Department Discussion Paper DDP1807

ISSN 1914-2838

NORMALIZED CES SUPPLY SYSTEMS: REPLICATION OF KLUMP, MCADAM, AND WILLMAN (2007)

Kenneth G. Stewart
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

Abstract

In light of the literature that it has spawned, Klump, McAdam, and Willman (Factor substitution and factor-augmenting technical progress in the United States: A normalized supply-side system approach. *Review of Economics and Statistics* 2007; 89(1): 183--192) can be seen to be an iconic empirical implementation that has helped resurrect interest in CES functional forms and parametric supply-side systems. This replication revisits their analysis using alternative software, confirming their results substantively and, in large measure, numerically. Contributions include a more explicit consideration of the nested testing structure than has appeared previously, and the appropriate means of imposing and testing the special case of logarithmic growth in technology. As well, plots of the likelihood serve to emphasize that maxima in the neighborhood of a unitary elasticity of substitution are often a spurious artifact of the singularity of the model at this point, of which empirical researchers should beware.

Keywords: capital-labour substitution, technological change, supply-side system

JEL Classifications: C22, E23, E25, O30, O51

Author Contact:
Kenneth Stewart, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; E-mail: kstewart@uvic.ca; Voice: (250) 721-8534.

Normalized CES Supply Systems:
Replication of Klump, McAdam, and Willman (2007)

Kenneth G. Stewart

Associate Professor
Department of Economics
University of Victoria
Victoria, British Columbia
Canada V8W 2Y2

email: kstewart@uvic.ca
phone: (250) 721-8534

Abstract

In light of the literature that it has spawned, Klump, McAdam, and Willman (Factor substitution and factor-augmenting technical progress in the United States: A normalized supply-side system approach. *Review of Economics and Statistics* 2007; 89(1): 183–192) can be seen to be an iconic empirical implementation that has helped resurrect interest in CES functional forms and parametric supply-side systems. This replication revisits their analysis using alternative software, confirming their results substantively and, in large measure, numerically. Contributions include a more explicit consideration of the nested testing structure than has appeared previously, and the appropriate means of imposing and testing the special case of logarithmic growth in technology. As well, plots of the likelihood serve to emphasize that maxima in the neighborhood of a unitary elasticity of substitution are often a spurious artifact of the singularity of the model at this point, of which empirical researchers should beware.

Acknowledgements

I thank Alpo Willman for providing the core RATS code that allowed me to double-check my TSP implementation, and the data.

In the absence of additional a priori structure on a production technology, it is not possible to separately identify the elasticity of substitution and distinct factor biases of technical change. This is the "non-identification" or "impossibility" theorem of Diamond, McFadden, and Rodriguez (1978), who establish

> . . . a non-identifiability of the elasticity and bias in the absence of a priori hypotheses on the structure of technical change. More precisely, given the time series of all observable market phenomena for a single economy with a classical aggregate production function, one finds that the same time series *could* have been generated by an alternative production function having an arbitrary elasticity or bias at the observed points . . . The identifiability of the elasticity and bias will depend on what is in fact true about the economy and on what the economist assumes a priori to be true (i.e., his maintained hypothesis, or model).

As such a model to overcome this non-identifiability, Klump, McAdam, and Willman (KMW) have studied the estimation of a constant elasticity of substitution (CES) production function joint with its implied factor demands. KMW (2007a) applied their model to annual US data 1953–98, while KMW (2007b) compared the results from annual US data 1953–2002 and quarterly euro-area data 1970–2003. KMW (2008) uses quarterly euro-area data 1970–2005, although calibrating one key parameter instead of estimating it.

Of course, the strategy of estimating a production function and its factor demands as a system is not original to KMW, dating at least to (empirically) Bodkin and Klein (1967) and even (conceptually) Marshak and Andrews (1944). However KMW introduced two innovations to their CES system that, they demonstrate convincingly, are critical to disentangling factor substitution and biases of technical change. First, they draw on earlier theoretical contributions—notably La Grandville (1989), Klump and La Grandville (2000), and Klump and Preissler (2000)—to formulate their empirical CES supply-side system in *normalized* form. Second, they use Box-Cox transformations of the factor-specific technology parameters to distinguish long-term from short-term biases of technical change. This is important because of Uzawa's (1961) famous result that balanced growth requires technical change to be labor augmenting. Given that it concerns balanced growth, Uzawa's steady state growth theorem does not preclude *any* capital bias to technological progress, only that it disappear in the long run, leaving labor-augmenting technical change as the sole long-run driver of growth in living standards. Remarkably, this key prediction of growth theory is testable in the KMW framework.

Subsequent work includes León-Ledesma, McAdam, and Willman (2010), who provide important simulation evidence demonstrating the ability of the model to identify these distinct elements, and the survey paper by KMW (2012). Applications of the framework to study broader macroeconomic and growth issues include McAdam and Willman (2013) and León-Ledesma, McAdam, and Willman (2015).

In view of this accumulated literature, the original KMW (2007a) estimation can now be seen to be an iconic empirical implementation. It is therefore useful to revisit it, investigating the sensitivity of the estimation results to alternative software. Whereas KMW used RATS, I use TSP, the numerics of which have been favorably evaluated by McCullough (1999). Henceforth I cite the principal article simply as KMW, referencing others by date.

As part of the replication, I lay out the model's complete nested testing structure more explicitly than has been done previously and use it to explore the incidence of multiple maxima of the likelihood function. As well, I clarify some nuances in the estimation of these systems that may be helpful to future researchers, such as the proper imposition and testing of the special case of logarithmic growth in technology. I also plot the loglikelihood of the system, something that does not appear to have been done in the published literature.[1] This allows me to conclude more strongly than KMW did that the multiple maxima that seem endemic to their model are of less practical importance than might appear, because they are typically an artifact of the singularity of the system at $\sigma = 1$.

1 The Klump-McAdam-Willman Model

The many issues surrounding the specification and estimation of CES supply-side systems, including normalization, are thoroughly treated in the articles cited above, of which KMW (2012) provides the best comprehensive overview. Here I merely summarize the essentials needed to understand my replication.

1.1 Growth Specifications of the Factor Efficiencies

The starting point is a constant returns to scale CES production function, one expression for which is

$$Y_t = \left[(E_t^N N_t)^{-\rho} + (E_t^K K_t)^{-\rho}\right]^{-1/\rho}. \tag{1}$$

Notation is for the most part standard, with N denoting labour, K capital, and ρ the substitution parameter. The labor- and capital-augmenting technology factors E_t^N and E_t^K grow along trajectories given by

$$E_t^i = E_0^i \mathrm{e}^{g_i(t)} \qquad (i = N, K),$$

[1] In their working paper KMW (2004) provide scatter plots of the local minima of the log determinant of the models they estimate. Their Graph 4.3 shows the two minima of the maintained model corresponding to the two maxima of the loglikelihood in my Figure 2.

with associated growth rates

$$\frac{\mathrm{d}\log E_t^i}{\mathrm{d}t} = \frac{\mathrm{d}g_i(t)}{\mathrm{d}t} \qquad (i = N, K).$$

The textbook case of constant growth at rate γ_i is $g_i(t) = \gamma_i t$. But even if this provides a good approximation to long run growth—itself an open question—it may be unwarranted in the short run. KMW use the Box-Cox transformation to allow more general growth trajectories, specifying $g_i(t)$ as

$$g_i(t) = \gamma_i \left(\frac{t^{\lambda_i} - 1}{\lambda_i} \right) \qquad (i = N, K) \tag{2}$$

so that the rates of technical change depend on the curvature parameters λ_i:

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = \gamma_i t^{\lambda_i - 1} = \begin{cases} \to \infty \text{ as } t \to \infty & \text{if } \lambda_i > 1 \quad \text{(accelerating growth);} \\ \gamma_i & \text{if } \lambda_i = 1 \quad \text{(constant growth);} \\ \to 0 \text{ as } t \to \infty & \text{if } \lambda_i < 1 \quad \text{(decelerating growth).} \end{cases}$$

Although this admits accelerating growth ($\lambda_i > 1$) in technology as a possibility, this is implausible empirically in the long run, as KMW find and I replicate.

Within the range $\lambda_i < 1$ lies the special case $\lambda_i = 0$. As $\lambda_i \to 0$ the Box-Cox function yields the logarithmic transformation

$$g_i(t) = \gamma_i \left(\frac{t^{\lambda_i} - 1}{\lambda_i} \right) \to \gamma_i \log t \quad \text{ as } \lambda_i \to 0$$

so that

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = \gamma_i \frac{\log t}{\mathrm{d}t} = \gamma_i \frac{1}{t} \to 0 \quad \text{ as } t \to \infty.$$

Thus the rate of factor-i-augmenting technological progress decelerates to zero if $\lambda_i < 1$. If $0 < \lambda_i < 1$ this deceleration is slower than when $g_i(t) = \gamma_i \log t$, while if $\lambda_i < 0$ it is faster. Hence, although $\lambda_i = 0$ might be regarded as a benchmark rate of deceleration, in terms of the qualitative properties of the growth trajectory it is of no special interest. If appropriate it does, however, simplify the numerics of nonlinear estimation by replacing the Box-Cox function with the log function, something that turns out to be relevant for capital K in this replication.

Within this KMW framework, notions of neutral technological progress that are prominent in the growth literature are well defined as restricted versions of the model. Consider a generic production function $\tilde{F}(N_t, K_t, A_t)$ where A_t is technology.

Technological progress is *Hicks neutral* if the function factors as $\tilde{F}(N_t, K_t, A_t) = A_t F(N_t, K_t)$. In the production function (1) with growth exponents (2) this implies the restrictions $\gamma_N = \gamma_K > 0$, $\lambda_N = \lambda_K$.

Technological progress is *Harrod neutral* if technology is solely labor-augmenting: $\tilde{F}(N_t, K_t, A_t) = F(A_t N_t, K_t)$. Harrod neutrality holds

- in both the short and long run if $\gamma_K = 0$, $\gamma_N > 0$, $\lambda_N \geq 1$;

- only in the long run if $\lambda_K < 1$, $\gamma_N > 0$, $\lambda_N \geq 1$.
- If Harrod neutrality is defined to preclude accelerating labor-augmenting technical change, then the restriction $\lambda_N \geq 1$ would specialize to $\lambda_N = 1$.

Technological progress is *Solow neutral* if technology is solely capital-augmenting: $\tilde{F}(N_t, K_t, A_t) = F(N_t, A_t K_t)$. Solow neutrality holds

- in both the short and long run if $\gamma_N = 0$, $\gamma_K > 0$, $\lambda_K \geq 1$;
- only in the long run if $\lambda_N < 1$, $\gamma_K > 0$, $\lambda_K \geq 1$.
- If Solow neutrality is defined to preclude accelerating capital-augmenting technical change, then the restriction $\lambda_K \geq 1$ would specialize to $\lambda_K = 1$.

The ability to distinguish empirically between short-term versus long-term biases in technical change is important. As the introduction noted, short-run capital-augmenting technical change does not violate Uzawa's (1961) steady state growth theorem as long as it disappears in the long run.

Of course, the CES specification that is the maintained hypothesis of the KMW model reduces to Cobb-Douglas under a unitary elasticity of substitution, in which case distinct factor efficiencies are not separately identified and all technological progress can be formulated as labor augmenting. Given that the elasticity of substitution is $\sigma = 1/(1+\rho)$, the restriction $\rho = 0$ or $\sigma = 1$ is also a sufficient condition for the steady state growth theorem to hold.

1.2 The Normalized System Specification

Whereas most of the *economic* issues embodied in the KMW model can be understood in terms of the single-equation CES function (1), identification issues complicate estimation of the model. The literature that bears on this goes back at least to La Grandville (1989) and is best surveyed by Klump, McAdam, and Willman (2012). Identification of the elasticity of substitution and the growth parameters is aided by two things: first, joint estimation of the production function with its implied factor demands; second, parameterization of this system in normalized form. The resulting three-equation system is:[2]

$$\log\left(\frac{w_t N_t}{p_t Y_t}\right) = \log\left(\frac{1-\pi}{1+\mu}\right) + \frac{1-\sigma}{\sigma}\left[\log\left(\frac{Y_t/\bar{Y}}{N_t/\bar{N}}\right) - \log\zeta - g_N(t,\bar{t})\right] \tag{3a}$$

$$\log\left(\frac{q_t K_t}{p_t Y_t}\right) = \log\left(\frac{\pi}{1+\mu}\right) + \frac{1-\sigma}{\sigma}\left[\log\left(\frac{Y_t/\bar{Y}}{K_t/\bar{K}}\right) - \log\zeta - g_K(t,\bar{t})\right] \tag{3b}$$

$$\log\left(\frac{Y_t}{N_t}\right) = \log\left(\frac{\zeta\bar{Y}}{\bar{N}}\right) + g_N(t,\bar{t}) - \frac{\sigma}{1-\sigma}\log\left\{\pi\exp\left[\frac{1-\sigma}{\sigma}\left(g_N(t,\bar{t}) - g_K(t,\bar{t})\right)\right]\left(\frac{K_t/\bar{K}}{N_t/\bar{N}}\right)^{(\sigma-1)/\sigma} + (1-\pi)\right\} \tag{3c}$$

[2] These are equations (6), (7), and (8) of KMW (2007a) with typesetting errors in the third equation corrected. The system appears correctly in KMW (2007b, equs. (3), (4), (5)).

The first two equations are the marginal productivity conditions while, under the maintained hypothesis of constant returns to scale, the third equation expresses the production function in labor intensive form. The expressions $g_i(t, \bar{t})$ are the normalized versions of the Box-Cox growth terms (2), defined as

$$g_i(t, \bar{t}) = \frac{\bar{t}\gamma_i}{\lambda_i}\left[\left(\frac{t}{\bar{t}}\right)^{\lambda_i} - 1\right] \qquad (i = N, K), \tag{4}$$

where $\bar{t}$ is the arithmetic mean of the time trend series. This normalization does not alter the economic interpretations of the growth parameters γ_i, λ_i. Notice, for example, that $\lambda_i = 1$ still yields constant growth, $g_i(t, \bar{t}) = \gamma_i(t - \bar{t})$, just with a redefined (i.e. normalized) time index. Similarly, in the special case of logarithmic growth ($\lambda_i = 0$)[3]

$$g_i(t, \bar{t}) = \frac{\bar{t}\gamma_i}{\lambda_i}\left[\left(\frac{t}{\bar{t}}\right)^{\lambda_i} - 1\right] \to \bar{t}\gamma_i \log(t/\bar{t}) \quad \text{as } \lambda_i \to 0. \tag{5}$$

In the Cobb-Douglas special case of $\sigma = 1$ the marginal productivity conditions (3a) and (3b) reduce to constant factor shares, as they should. This special case of the production function (3c) as $\sigma \to 1$ (or, equivalently, $\rho \to 0$) involves the usual application of L'Hôpital's rule. As a result, the system as a whole has a singularity at $\sigma = 1$: the likelihood function behaves anomalously as $\sigma \to 1$ and the third term of (3c) becomes explosive, contributing to the multiple maxima that KMW found and I verify.

In addition to being parameterized in terms of the elasticity of substitution σ instead of ρ, several parameters appear in the system (3) that do not appear in the original production function (1). The distribution parameter π is the share of capital in total factor payments and so should roughly correspond to the sample mean of $q_tK_t/(w_tN_t + q_tK_t)$, which is 0.224 in the KMW sample. Indeed, for data sets for which nonlinear estimation proves problematic, convergence can be aided by setting π to this sample mean. Like KMW I did not find this necessary, and allow π to be freely estimated.[4]

The parameter μ is a markup that provides for a wedge between GDP and factor payments, allowing for imperfect competition, and should roughly correspond to the mean profit share of $(p_tY_t - w_tN_t - q_tK_t)/p_tY_t$, which is 0.032 in the KMW sample.

Finally, in principle the point of normalization should be a "fixed point" at which baseline values of the factors N and K yield a baseline value of production Y. In practice the geometric means $\bar{N}$, $\bar{K}$, $\bar{Y}$ are used as these baseline values, but the nonlinearity of the system means that $\bar{N}$, $\bar{K}$ will yield $\bar{Y}$ only approximately, not exactly. The "normalization constant" ζ treats this discrepancy. Although it has no particular economic interpretation, ζ will be closer to unity the better the approximation that the sample means provide to a true fixed point of the estimated model.

[3]Footnote a of Table 1 in KMW (2007a) appears to be in error in indicating that this limit is $\gamma_K \log(t - \bar{t})$, which would not be defined for the half of the sample in which $t - \bar{t} < 0$. In the same footnote they indicate that, in estimation, they imposed the restriction $\lambda_i = 0$ by the approximation of setting $\lambda_i = -0.001$ in the Box-Cox function (4). In my replication I instead replace the general expression (4) with its logarithmic special case $\bar{t}\gamma_i \log(t/\bar{t})$.

[4]KMW (2008) and León-Ledesma, McAdam, and Willman (2015) are examples of analyses that experiment with estimating the other parameters using a calibrated π. The latter paper finds that this makes "minimal difference" to the estimates.

2 Results

Figure 1 shows the nested testing structure that arises from successive imposition on the maintained KMW system (3) of restrictions having an economic motivation. Comparing with the KMW estimation results, the maintained model M0 corresponds to column 1.4 of their Table 1, model M45 to their columns 1.1 and 1.2, and the case of *Cobb-Douglas constant growth* to column 1.3.[5] For each model in the structure, Figure 1 reports my estimate of the elasticity of substitution σ and the associated loglikelihood function value $\mathscr{L}$. Both local and global maxima are reported for models in which they were revealed by my searches over alternative starting values. As well, I report Schwarz's Bayesian information criterion (BIC) for the global maximum, which may be useful in comparing nonnested specifications. Nested specifications are most naturally compared with likelihood ratio tests.

2.1 Overview of the Nested Testing Structure

For the most part I was successful in replicating KMW's results, at least in substance. For example, for M45 (the special case of constant rates of technical change for both factors, $\lambda_N = \lambda_K = 1$) KMW report local and global maxima at σ estimates of 0.509 and 0.998, as do I (although I found another local maximum at 1.044, which they may not have reported given that its loglikelihood value is well below the others.)

Even before considering the full estimation results for any one model, several broad conclusions emerge. First, M45 is not the only model having multiple maxima; so do M0, M4, and M5. Furthermore, for each of the multiple maxima models M0, M4, M5, and M45, all but one of the maxima are in the immediate neighborhood of $\sigma = 1$. This suggests that the multiple maxima arise from the singularity of the model at that point, something I study in more detail below. Setting aside the maxima in the neighborhood of $\sigma = 1$, models M0, M4, M5, and M45 yield estimates in the fairly narrow range of 0.50938–0.585358. The estimates of σ yielded by the unique maximum models fall in a considerably broader range, although always below unity.

Second, more complicated models (i.e. the ones embodying the fewest restrictions and therefore having the most parameters to be estimated, beginning with the maintained model M0) tend to be the ones with multiple maxima. The simpler models (i.e. the ones embodying the most restrictions and therefore having the fewest parameters to be estimated, such as the Cobb-Douglas special cases) tend to be the ones having a unique maximum. Nevertheless there are exceptions to this tendency: M45 imposes the strong restrictions of $\lambda_N = \lambda_K = 1$, yet has three maxima.

Third, likelihood ratio tests reject almost all the restricted models. For example, a Cobb-Douglas

[5] Given the evidence in León-Ledesma, McAdam, and Willman (2010) that the system based on the Kmenta approximation is unsuccessful in identifying key parameters, replication results for the Kmenta parameterization are of little interest and so are not reported. Following KMW, I do use the Kmenta approximation to replicate the average TFP growth rates reported in my Tables 1 and 2.

system ($\sigma = 1$, although leaving the unique growth parameters γ and λ unconstrained) eliminates distinct γ_N, γ_K and λ_N, λ_K from the model as well as setting $\sigma = 1$, and so reduces the dimension of the parameter space by three. Relative to the global maximum of the maintained model M0 the likelihood ratio statistic is $LR = 2(253.270 - 235.872) = 34.8$, very strongly rejecting this special case ($\chi^2_{0.01}(3) = 11.34$).

The exception to these rejections is M6, which imposes the single restriction $\lambda_N = \lambda_K$ (a common Box-Cox parameter in the growth functions (4), although still permitting distinct growth parameters γ_N, γ_K). Relative to M0 this yields $LR = 2(253.270 - 252.135) = 2.27$, which does not reject at conventional significance levels ($\chi^2_{0.10}(1) = 2.71$). Given this result, I report my estimation results for M6 in Table 2 and discuss them further below.

Based on a comparison of their global maxima, it might also seem that M45 ($\lambda_N = \lambda_K$=1) is not strongly rejected relative to M0: $LR = 2(253.270 - 250.575) = 5.39$ rejects at a 10% level of significance but not 5%. However I conclude below that maxima in the neighborhood of the point of singularity $\sigma = 1$ should be discounted. Using instead the local maximum of $\mathscr{L} = 226.574$ at $\hat{\sigma} = 0.509$, M45 is decisively rejected—a rejection I comment further on below in considering the estimation results for M6.

Finally, it sometimes doesn't take much in the way of additional restrictions to change the estimate of σ associated with the global maximum quite dramatically. For example, M5 ($\lambda_K = 1$) and M6 ($\lambda_N = \lambda_K$) yield estimates of σ in the range 0.55–0.56, yet imposing these restrictions jointly (model M45) yields $\hat{\sigma} = 0.998$ at the global maximum. This sensitivity seems to reflect the multiple maxima that tend to arise from the singularity of the model at $\sigma = 1$, a conjecture that can be pursued by examining the detailed results for particular models.

2.2 The Maintained Model

Table 1 compares the KMW estimation results for the maintained model M0 (column 1.4 of their Table 1) with those for the two maxima I found. Focusing initially on my global maximum, for the most part I replicate their results, substantively if not identically. I obtain an elasticity of substitution of 0.5458 in comparison with their 0.556, and both are significantly less than unity according to the respective standard errors. The only notable difference is for the Box-Cox parameter λ_K that governs the growth rate of capital-augmenting technology: my estimate is 0.2038 in comparison with their −0.118. However neither of these estimates is remotely statistically significant, and so both sets of results are consistent with logarithmic growth $\lambda_K = 0$. All other coefficients are easily statistically significant at conventional levels in both my and their results.

Few macroeconomists would find believable an aggregate elasticity of substitution between capital and labor much less than about 0.5, so our estimates of around 0.55 are at the low end of the plausible range. For Uzawa's steady state growth theorem to hold in this circumstance, labor-augmenting technology must dominate in the long run. Is this true? Both labor- and capital-augmenting technology grow at positive rates (my estimates of the growth parameters are $\hat{\gamma}_N = 0.0154$, $\hat{\gamma}_K = 0.0036$) that are statistically significant and intuitively plausible. However my estimated Box-Cox parameters of $\hat{\lambda}_N = 0.5736$ and $\hat{\lambda}_K = 0.2038$ are both significantly below unity, indicating decelerating rather than sustained growth in each of the technology factors. Although there are minor differences in these estimates from those of KMW, the qualitative results are the same. Whereas growth in labor-augmenting technology dominates that of capital-augmenting technology ($\hat{\gamma}_N > \hat{\gamma}_K$ and $\hat{\lambda}_N > \hat{\lambda}_K$), and to this extent is consistent with Uzawa's theorem, neither is sustained in the long run.

How seriously should we take this conclusion? In addition to its global maximum of $\mathscr{L} = 253.270$ at $\hat{\sigma} = 0.546$, the maintained model has a local maximum of $\mathscr{L} = 253.031$ at the very different $\hat{\sigma} = 0.987$. It would therefore seem that the data do not strongly favor one set of estimates over the other, and we should consider whether those of the local maximum alter our conclusions.

2.3 The Local Maximum of the Maintained Model

To investigate the multiple maxima of the maintained model in more detail, Figure 2 plots the loglikelihood over the range $0.5 < \sigma < 1.5$. It reveals something that the point estimates do not: the local maximum at $\hat{\sigma} = 0.987$—essentially the Cobb-Douglas special case—is associated with an almost-discrete jump in the loglikelihood as $\sigma \to 1$ from below. This anomalous behavior in the neighborhood of $\sigma = 1$ clearly arises from the singularity of the model at this point.

Consistent with the absence of separately identifiable factor-augmenting technologies that is intrinsic to the Cobb-Douglas functional form, the parameters γ_N, λ_N, γ_K, and λ_K are not well estimated at this local maximum: the estimates reported in Table 1 are all well within one standard error of zero. Furthermore a literal imposition of the Cobb-Douglas special case (but with unrestricted values for the unique growth parameters γ and λ) is plainly rejected by the data: it yields $\mathscr{L} = 235.872$, far below $\mathscr{L}$ for the maintained model.

Given this rejection of $\sigma = 1$, the conclusion is clear that there is no point dwelling on maxima—local or global—in the immediate neighborhood of $\sigma = 1$. Instead they should be seen to be artifacts of the singularity of the CES function at $\sigma = 1$ and the lack of identification of distinct γ_N, γ_K and λ_N, λ_K that accompanies it. Of course this conclusion is specific to this data set, but it serves as a general lesson in other applications of the model: researchers should be wary of maxima in the neighborhood of the point of singularity. This lesson is reinforced by the special case of logarithmic growth in capital-augmenting technology, considered shortly.

In other applications, how should a researcher differentiate between such an artifact of singularity and a case where an elasticity of substitution close to unity is indeed the superior model? In such a case the Cobb-Douglas version of the system should yield a loglikelihood value that is not dramatically different from that of the general model, in contrast to what Figure 2 shows for the KMW data set. Of course, in this instance the general model is unlikely to yield precise estimates of the distinct growth parameters $\gamma_N, \gamma_K, \lambda_N, \lambda_K$, as illustrated by the KMW local maximum.

2.4 The Restricted Model M6: $\lambda_N = \lambda_K$

Are any of our substantive findings altered by adding information to the estimation? We found in the discussion of Figure 1 that, for the most part, the *economic* special cases delineated in that nested testing structure are not particularly supported by the data. The exception is M6, $\lambda_N = \lambda_K$.

My estimation results for M6 are reported in the final column of Table 2, and show that the substantive economic implications of the model are unaltered. (KMW did not estimate this model, and so there are no results of theirs to provide for comparison.) As in the maintained model, the estimates $\hat{\gamma}_N = 0.0152$ and $\hat{\gamma}_K = 0.0042$ are both statistically significant and attach the larger role to labor-augmenting technical change. Nevertheless the joint estimate $\hat{\lambda}_N = \hat{\lambda}_K = 0.5292$ continues to imply decelerating growth in both technology factors, so that growth is not sustained in the long run. The elasticity of substitution of $\hat{\sigma} = 0.5542$ is essentially that of the maintained model.

The joint estimate $\hat{\lambda}_N = \hat{\lambda}_K = 0.5292$ has a standard error of 0.0421, and so neither $\lambda_N = \lambda_K = 0$ (logarithmic growth in both technology factors) nor $\lambda_N = \lambda_K = 1$ (constant growth in both technology factors) is supported. The latter constitutes a more definitive rejection of the special case M45 than may have been evident from a likelihood ratio test, given the multiple maxima that I found for that model.

2.5 Logarithmic Growth in Capital-Augmenting Technology

Although logarithmic growth in both technology factors ($\lambda_N = \lambda_K = 0$) is rejected, the hypothesis is not rejected individually for capital in either my or the original KMW results of Table 1. As discussed in connection with the Box-Cox growth expression (2), the special case $\lambda_K = 0$ has no particular *economic* motivation, and so does not appear in Figure 1. Nevertheless it offers the useful *numerical* simplification of replacing the Box-Cox function (4) with the simpler log expression (5).

The first three columns of Table 2 compare the KMW results for this special case with the global and local maxima that I found. Figure 3 graphs the loglikelihood, which is similar to that of the maintained model (Figure 2) in that there are two maxima, one at $\hat{\sigma} = 0.5458$ and the other at $\hat{\sigma} = 0.9887$. However now it is the latter that is the global maximum, although only very marginally: $\mathscr{L} = 253.026$ versus $\mathscr{L} = 252.727$. Nevertheless, for the same reasons that we dismissed the estimation results for the maintained model M0 in the neighborhood of $\sigma = 1$, it seems sensible to do so again here. The maximum in the neighborhood of $\sigma = 1$ appears to be an artifact of the poor identification of the growth parameters γ_N, λ_N, and γ_K at this point of singularity: only the estimate of γ_K is more than twice its standard error. As well, imposing $\sigma = 1$ yields the

Cobb-Douglas special case that, even for unrestricted γ and λ, has $\mathscr{L} = 235.872$, indicating a clear rejection of this special case.

Focusing then on the local maximum at $\hat{\sigma} = 0.5458$, the coefficient estimates for this model come very close to replicating those reported by KMW. All estimates are statistically significant, so there is no basis for eliminating additional parameters that would simplify the model further. As well, my estimates $\hat{\gamma}_N = 0.0156$ and $\hat{\gamma}_K = 0.0032$ imply sensible rates for each factor-augmenting growth in technology, with that for labor being larger than that for capital. This is particularly so given that growth in capital-augmenting technology decelerates more rapidly ($\lambda_K = 0$) than does labor ($\lambda_N = 0.5923$), and so labor-augmenting technical change is the principal driver of economic growth. Nevertheless that growth is not sustained: $\lambda_N = 0.5923$ is well below two standard errors of unity.

3 Conclusions

For the most part I have found the results of KMW (2007a) to be robust to replication using alternative software, both numerically and substantively. Their key economic finding that labor-augmenting technical change dominates capital-augmenting change, but that neither is sustained in the long run, stands up to variations in the analysis such as the imposition of restrictions supported by the data ($\lambda_N = \lambda_K$ or $\lambda_K = 0$). So does their estimate of the elasticity of substitution of around $\sigma = 0.56$.

Also replicated is their finding that there is not much support for restricted versions of the model. Most of the economic special cases delineated in my Figure 1 are rejected. This confirms that the KMW generalization of CES supply-side systems to permit factor-specific technical change, in a way that distinguishes between short- and long-term effects, is indeed called for by the data and hence a valuable addition to the literature.

One contribution of this replication has been to investigate in some detail the behavior of the likelihood as σ varies over its plausible range. Researchers employing the model should be aware that maxima in the neighborhood of the point of singularity $\sigma = 1$ are often suspect: they tend to be an artifact of the lack of identification of the growth parameters γ_N, γ_K, λ_N, and λ_K. This is easily revealed in empirical application by estimating the Cobb-Douglas special case that imposes $\sigma = 1$, a model that is typically strongly rejected.

References

Bodkin, R., and L. Klein (1967) "Nonlinear estimation of aggregate production functions," *Review of Economics and Statistics* 29, 28–44.

La Grandville, O. de (1989) "In quest of the Slutzky diamond," *American Economic Review* 79, 468–481.

Diamond, P., McFadden, D., and M. Rodriguez (1978) "Measurement of the elasticity of factor substitution and bias of technical change," Chapter IV.2 in *Production Economics: A Dual Approach to Theory and Applications,* ed. by M. Fuss and D. McFadden (North-Holland).

Klump, R., and O. de La Grandville (1989) "Economic growth and the elasticity of substitution: Two theorems and some suggestions," *American Economic Review* 90, 281–291.

Klump, R., McAdam, P., and A. Willman (2004) "Factor substitution and factor-augmenting technical progress in the US: A normalized supply-side system approach," European Central Bank Working Paper no. 367.

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

Klump, R., McAdam, P., and A. Willman (2007b) "The long term sucCESs of the neoclassical growth model," *Oxford Review of Economic Policy* 23, 94–114.

Klump, R., McAdam, P., and A. Willman (2008) "Unwrapping some euro area growth puzzles: Factor substitution, productivity, and unemployment," *Journal of Macroeconomics* 30, 645–666.

Klump, R., McAdam, P., and A. Willman (2012) "The normalized CES production function: Theory and empirics," *Journal of Economic Surveys* 26, 769–799.

León-Ledesma, M., McAdam, P., and A. Willman (2010) "Identifying the elasticity of substitution with biased technical change," *American Economic Review* 100, 1330–1357.

León-Ledesma, M., McAdam, P., and A. Willman (2015) "Production technology estimates and balanced growth," *Oxford Bulletin of Economics and Statistics* 77, 40–65.

Marshak, J., and W. Andrews (1944) "Random simultaneous equations and the theory of production," *Econometrica* 12, 143–205.

McAdam, P., and A. Willman (2013) "Medium run redux," *Macroeconomic Dynamics* 17, 695–727.

McCullough, B.D. (1999) “Econometric Software Reliability: EViews, LIMDEP, SHAZAM, and TSP,” *Journal of Applied Econometrics* 14, 191–202.

Uzawa, H. (1961) “Neutral inventions and the stability of growth equilibrium,” *Review of Economic Studies* XXVIII, 117-124.

Table 1: Maintained Model M0

		Replication	
Parameter	KMW Table 1 Column 1.4	Global maximum	Local maximum
ζ	1.029	1.0269	1.0359
	(0.006)	(0.0041)	(0.0037)
π	0.221	0.2219	0.2182
	(0.009)	(0.0068)	(0.0068)
γ_N	0.015	0.0154	−0.0161
	(0.000)	(0.0004)	(0.0863)
λ_N	0.439	0.5736	−0.3361
	(0.076)	(0.0539)	(0.6994)
γ_K	0.004	0.0036	0.1170
	(0.001)	(0.0006)	(0.3107)
λ_K	−0.118	0.2038	0.0274
	(0.336)	(0.2335)	(0.5043)
σ	0.556	0.5458	0.9866
	(0.018)	(0.0210)	(0.0400)
$1+\mu$	1.042	1.0414	1.0412
	(0.011)	(0.0095)	(0.0097)
Average TFP growth rate	0.013	0.0151	0.0135
Loglikelihood	255.400[a]	253.270	253.031
ADF_N[b]	−4.310	−3.447	−2.922
ADF_K[b]	−3.580	−3.525	−3.454
ADF_Y[b]	−3.960	−3.171	−3.916

[a] Calculated from the log determinant of −19.618 reported by KMW. For $T = 46$ observations and $n = 3$ equations the relationship is $\mathscr{L} = -T[n(1 + \log 2\pi) + (-19.618)]/2 = 255.400$.

[b] ADF regressions include an intercept, no trend, and (in view of the data being annual) one augmenting lag. For 50 observations the associated ADF critical values are −2.93 (5%) and −2.60 (10%), so the equations of the KMW model are generally successful in yielding stationary residuals.

Table 2: Restricted Models Supported by the Data

	Logarithmic growth in capital-augmenting technical change: $\lambda_K = 0$			
		Replication		
Parameter	KMW Table 1 Column 1.5	Local maximum	Global maximum	M6 $\lambda_N = \lambda_K$
ζ	1.029	1.0265	1.0359	1.0279
	(0.006)	(0.0041)	(0.0037)	(0.0040)
π	0.222	0.2213	0.2182	0.2230
	(0.009)	(0.0067)	(0.0068)	(0.0068)
γ_N	0.015	0.0156	−0.0213	0.0152
	(0.000)	(0.0004)	(0.0163)	(0.0004)
λ_N	0.427	0.5923	−0.3059	0.5292
	(0.083)	(0.0493)	(0.1874)	(0.0421)
γ_K	0.004	0.0032	0.1356	0.0042
	(0.000)	(0.0003)	(0.0592)	(0.0005)
λ_K	0[a]	0[b]	0[b]	0.5292
				(0.0421)
σ	0.557	0.5458	0.9887	0.5542
	(0.018)	(0.0208)	(0.0052)	(0.0223)
$1 + \mu$	1.042	1.0414	1.0412	1.0414
	(0.012)	(0.0095)	(0.0097)	(0.0095)
Average TFP growth rate	0.013	0.0151	0.0135	0.0151
Loglikelihood	255.308[c]	252.727	253.026	252.135
ADF_N[d]	−4.360	−3.430	−2.925	−3.485
ADF_K[d]	−3.580	−3.515	−3.453	−3.531
ADF_Y[d]	−3.970	−3.150	−3.921	−3.219

[a] Imposed approximately by setting $\lambda_K = -0.001$.

[b] Imposed by replacing the general Box-Cox function with $\bar{t}\gamma_K \log(t/\bar{t})$, as in equation (5) of the text.

[c] Calculated from the log determinant of −19.614 reported by KMW. For $T = 46$ observations and $n = 3$ equations the relationship is $\mathscr{L} = -T[n(1 + \log 2\pi) + (-19.614)]/2 = 255.308$.

[d] ADF regressions include an intercept, no trend, and (in view of the data being annual) one augmenting lag. For 50 observations the associated ADF critical values are −2.93 (5%) and −2.60 (10%), so the equations of the KMW model are generally successful in yielding stationary residuals.

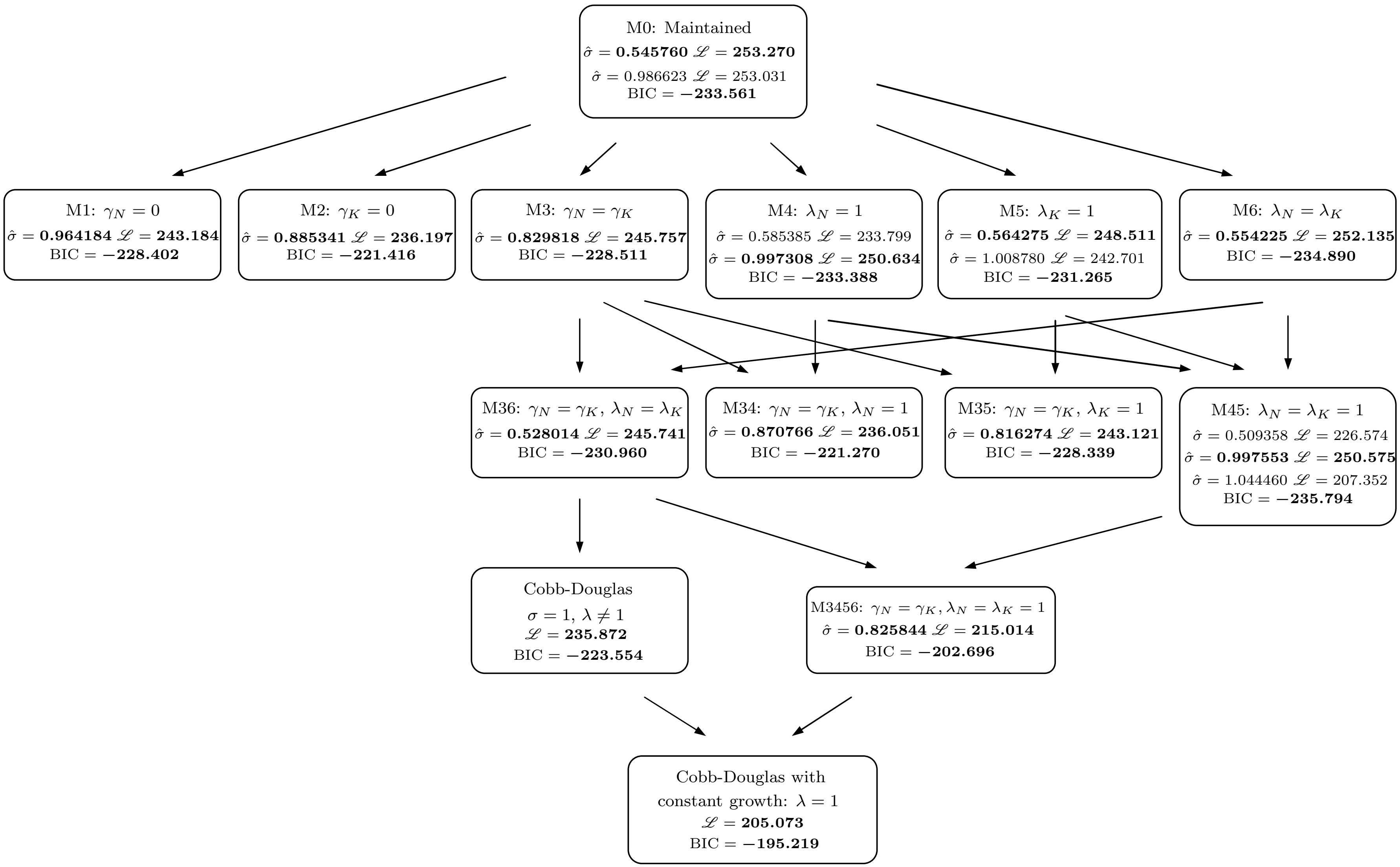

Figure 1: The KMW nested testing structure (unique or global maxima are bolded) [In hardcopy, the readability of this Figure can be improved by setting Orientation to Auto portrait/landscape when printing the paper, which will print this page in landscape format.]

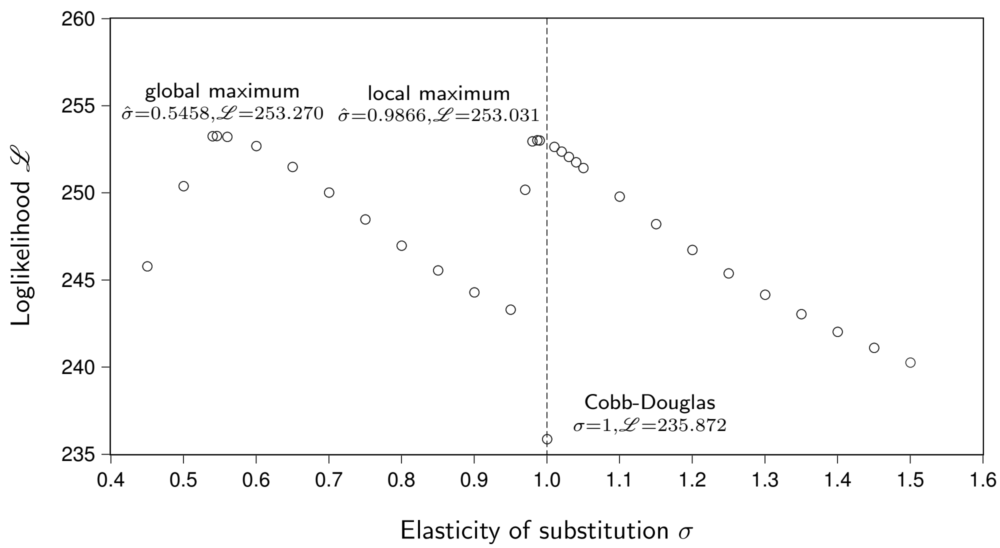

Figure 2: Scatter plot of loglikelihood values for the maintained model M0

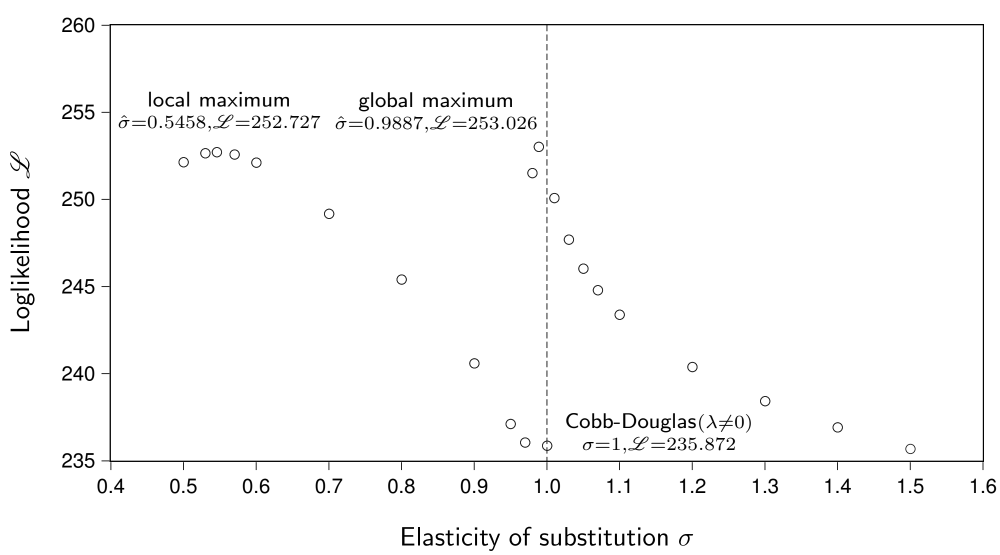

Figure 3: Scatter plot of loglikelihood values for the restricted model: $\lambda_K = 0$