Econ Journal Watch,
Volume 4, Number 1,
January 2007, pp 71-78.

Handling Economic Freedom in Growth Regressions: Suggestions for Clarification

JULIO H. COLE [*] AND ROBERT A. LAWSON [**]

A RECENT EXCHANGE BETWEEN LAWSON (2006) AND DE HAAN and Sturm (2006) highlights an important methodological issue in empirical studies of the connection between economic freedom (*EF*) and economic growth.

In a major survey article on the measurement and applications of economic freedom indicators (de Haan, Lundström, and Sturm 2006), the authors criticized, among other things, the tendency in many applied studies to use both the level *and* the change in the *EF* index as regressors in cross-country growth-regressions, advocating instead a specification in which *only* the change in the *EF* index is included. In their view, "studies that jointly employ the level and the change of EF as regressors are suspect" (177).[1]

In his comment, Lawson argues that using the level of the *EF* index (in addition to the change) should not be ruled out *a priori*, but should be decided empirically. He also makes a theoretical case for including both the level *and* the change in *EF* in an initial equation specification. The purpose of this note is to help clarify some of the issues involved in this discussion.

[*] Department of Economics, Universidad Francisco Marroquín. Guatemala City, Guatemala 01010.
[**] Department of Economics, Capital University. Columbus, Ohio 43209-2394.
The authors wish to thank Barry Hirsch and Jac Heckelman for helpful comments.
[1] Full disclosure: Both authors have published papers that de Haan et al. criticize in this regard.

THE SPECIFICATION ISSUE

Consider the following specifications:

(1) $$GROWTH = a_0 + a_1 EF_0 + a_2 \Delta EF + a_3 Z$$

(2) $$GROWTH = b_0 + b_1 EF_0 + b_2 EF_1 + b_3 Z$$

(3) $$GROWTH = c_0 + c_1 EF_1 + c_2 Z$$

(4) $$GROWTH = d_0 + d_1 \Delta EF + d_2 Z$$

where *GROWTH* is the rate of economic growth over some period, EF_0 is the economic freedom index at the beginning of the period, EF_1 is the economic freedom index at the end of the period, *Z* is a matrix of control variables (i.e., other variables affecting *GROWTH* over the period), and $\Delta EF = EF_1 - EF_0$ by definition. Equation (1) is Lawson's preferred specification, while de Haan et al. favor Equation (4).

The theoretical case for Equation (1) is straightforward: If *EF* matters at all for economic growth, then we would expect that, other things equal (i.e., holding ΔEF and the *Z* variables constant), countries with higher initial levels of *EF* should grow faster than countries with lower initial levels. This would show up as a positive and significant estimate for a_1 in Equation (1). On the other hand, two countries might have the same initial values for *EF* (and similar values for the *Z* variables), except that in one country *EF* is increasing over the sample period ($\Delta EF > 0$) while in the other it is decreasing ($\Delta EF < 0$). In that case, and if *EF* indeed matters for growth, then we would expect the former country to have better economic performance than the latter. This would show up as a positive and significant estimate for a_2 in Equation (1).

Both Lawson and de Haan et al. agree that $a_2 > 0$. The difference is in the treatment of a_1. Lawson favors including EF_0 in the regression (and predicts that $a_1 > 0$), while de Haan et al. insist on excluding EF_0 (thereby implicitly assuming that $a_1 = 0$).[2]

Given that $\Delta EF = EF_1 - EF_0$, Equation (2) is formally identical to Equation (1), with $b_1 = a_1 - a_2$ and $b_2 = a_2$. There is no difference between

[2] It is not entirely clear that de Haan et al. actually believe that initial *EF* has no growth-impact whatsoever, or whether their favored approached is premised exclusively on grounds of econometric methodology.

estimating (1) or (2). Estimated coefficients for the two regressions will satisfy these relationships, while the constant and the coefficients for the other regressors will be identical in both regressions, as well as the R^2 and other regression statistics. The only difference is in the standard error for b_1, the coefficient for EF_0 in Equation (2), which is of course related to the results for Equation (1) by

$$Var(b_1) = Var(a_1 - a_2) = Var(a_1) + Var(a_2) - 2Cov(a_1, a_2)$$

where $Var(b_1)$ is the sample variance for b_1, $Var(a_1)$ and $Var(a_2)$ are the sample variances for a_1 and a_2 from Equation (1), and $Cov(a_1,a_2)$ is the covariance between a_1 and a_2. From a strictly econometric point of view there is nothing to choose between these two regressions, because *they are the same regression*.

There seems to be some confusion about this point when Lawson (2006) writes:

> I agree that Equation (2) is still a problematic specification because the level of EF at the beginning of the period is likely to be highly collinear with the level of EF at the end of the period. This high degree of collinearity between EF_0 and EF_1 will make the coefficient estimates in Equation (2) difficult to interpret and will bias the standard errors. (403)

Equation (2) might indeed be more difficult to interpret, but only because it is a somewhat unorthodox way to express Equation (1), and not because of collinearity or any other statistical problems present in (2) but not in (1).[3] If classical OLS assumptions are satisfied in one equation, they will also be satisfied in the other, and both equations will yield unbiased estimates of their respective coefficients.[4]

---

[3] Lawson agrees entirely with this. The primary point was that Equation (2) is simply difficult to interpret economically relative to Equation (1).

[4] If a_1 and a_2 are about the same order of magnitude, then b_1 in Equation (2) might be small (or even negative). At first glance, one might think that this is due to the fact that $EF_1 = EF_0 + \Delta EF$, and therefore EF_0 does not add any additional information to the regression. This is not, however, a matter of statistics, but of simple numerics. The results for EF_0 in Equation (2) depend upon the sign of the coefficient for ΔEF in Equation (1). If, in some other context, ΔEF happened to have a negative coefficient in a regression similar in form to Equation (1), then EF_0 would have a "large" coefficient in the equivalent Equation (2) form. In any case, the fact that EF_0 has a small (or negative) coefficient in Equation (2) does not

The fact that (1) and (2) are formally identical seems to be the essence of the de Haan-Sturm case against including the level of *EF* in a growth-regression. Their argument is, in effect, a two-step one: First discredit (2), and then discredit (1) by implication.

The problem, in their view, is that (2) includes EF_1, the value of *EF* at the end of the sample period. To understand why they think this is a problem, first consider Equation (3), in which *GROWTH* depends upon end-period *EF* but not upon initial *EF*. For some reason, there seems to be unanimous agreement that (3) is an inadmissible specification. Lawson states:

> Clearly, [Equation (3)] is inappropriate as it is logically impossible for the level of economic freedom at the end of the period to affect economic growth in the previous period. Something occurring today cannot determine what happened yesterday. (402)

And de Haan and Sturm concur:

> We all agree that [Equation (3)] does not make theoretical sense: the level of economic freedom at the end of the sample period cannot explain economic growth experienced over the sample. (409)

Hence, they conclude, "if [Equation (3)] is to be considered nonsensical, then . . . Lawson's proposed equation is as well, and for the same reason" (410). That is, if (3) makes no sense, then (2) makes no sense either, since it also includes EF_1 as a regressor, and therefore we should reject (1) as well, since it is formally identical to (2). That leaves us with Equation (4), which includes only ΔEF.

But is it really true that (3) is "logically impossible" and "does not make theoretical sense"? After all, EF_1 is simply EF_0 plus ΔEF, so (3) equals

$$GROWTH = c_0 + c_1 EF_0 + c_1 \Delta EF + c_2 Z$$

imply that this variable has a small (or negative) effect on *GROWTH*, since the full effect of cross-country differences in EF_0 equals $b_1 + b_2$ (= a_1 in Equation (1)).

and is therefore simply a restricted version of Equation (1) with $a_1 = a_2$. Thus, (3) simply states that a one-unit cross-country difference in initial *EF* and a one-unit cross-country difference in ΔEF have the same effect on *GROWTH*. Now, this restriction seems arbitrary, and might well be an invalid assumption, but that does not make it somehow "nonsensical," and in any case it is something that should be decided empirically, not on theoretical grounds.[5]

Nor is the de Haan-Sturm "solution" any better, since Equation (4) can be expressed as

$$GROWTH = d_0 - d_1 EF_0 + d_1 EF_1 + d_2 Z$$

which, as Lawson rightly notes (404), is simply a restricted version of Equation (2) with $b_1 = -b_2$. Thus, if de Haan and Sturm reject Equation (2)—and by implication, Equation (1)—because it includes EF_1 as a regressor, then they should reject their own preferred equation as well, since it includes EF_1 too.

The point is that the proper specification for a growth regression that includes economic freedom should be driven first by theory, and any restrictions placed on the regression should be subject to empirical testing. We still maintain that the de Haan-Sturm specification implies a restriction that cannot be justified statistically and that the omission of the level of *EF* represents a potentially serious omitted variable problem.

THE ENDOGENEITY ISSUE

To be fair, we are perhaps talking at cross-purposes here. Lawson's primary issue in his comment was about the proper *specification* of the growth regression. De Haan and Sturm in contrast make much of *reverse-causation* and *endogeneity*—issues that affect all cross-country growth regressions of the sort de Haan, Lundström and Sturm (2006) surveyed. The problem is that a strong empirical correlation between, say, *GROWTH*

[5] Lawson now concedes that the problem isn't the logical invalidity of (3) but rather its statistical invalidity in that the implied restriction cannot be supported with standard statistical testing.

and ΔEF, does not necessarily imply that the direction of causality is $\Delta EF \rightarrow GROWTH$, since it could be the other way around (reverse-causality), or both ΔEF and $GROWTH$ could be responding to some other factor (endogeneity).

What we fail to see is why this should be less of a problem in Equation (4) than in the other three specifications since ΔEF (and implicitly EF_0 and EF_1) is in Equation (4) as well. If there is a reverse-causality problem in Equation (1) or (2), then it is likely still to be present in Equation (4). Dropping a single variable, EF_0, is not likely to eliminate magically any reverse-causality or endogeneity problems associated with ΔEF and $GROWTH$.

If reverse-causality and/or endogeneity are the real bone of contention, then we see three solutions, each with problems of their own. (1) We could discard ΔEF altogether and estimate a model containing only the initial-period EF; after all, nobody is suggesting that economic growth over the sample period somehow "determines" initial EF. The problem here, however, is that if Equation (1) is really the true model, as Lawson suggests, then discarding ΔEF will create an omitted variable problem, biasing the estimated coefficient for EF_0. (2) We could use the level of EF and the ΔEF from a period *before* the growth period under investigation to test for reverse causality (see Gwartney et al., 2006). This approach faces data limitations and still fails to account for possible endogeneity. (3) We could dispense with single-equation estimation altogether and move to an instrumental variables (IV) approach. While this method may deal with the problem, IV models themselves invite a whole host of criticisms regarding what the proper instruments should be, and the results are especially fragile to the choice of instruments.

Endogeneity and reverse-causation are, to be sure, important and thorny issues, but they are present in practically any kind of econometric analysis, especially in these reduced-form cross-country growth regressions. Indeed, that "correlation does not imply causation" is something that statisticians and econometricians have known for ages and not something that we all suddenly realized after we learned how to say "endogeneity." This just means that regression models only provide measures of the degree of statistical association between variables, and that inferences regarding causality require an *interpretation* of the results in terms of prior theory. Econometric technique *per se* will often provide little guidance in this respect, especially in the kind of cross-section studies we are concerned with here.

REFERENCES

De Haan, Jakob, Susanna Lundström and Jan-Egbert Sturm. 2006. Market-Oriented Institutions and Policies and Economic Growth: A Critical Survey. *Journal of Economic Surveys* 20(2): 157-181.

De Haan, Jakob and Jan-Egbert Sturm. 2006. How to Handle Economic Freedom: Reply to Lawson. *Econ Journal Watch* 3(3): 407-410.

Gwartney, James, Randall Holcombe, and Robert Lawson. 2006. Institutions and the Impact of Investment on Growth. *Kyklos* 59(2): 255-76.

Lawson, Robert A. 2006. On Testing the Connection between Economic Freedom and Growth. *Econ Journal Watch* 3(3): 398-406.

ABOUT THE AUTHORS

Julio H. Cole is Professor of Economics at Universidad Francisco Marroquín, in Guatemala, and editor of the journal *Laissez-Faire*. He earned both his undergraduate degree (*licenciatura*) and his Ph.D. in economics from Universidad Francisco Marroquín. He is the author of several books, including *Latin American Inflation: Theoretical Interpretations and Empirical Results* (1987), and *Elementos de Econometría Aplicada* (2nd ed., 2006), and has published articles in *Journal of Economic Growth*, *Banca Central*, *Journal of Libertarian Studies*, *Cato Journal*, *Ciencias Económicas* and *Independent Review*. Email: jhcole@ufm.edu.gt.

Robert A. Lawson is Professor of Economics and George H. Moor Chair in the School of Management at Capital University in Columbus, Ohio. He earned his B.S. in Economics from the Honors Tutorial College at Ohio University and his M.S. and Ph.D. in economics from Florida State University. Lawson is co-author of the *Economic Freedom of the World* reports and has professional publications in *Public Choice*, *Cato Journal*, *Kyklos*, *Journal of Labor Research*, *Journal of Institutional and Theoretical Economics*, and *European Journal of Political Economy*. Email: rlawson@capital.edu.

GO TO SECOND REPLY BY DE HAAN AND STURM (2007)
GO TO ORIGINAL COMMENT BY LAWSON (2006)
GO TO FIRST REPLY BY DE HAAN AND STRUM (2006)