. (1997): "The Medium Run." *Brookings Papers on Economic Activity* **1997(2)**: pp. 89–141.

Bodkin, R. G. & L. R. Klein (1967): "Nonlinear Estimation of Aggregate Production Functions." *Review of Economic and Statistics* **49(1)**: pp. 28–44.

Brown, M. (1966): "A Measure of the Change in Relative Exploitation of Capital and Labor." *Review of Economics and Statistics* **48(2)**: pp. 182–192.

Brown, M. & J. S. De Cani (1963): "Technological Change and the Distribution of Income." *International Economic Review* **4(3)**: pp. 289–309.

Brox, J. A. & C. A. Fader (2005): "Infrastructure Investment and Canadian Manufacturing Productivity." *Applied Economics* **37(11)**: pp. 1247–1256.

Bruno, M. & J. Sachs (1982): "Input Price Shocks and the Slowdown in Economic Growth: The Case of UK Manufacturing." *Review of Economic Studies* **49(5)**: pp. 679–705.

Caballero, R. J. (1994): "Small Sample Bias and Adjustment Costs." *Review of Economics and Statistics* **76(1)**: pp. 52–58.

Chetty, V. K. & U. Sankar (1969): "Bayesian Estimation of the CES Production Function." *Review of Economic Studies* **36(3)**: pp. 289–294.

Chirinko, R. S., S. M. Fazzari, & A. P. Meyer (1999): "How Responsive Is Business Capital Formation to its User Cost? An Exploration with Micro Data." *Journal of Public Economics* **74(1)**: pp. 53–80.

Chirinko, R. S., S. M. Fazzari, & A. P. Meyer (2011): "A New Approach to Estimating Production Function Parameters: The Elusive Capital–Labor Substitution Elasticity." *Journal of Business & Economic Statistics* **29(4)**: pp. 587–594.

Chirinko, R. S. & D. Mallick (2017): "The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model." *American Economic Journal: Macroeconomics* **9(4)**: pp. 225–53.

Chwelos, P., R. Ramirez, K. L. Kraemer, & N. P. Melville (2010): "Does Technological Progress Alter the Nature of Information Technology as a Production Input? New Evidence and New Results." *Information Systems Research* **21(2)**: pp. 392–408.

Clark, P. K. & D. E. Sichel (1993): "Tax Incentives and Equipment Investment." *Brookings Papers on Economic Activity* **1993(1)**: pp. 317–347.

Claro, S. (2003): "A Cross-Country Estimation of the Elasticity of Substitution between Labor and Capital in Manufacturing Industries." *Cuadernos de economía* **40(120)**: pp. 239–257.

Cummins, J. G. & K. A. Hassett (1992): "The Effects of Taxation on Investment: New Evidence from Firm Level Panel Data." *National Tax Journal* **45(3)**: pp. 243–251.

Cummins, J. G., K. A. Hassett, R. G. Hubbard, R. E. Hall, & R. J. Caballero (1994): "A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments." *Brookings Papers on Economic Activity* **1994(2)**: pp. 1–74.

Daniels, M. R. (1969): "Differences in Efficiency among Industries in Developing Countries." *American Economic Review* **59(1)**: pp. 159–171.

David, P. A. & T. Van de Klundert (1965): "Biased Efficiency Growth and Capital-Labor Substitution in the US, 1899-1960." *American Economic Review* **55(3)**: pp. 357–394.

Dhrymes, P. J. (1965): "Some Extensions and Tests for the CES Class of Production Functions." *Review of Economics and Statistics* **47(4)**: pp. 357–366.

Dissou, Y. & R. Ghazal (2010): "Energy Substitutability in Canadian Manufacturing Econometric Estimation with Bootstrap Confidence Intervals." *Energy Journal* **31(1)**: pp. 121–148.

Dissou, Y., L. Karnizova, & Q. Sun (2015): "Industry-Level Econometric Estimates of Energy-Capital-Labor Substitution with a Nested CES Production Function." *Atlantic Economic Journal* **43(1)**: pp. 107–121.

Donges, J. B. (1972): "Returns to Scale and Factor Substitutability in the Spanish Industry." *Review of World Economics* **108(4)**: pp. 597–608.

Duffy, J. & C. Papageorgiou (2000): "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification." *Journal of Economic Growth* **5(1)**: pp. 87–120.

Dwenger, N. (2014): "User Cost Elasticity of Capital Revisited." *Economica* **81(321)**: pp. 161–186.

Easterly, W. & S. Fischer (1995): "The Soviet Economic Decline." *World Bank Economic Review* **9(3)**: pp. 341–371.

Eisner, R. (1967): "Capital and Labor in Production: Some Direct Estimates." In "The Theory and Empirical Analysis of Production," pp. 431–475. NBER.

Eisner, R. (1969): "Tax Policy and Investment Behavior: Comment." *American Economic Review* **59(3)**: pp. 379–388.

Eisner, R. & M. I. Nadiri (1968): "Investment Behavior and Neo-Classical Theory." *Review of Economics and Statistics* **50(3)**: pp. 369–382.

Elbers, C., J. W. Gunning, & B. Kinsey (2007): "Growth and Risk: Methodology and Micro Evidence." *World Bank Economic Review* **21(1)**: pp. 1–20.

Ellis, C. & S. Price (2004): "UK Business Investment and the User Cost of Capital." *Manchester School* **72**: pp. 72–93.

Feldstein, M. S. (1967): "Specification of the Labour Input in the Aggregate Production Function." *Review of Economic Studies* **34(4)**: pp. 375–386.

Feldstein, M. S. & J. S. Flemming (1971): "Tax Policy, Corporate Saving and Investment Behaviour in Britain." *Review of Economic Studies* **38(4)**: pp. 415–434.

Felipe, J. & J. McCombie (2009): "Are Estimates of Labour Demand Functions Mere Statistical Artefacts?" *International Review of Applied Economics* **23(2)**: pp. 147–168.

Ferguson, C. E. (1965): "Time-Series Production Functions and Technological Progress in American Manufacturing Industry." *Journal of Political Economy* **73(2)**: pp. 135–147.

Fishelson, G. (1979): "Elasticity of Factor Substitution in Cross-Section Production Functions." *Review of Economics and Statistics* **61(3)**: pp. 432–436.

Fitchett, D. A. (1976): "Capital-Labor Substitution in the Manufacturing Sector of Panama." *Economic Development and Cultural Change* **24(3)**: pp. 577–592.

Fuchs, V. R. (1963): "Capital-Labor Substitution: A Note." *Review of Economics and Statistics* **45(4)**: pp. 436–438.

Griliches, Z. (1964): "Research Expenditures, Education, and the Aggregate Agricultural Production Function." *American Economic Review* **54(6)**: pp. 961–974.

Griliches, Z. (1967): "Production Functions in Manufacturing: Some Preliminary Results." In "The Theory and Empirical Analysis of Production," pp. 275–340. NBER.

Herrendorf, B., C. Herrington, & A. Valentinyi (2015): "Sectoral Technology and Structural Transformation." *American Economic Journal: Macroeconomics* **7(4)**: pp. 104–33.

Hijzen, A. & P. Swaim (2010): "Offshoring, Labour Market Institutions and the Elasticity of Labour Demand." *European Economic Review* **54(8)**: pp. 1016–1034.

Hossain, S. I. (1987): "Allocative and Technical Efficiency: A Study of Rural Enterprises in Bangladesh." *Developing Economies* **25(1)**: pp. 56–72.

Humphrey, D. B. & J. R. Moroney (1975): "Substitution among Capital, Labor, and Natural Resource Products in American Manufacturing." *Journal of Political Economy* **83(1)**: pp. 57–82.

Iqbal, M. (1986): "Substitution of Labour, Capital and Energy in the Manufacturing Sector of Pakistan." *Empirical Economics* **11(2)**: pp. 81–95.

Jalava, J., M. Pohjola, A. Ripatti, & J. Vilmunen (2006): "Biased Technical Change and Capital-

Labour Substitution in Finland, 1902-2003." *Topics in Macroeconomics* **6(1)**: pp. 1–20.

Jones, D. C. & D. K. Backus (1977): "British Producer Cooperatives in the Footware Industry: An Empirical Evaluation of the Theory of Financing." *Economic Journal* **87(347)**: pp. 488–510.

Judzik, D. & H. Sala (2015): "The Determinants of Capital Intensity in Japan and the US." *Journal of the Japanese and International Economies* **35**: pp. 78–98.

Juselius, M. (2008): "Long-Run Relationships between Labor and Capital: Indirect Evidence on the Elasticity of Substitution." *Journal of Macroeconomics* **30(2)**: pp. 739–756.

Kalt, J. P. (1978): "Technological Change and Factor Substitution in the United States: 1929-1967." *International Economic Review* **19(3)**: pp. 761–775.

Karabarbounis, L. & B. Neiman (2013): "The Global Decline of the Labor Share." *Quarterly Journal of Economics* **129(1)**: pp. 61–103.

Kilponen, J. & M. Viren (2010): "Why Do Growth Rates Differ? Evidence from Cross-Country Data on Private Sector Production." *Empirica* **37(3)**: pp. 311–328.

Kislev, Y. & W. Peterson (1982): "Prices, Technology, and Farm Size." *Journal of Political Economy* **90(3)**: pp. 578–595.

Klump, R., P. McAdam, & A. Willman (2007): "Factor Substitution and Factor-Augmenting Technical Progress in the United States: A Normalized Supply-Side System Approach." *Review of Economics and Statistics* **89(1)**: pp. 183–192.

Klump, R., P. McAdam, & A. Willman (2008): "Unwrapping Some Euro Area Growth Puzzles: Factor Substitution, Productivity and Unemployment." *Journal of Macroeconomics* **30(2)**: pp. 645–666.

Kmenta, J. (1967): "On Estimation of the CES Production Function." *International Economic Review* **8(2)**: pp. 180–189.

Krusell, P., L. E. Ohanian, J.-V. Roos-Rull, & G. L. Violante (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* **68(5)**: pp. 1029–1053.

Lee, M. & M. Tcha (2004): "The Color of Money: The Effects of Foreign Direct Investment on Economic Growth in Transition Economies." *Review of World Economics* **140(2)**: pp. 211–229.

León-Ledesma, M. A., P. McAdam, & A. Willman (2015): "Production Technology Estimates and Balanced Growth." *Oxford Bulletin of Economics and Statistics* **77(1)**: pp. 40–65.

Leung, D. & T. Yuen (2010): "Do Exchange Rates Affect the Capital–Labour Ratio? Panel Evidence from Canadian Manufacturing Industries." *Applied Economics* **42(20)**: pp. 2519–2535.

Lianos, T. P. (1971): "The Relative Share of Labor in United States Agriculture, 1949–1968." *American Journal of Agricultural Economics* **53(3)**: pp. 411–422.

Lianos, T. P. (1975): "Capital-Labor Substitution in a Developing Country: The Case of Greece." *European Economic Review* **6(2)**: pp. 129–141.

Lin, W. T. & B. B. Shao (2006): "The Business Value of Information Technology and Inputs Substitution: The Productivity Paradox Revisited." *Decision Support Systems* **42(2)**: pp. 493–507.

Lovell, C. K. (1973a): "CES and VES Production Functions in a Cross-Section Context." *Journal of Political Economy* **81(3)**: pp. 705–720.

Lovell, C. K. (1973b): "Estimation and Prediction with CES and VES Production Functions." *International Economic Review* **14(3)**: pp. 676–692.

Luoma, A. & J. Luoto (2010): "The Aggregate Production Function of the Finnish Economy in the Twentieth Century." *Southern Economic Journal* **76(3)**: pp. 723–737.

Mallick, D. (2012): "The Role of the Elasticity of Substitution in Economic Growth: A Cross-Country Investigation." *Labour Economics* **19(5)**: pp. 682–694.

Martin, S. A., R. McHugh, & S. R. Johnson (1993): "The Influence of Location on Productivity: Manufacturing Technology in Rural and Urban Areas." *Growth and Change* **24(4)**: pp. 459–486.

Masanjala, W. H. & C. Papageorgiou (2004): "The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity." *Journal of Applied Econometrics* **19(2)**: pp. 171–201.

McAdam, P. & A. Willman (2004): "Production, Supply and Factor Shares: An Application to Estimating German Long-Run Supply." *Economic Modelling* **21(2)**: pp. 191–215.

McCallum, J. (1985): "Wage Gaps, Factor Shares and Real Wages." *Scandinavian Journal of Economics* **87(2)**: pp. 436–459.

McKinnon, R. I. (1962): "Wages, Capital Costs, and Employment in Manufacturing: A Model Applied to 1947-58 US Data." *Econometrica* **30(3)**: pp. 501–521.

McLean-Meyinsse, P. E. & A. A. Okunade (1988): "Factor Demands of Louisiana Rice Producers: An Econometric Investigation." *Journal of Agricultural and Applied Economics* **20(2)**: pp. 127–136.

Meller, P. (1975): "Production Functions for Industrial Establishments of Different Sizes: The Chilean Case." In "Annals of Economic and Social Measurement, Volume 4, number 4," pp. 595–634. NBER.

Minasian, J. R. (1961): "Elasticities of Substitution and Constant-Output Demand Curves for Labor." *Journal of Political Economy* **69(3)**: pp. 261–270.

Mohabbat, K. A. & A. J. Dalai (1983): "Factor Substitution and Import Demand for South Korea: A Translog Analysis." *Review of World Economics* **119(4)**: pp. 709–723.

Mohabbat, K. A., A. Dalal, & M. Williams (1984): "Import Demand for India: A Translog Cost Function Approach." *Economic Development and Cultural Change* **32(3)**: pp. 593–605.

Moroney, J. R. (1966): "Time-Series Elasticities of Substitution and Labor's Share in US Manufacturing: The Postwar Period." *Southern Economic Journal* **32(4)**: p. 474.

Moroney, J. R. (1970): "Identification and Specification Analysis of Alternative Equations for Estimating the Elasticity of Substitution." *Southern Economic Journal* **36(3)**: pp. 287–299.

Moroney, J. R. & B. T. Allen (1969): "Monopoly Power and the Relative Share of Labor." *ILR Review* **22(2)**: pp. 167–178.

Moroney, J. R. & A. L. Toevs (1977): "Factor Costs and Factor Use: An Analysis of Labor, Capital, and Natural Resource Inputs." *Southern Economic Journal* **44(2)**: pp. 222–239.

Nadiri, M. I. (1968): "The Effects of Relative Prices and Capacity on the Demand for Labour in the US Manufacturing Sector." *Review of Economic Studies* **35(3)**: pp. 273–288.

Panik, M. J. (1976): "Factor Learning and Biased Factor-Efficiency Growth in the United States, 1929-1966." *International Economic Review* **17(3)**: pp. 733–739.

Parks, R. W. (1971): "Price Responsiveness of Factor Utilization in Swedish Manufacturing, 1870-1950." *Review of Economics and Statistics* **53(2)**: pp. 129–139.

Pollak, R. A., R. C. Sickles, & T. J. Wales (1984): "The CES-Translog: Specification and Estimation of a New Cost Function." *Review of Economics and Statistics* **66(4)**: pp. 602–607.

Raurich, X., H. Sala, & V. Sorolla (2012): "Factor Shares, the Price Markup, and the Elasticity of Substitution between Capital and Labor." *Journal of Macroeconomics* **34(1)**: pp. 181–198.

Roskamp, K. W. (1977): "Labor Productivity and the Elasticity of Factor Substitution in West German Industries 1950-1960." *Review of Economics and Statistics* **59(3)**: pp. 366–371.

Sahota, G. S. (1966): "The Sources of Measured Productivity Growth: United States Fertilizer Mineral Industries, 1936-1960." *Review of Economics and Statistics* **48(2)**: pp. 193–204.

Salvanes, K. G. (1989): "The Structure of the Norwegian Fish Farming Industry: An Empirical Analysis of Economies of Scale and Substitution Possibilities." *Marine Resource Economics* **6(4)**: pp. 349–373.

Sankar, U. (1972): "Investment Behavior in the US Electric Utility Industry, 1949-1968." *Bell Journal of Economics* **3(2)**: pp. 645–664.

Sapir, A. (1980): "Economic Growth and Factor Substitution: What Happened to the Yugoslav Miracle?" *Economic Journal* **90(358)**: pp. 294–313.

Sato, K. (1977): "A note on Factor Substitution and Efficiency." *Review of Economics and Statistics* **59(3)**: pp. 360–366.

Sato, R. & R. F. Hoffman (1968): "Production Functions with Variable Elasticity of Factor Substitution: Some Analysis and Testing." *Review of Economics and Statistics* **50(4)**: pp. 453–460.

Saxonhouse, G. (1977): "Productivity Change and Labor Absorption in Japanese Cotton Spinning 1891–1935." *Quarterly Journal of Economics* **91(2)**: pp. 195–219.

Schaller, H. (2006): "Estimating the Long-Run User Cost Elasticity." *Journal of Monetary Economics* **53(4)**: pp. 725–736.

Schmitz, M. (1981): "The Elasticity of Substitution in 19th-Century Manufacturing." *Explorations in Economic History* **18(3)**: p. 290.

Semieniuk, G. (2017): "Piketty's Elasticity of Substitution: A Critique." *Review of Political Economy* **29(1)**: pp. 64–79.

Shahe Emran, M., F. Shilpi, & M. I. Alam (2007): "Economic Liberalization and Price Response of Aggregate Private Investment: Time Series Evidence from India." *Canadian Journal of Economics* **40(3)**: pp. 914–934.

Smith, J. (2008): "That Elusive Elasticity and the Ubiquitous Bias: Is Panel Data a Panacea?" *Journal of Macroeconomics* **30(2)**: pp. 760–779.

Solow, R. (1964): "Capital, Labor, and Income in Manufacturing." In "The Behavior of Income Shares: Selected Theoretical and Empirical Issues," pp. 101–142. Princeton University Press.

Tevlin, S. & K. Whelan (2003): "Explaining the Investment Boom of the 1990s." *Journal of Money, Credit and Banking* **35(1)**: pp. 1–22.

Tsang, H. H. & J. J. Persky (1975): "On the Empirical Content of CES Production Functions." *Economic Record* **51(4)**: pp. 539–548.

Weitzman, M. L. (1970): "Soviet Postwar Economic Growth and Capital-Labor Substitution." *American Economic Review* **60(4)**: pp. 676–692.

Van der Werf, E. (2008): "Production Functions for Climate Policy Modeling: An Empirical Analysis." *Energy Economics* **30(6)**: pp. 2964–2979.

Williams, M. & P. S. Laumas (1984): "Economies of Scale for Various Types of Manufacturing Production Technologies in an Underdeveloped Economy." *Economic Development and Cultural Change* **32(2)**: pp. 401–412.

Young, A. T. (2013): "US Elasticities of Substitution and Factor Augmentation at the Industry Level." *Macroeconomic Dynamics* **17(4)**: pp. 861–897.

Zarembka, P. (1970): "On the Empirical Relevance of the CES Production Function." *Review of Economics and Statistics* **52(1)**: pp. 47–53.