BIASES IN APPROXIMATING LOG PRODUCTION†
KAI SUN, DANIEL J. HENDERSON* AND SUBAL C. KUMBHAKAR
Department of Economics, State University of New York, Binghamton, NY, USA

SUMMARY
Most empirical work in economic growth assumes either a Cobb–Douglas production function expressed in
logs or a log-approximated constant elasticity of substitution specification. Estimates from each are likely
biased due to logging the model and the latter can also suffer from approximation bias. We illustrate this
with a successful replication of Masanjala and Papagerogiou (The Solow model with CES technology:
nonlinearities and parameter heterogeneity, Journal of Applied Econometrics 2004; 19: 171–201) and then
estimate both models in levels to avoid these biases. Our estimation in levels gives results in line with
conventional wisdom. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION
There has been a long-standing tradition of estimating production functions in logs when studying
economic growth. For example, Mankiw et al. (1992) derive the steady state level of output per
worker and estimate the parameters of their model using a log-linearized version of the standard
Cobb–Douglas (CD) production function. Mankiw et al. (1992), and other pioneering papers in
this field, have literally been cited thousands of times and an overwhelming majority of these
papers estimate the parameters of their models via ordinary least squares (OLS), which is possible
when the CD model is expressed in logarithmic form. Although the CD production function is
relatively standard in this literature, Duffy and Papageorgiou (2000) and others suggest using a
constant elasticity of substitution (CES) production function in order to capture nonlinearities in
the growth process and potential heterogeneity across countries.1 CES production functions are
inherently nonlinear, but the authors take logs and linearize the CES production function using a
Taylor series expansion as in Kmenta (1967) in order to employ OLS.2

Both the CD and CES models estimated in logs can produce biased estimates because the
expected value of the logarithm of the error term generally depends upon the regressors. This
point has been argued compellingly by Santos Silva and Tenreyro (2006) in the context of the

† All data and R code for this paper are available on the JAE Data Archive.
1 The Journal of Macroeconomics in 2008 devoted a special issue (Volume 30, Issue 2) to the CES production function
and its impact on the theory and empirics of economic growth.
2 The authors also estimate nonlinear versions of their models using nonlinear least squares. These models do not use
Taylor series expansions, but in their paper they are estimated in logs.

Copyright © 2010 John Wiley & Sons, Ltd.
gravity model of trade (for related discussions see Papke and Wooldridge, 1996; Manning and Mullahy, 2001). A second bias can also arise with CES models when they are approximated. It is well known that the Kmenta approximation can suffer from omitted variable bias. Specifically, this approximation does not guarantee that the underlying CES parameters are consistently estimated because it is a Taylor series expansion and the remainder term becomes an omitted variable in the regression. This point was laid out in Thursby and Lovell (1978) and by others. While these past papers have made an impact in the empirical international trade and productivity literatures, these potential biases appear to be ignored or are not considered to be too problematic by many growth economists (for an exception see Temple, 2001, pp. 914–915).

We showcase each of these biases by replicating Masanjala and Papagerogiou (2004), hereafter MP, who estimate both CD and CES models in logs. We then estimate their models in levels, both with and without using an approximation, via nonlinear least squares (NLLS) and Poisson pseudo-maximum likelihood estimation (PPML). The approximated models avoid the bias due to logging. When we estimate the models in levels without an approximation, we avoid both the bias due to estimating the model in logs as well as the approximation bias.

We should point out that we are making several assumptions when claiming consistency of the estimated parameters of the model. First, we are assuming that the CD or CES model is the correct parametric specification. That being said, we perform functional form specification tests and are unable to reject the models estimated in levels. Second, we assume that there are no omitted variables which may bias our results. In other words, we are completely ignoring the issue of endogeneity in this article. While our proposed estimation strategy may help with biases, estimating cross-country production functions in levels (as opposed to growth rates) potentially amplifies the endogeneity problem, which is severe in this line of work. Third, although we allow for heterogeneity in the shares of physical and human capital in the CES model, we do not allow for parameter heterogeneity. There is good reason to believe that the parameters of the model may vary drastically across different groups of countries. We urge readers to consider the potential impact of each of these caveats, and perhaps others, when interpreting our results. In spite of these potential problems in our estimates, we still advocate for estimating models in levels.

2. BIASES IN GROWTH MODELS

Approximating a function using a Taylor series expansion is well known to economists. What is often ignored in many applications is the bias associated with the remainder term. Since this term is omitted in estimation, the resulting estimators will likely suffer from omitted variable bias (similar to Heckman’s selectivity bias). Consider the CES production function in MP. Given that the CES production function is nonlinear, MP use the Kmenta approximation in order to estimate several of their models. This approximation is based on the elasticity of substitution ($\sigma$) being equal to unity (equations (4) and (6) in MP). The approximation error here (the remainder term in the Taylor series expansion) is non-zero (function of $x$) and likely grows when $\sigma$ deviates from 1. The other source of bias is due to the non-constant (function of $x$) mean of the error term when using logs.

We address biases in the estimates of CES production functions (although the arguments provided are also applicable to other parametric functions) arising from both problems. To

---

3 The remainder term will also vary with the values of the regressors (Thursby and Lovell, 1978).
4 Noting that CD is a special case of CES when $\sigma = 1$, the CD model will not suffer from approximation bias as it does not require the Kmenta approximation.
distinguish between the models in level and logarithmic form, with and without approximation error, we write the CES production function in levels as

\[ y = f(x, \theta) + u \]  \hspace{1cm} (1)

where \( y \) is a scalar output, \( x \) is a vector of inputs and \( \theta \) is a vector of parameters. We assume that \( f(x, \theta) \) has a CES form and that \( E(u|x) = 0 \). The error term \( u \) can be either homoskedastic or heteroskedastic. We can rewrite (1) as \( y \equiv f(x, \theta)\eta \) and express it in logarithmic form:

\[ \ln y = \ln f(x, \theta) + \ln \eta \]  \hspace{1cm} (2)

where \( \eta = (1 + u/f(x, \theta)) \), and \( E(\eta|x) = 1 \). We can also use a linear or quadratic approximation of \( f(x, \theta) \) at \( \sigma = 1 \) and write it as

\[ y = f^0(x, \theta) + r(x, \theta) + u \]  \hspace{1cm} (3)

where \( f^0(x, \theta) \) is a linear or quadratic approximation of \( f(x, \theta) \) at \( \sigma = 1 \) and \( r(x, \theta) \) is the approximation error (it contains higher-order terms in the Taylor series expansion). The models in (1) and (3) are referred to as models in levels.

We can take a linear or quadratic approximation of (2) to get

\[ \ln y = \ln f^0(x, \theta) + R(x, \theta) + \ln \eta \]  \hspace{1cm} (4)

where \( \ln f^0(x, \theta) \) is a log-linear or log-quadratic approximation of \( \ln f(x, \theta) \) at \( \sigma = 1 \) and \( R(x, \theta) \) is the approximation error. We refer to the models in (2) and (4) as the log and log-approximated models, respectively.

There are two sources of bias in estimating (4): (i) the deterministic terms \( R(x, \theta) \) might not be zero (and might not converge to a constant as the sample size increases); and (ii) the expectation of the random term \( E(\ln \eta) \) is almost always a function of \( x \). Each will cause bias and inconsistency if (4) is estimated using least-squares methods. If the approximated non-log model in (3) is estimated using NLLS, it will obviously not suffer a bias due to logging the model, but it will suffer from omitted variable bias (due to excluding the \( r(x, \theta) \) term). Using NLLS or PPML on (1) will avoid biases from both sources.\(^6\)

3. ESTIMATION IN LOGS

We were able to successfully replicate all of the tables and figures in MP. Given the specific interest of this paper, we solely present the replication results for the input shares obtained from the basic Solow growth model (without human capital) and its extended counterpart (with human capital).\(^7\) In the CD version of the model, \( \alpha \) and \( \beta \) are the actual shares of physical and human capital.

---

\(^5\) Expanding \( \ln \eta = \ln[1 + u/f(x, \theta)] = \left[ u/f(x, \theta) - \frac{1}{2} u^2/((f(x, \theta))^2 + \frac{1}{4} u^3/(f(x, \theta))^3 - \ldots \right] \) and then taking expectations of both sides shows that \( E(\ln \eta) \) is a function of \( x \). In the special case where \( u \) is heteroskedastic and the form is such that \( u = f(x, \theta)v \), where \( v \) has zero mean and higher order moments that do not depend on \( x \), then \( E(\ln \eta) \) will be a constant. Otherwise, \( E(\ln \eta) \) will be a function of \( x \) and we will refer to this as the bias associated with logging the model.

\(^6\) Although the consistency of both estimators does not depend on the form of heteroskedasticity of \( u \), the primary difference between the two is that PPML assumes that the variance of \( u \) is proportional to the conditional mean, whereas the NLLS estimator assumes that \( u \) is homoskedastic (Gourieroux et al., 1984).

\(^7\) These correspond to the estimates of equations (1), (2), (4) and (6) in Table I of MP.
Table I. Basic and extended growth models

<table>
<thead>
<tr>
<th>Method</th>
<th>Implied $\alpha$</th>
<th>Implied $\beta$</th>
<th>Extended Method</th>
<th>Implied $\alpha$</th>
<th>Implied $\beta$</th>
<th>Implied $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD OLS (log)</td>
<td>0.5981***</td>
<td></td>
<td>CD NLLS</td>
<td>0.3082***</td>
<td></td>
<td>0.2743***</td>
</tr>
<tr>
<td></td>
<td>(0.0170)</td>
<td></td>
<td></td>
<td>(0.0465)</td>
<td></td>
<td>(0.0356)</td>
</tr>
<tr>
<td>NLLS</td>
<td>0.5907***</td>
<td></td>
<td>NLLS NLLS</td>
<td>0.2845***</td>
<td>0.3091***</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(0.0254)</td>
<td></td>
<td></td>
<td>(0.0831)</td>
<td></td>
<td>(0.0751)</td>
</tr>
<tr>
<td>PPML</td>
<td>0.6057***</td>
<td></td>
<td>PPML PPML</td>
<td>0.3028***</td>
<td>0.3005***</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(0.0165)</td>
<td></td>
<td></td>
<td>(0.0545)</td>
<td></td>
<td>(0.0464)</td>
</tr>
</tbody>
</table>

**CES**

<table>
<thead>
<tr>
<th>Method</th>
<th>Implied $\alpha$</th>
<th>Implied $\beta$</th>
<th>Extended Method</th>
<th>Implied $\alpha$</th>
<th>Implied $\beta$</th>
<th>Implied $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS (log)</td>
<td>0.4984***</td>
<td>1.5425</td>
<td>NLLS (log)</td>
<td>0.2395***</td>
<td>0.3582***</td>
<td>1.1894***</td>
</tr>
<tr>
<td></td>
<td>(0.0500)</td>
<td></td>
<td></td>
<td>(0.0484)</td>
<td></td>
<td>(0.0512)</td>
</tr>
<tr>
<td>NLLS (approximated)</td>
<td>0.7009***</td>
<td>0.9499***</td>
<td>NLLS (approximated)</td>
<td>0.4295***</td>
<td>0.2773***</td>
<td>0.9219***</td>
</tr>
<tr>
<td></td>
<td>(0.0229)</td>
<td></td>
<td></td>
<td>(0.0693)</td>
<td></td>
<td>(0.0611)</td>
</tr>
<tr>
<td>PPML (approximated)</td>
<td>0.6376***</td>
<td>0.9623</td>
<td>PPML (approximated)</td>
<td>0.2796***</td>
<td>0.3140***</td>
<td>1.0590</td>
</tr>
<tr>
<td></td>
<td>(0.0936)</td>
<td></td>
<td></td>
<td>(0.0649)</td>
<td></td>
<td>(0.0564)</td>
</tr>
<tr>
<td>NLLS</td>
<td>0.7486***</td>
<td>0.8354</td>
<td>NLLS</td>
<td>0.3928***</td>
<td>0.2851***</td>
<td>0.8381*</td>
</tr>
<tr>
<td></td>
<td>(0.1192)</td>
<td></td>
<td></td>
<td>(0.1185)</td>
<td></td>
<td>(0.0705)</td>
</tr>
<tr>
<td>PPML</td>
<td>0.6172***</td>
<td>0.9813</td>
<td>PPML</td>
<td>0.2851***</td>
<td>0.3105***</td>
<td>1.0369</td>
</tr>
<tr>
<td></td>
<td>(0.0457)</td>
<td></td>
<td></td>
<td>(0.0650)</td>
<td></td>
<td>(0.0551)</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are heteroskedastic robust standard errors. Asterisks indicate statistical difference from 0 and stars for statistical difference from 1 at the * (*), ** (**), *** (***); 10%, 5%, and 1% level, respectively.

capital, respectively. In the CES case, $\alpha$ and $\beta$ are distribution parameters which can be used to calculate the shares of physical and human capital (MP, pp. 177–179). In the CD model, the elasticity of substitution parameter $\sigma$ is unity and this parameter is allowed to differ from one in the CES specification.

Table I presents the results for the basic and extended Solow models. The replication results correspond to the rows which are listed as ‘OLS (log)’ or ‘NLLS (log)’. Specifically, we report the implied $\alpha$, $\beta$ and $\sigma$. Each of our parameter estimates for the log models are identical to those
Figure 2. Physical and human capital shares in the extended CES model

reported in MP. By construction, the shares for the CD production function are constant across countries. On the other hand, the shares for the CES production function are heterogeneous. MP report the shares for each country for each estimation method. We simply plot the kernel densities of these shares in Figures 1 and 2 for the basic Solow and the extended Solow models, respectively. In Figure 1, the solid line represents the kernel density for the shares of physical capital from the log-approximated CES production function. The estimated density is relatively flat and roughly covers the range 0.3–1.0. The wide range likely represents heterogeneity across countries as the authors argue, but it may also be due to biases in their setup. In Figure 2 we see that the shares for physical capital have a mode near 0.26. Conventional wisdom suggests that these shares should be near one-third and this downward bias is perhaps due to the problems associated with logging the model and the Taylor series expansion as discussed above. For human capital, the log-approximated density extends from roughly 0.2 to 0.4. This partially overlaps the range of 1/3 to 1/2 which some authors consider to be ‘sensible’ (Mankiw et al., 1992).

To get a feel of whether either model is correctly specified, we performed Ramsey RESET tests on each logged model. In each case we reject the null that the proposed model is correctly specified. The results of this section suggest that there may be room for improvement and hence we now consider estimation of the CD and CES production functions in levels.

4. ESTIMATION IN LEVELS

4.1. Cobb–Douglas

In the basic Solow model, the implied share of physical capital ($\alpha$) is relatively constant across estimation procedures. Our initial expectation was that the NLLS and PPML estimates would be significantly different from the log-linear estimates because $E(\ln \eta|x)$ is likely a function of $x$. This unexpected result likely occurs because either $E(\ln \eta|x)$ is close to a constant, the heteroskedasticity of $u$ is close to the specific form mentioned above, or some combination thereof. Given that the parameter estimates from the logged model are similar to those from the NLLS and PPML estimators, some may take this to mean that the CD model estimated in logs is appropriate. We
do not oppose this possibility, but note that it would be impossible to determine this in practice unless the level model is also estimated. Formal tests should be performed before making this conclusion. We also performed the RESET test on the CD model estimated via NLLS and PPML and were unable to reject the null in either case. This suggests that the CD model may be correctly specified and that we should estimate it in levels.

4.2. Constant Elasticity of Substitution

For the basic Solow model, we see large differences between the logged and level methods. The implied value of $\alpha$ for the logged estimator is much smaller than that of any other estimator. It appears that the combination of logging the model and the Taylor series expansion cause biases in these estimates. This is in contrast to the CD production function, where logging the function did not make major differences in the parameter estimates.

Figure 1 clearly shows the difference in each of the estimators. The bias due to logging is large here and the approximation bias appears to be minimal. The lack of a large approximation bias is likely due to the estimated value of $\sigma$ being near unity. Thursby and Lovell (1978) show that the bias will grow as the estimated value of $\sigma$ deviates from unity (the PPML estimate of $\sigma$ is 1.037).

We see larger differences between the NLLS and PPML estimators (approximated or not) in the CES case. However, none of these parameter estimates (across estimation methods) are statistically different from one another. Another interesting result is that we fail to reject the null that $\sigma$ is equal to unity in almost every model (NLLS approximated being the exception). In other words, our models point to CD as the preferred functional form. The most compelling results come from the PPML estimator. The PPML CD and CES implied $\alpha$ and $\sigma$ are not statistically different from one another and are nearly identical.

The models estimated in levels, approximated or not, do not have significant differences in their parameter estimates. The only qualitative difference is that the NLLS estimators reject the null that $\sigma = 1$ and the PPML estimators do not. Thus the evidence in favor of CD is mixed in the extended CES case. The RESET test does not help us pick between level estimators as it again fails to reject any of the models estimated in levels.

In Figure 2 we see that the density of estimated shares obtained from the log-linearized method has a mode less than the hypothesized 1/3 and the approximated NLLS estimator gives physical capital shares larger than one-third. On the other hand, the mode for the shares of physical capital from the level NLLS estimator is centered around one-third. The level and approximated PPML physical capital shares are slightly less than 1/3 and show much less variation. It is unclear which level estimator is preferable in this situation without examining mean squared error, but it should be obvious that both level estimators are intuitively more plausible than the logged model. In contrast, the shares of human capital show a substantial amount of overlap amongst estimators.

In conclusion, for the CD case, for this particular sample of data, given the estimation strategies we have employed, we recommend estimating the CD model in levels with human capital included in the model. In the CES case, the log-approximated model introduces bias. Estimation of CES

---

8 Although the parameter estimates do not differ significantly, there is a substantial decrease in the variability of the physical capital shares when we switch from the NLLS (approximated) to the NLLS estimator.

9 It is worth noting that the PPML estimator has the largest $p$-value for the RESET test.
should, where possible, use a level estimator. We have mixed evidence that the elasticity of substitution is equal to unity here. If it is not different from one, we should switch to the CD model. It would be interesting to see how these results translate to a larger sample with more up-to-date data.

5. CONCLUSION

In this paper we were able to successfully replicate Masanjala and Papageorgiou (2004). We pointed out known potential biases due to logging the model and using Taylor series expansions to approximate functional forms. To avoid these biases, we estimated their models in levels. We showed that with this particular paper elimination of these biases led to conclusions which were in line with conventional wisdom.

We again note that logging a model and approximating a functional form need not lead to significant biases in practice. However, given the computing speed and canned statistical packages available today, the cost of estimating models in levels has been greatly reduced. Therefore, we advocate estimating cross-country production functions in levels as a way to avoid potential and unnecessary biases.

It is, however, worth noting again that our results depend upon several assumptions. When estimating models in levels authors should still be concerned with correct functional form specification, omitted variable bias, parameter heterogeneity as well as other possible biases. Controlling for each of these could potentially change the results of this experiment.

ACKNOWLEDGEMENTS

The authors would like to thank three anonymous referees, the editor Badi Baltagi, Chris Papageorgiou, João Santos Silva and Jon Temple for useful comments which led to an improved version of this article. All remaining errors are the authors’ sole responsibility.

REFERENCES


