



On Whether the Size of Government Belongs in Economic Freedom Indices

Ryan H. Murphy¹

[LINK TO ABSTRACT](#)

Jan Ott (2018) proposes that measures of economic freedom can be “improved” by ceasing to include measures of the size of government. Applying the concept of convergent validity, he shows that including size of government when measuring economic freedom yields a value of Cronbach’s alpha that he finds disturbingly low. Convergent validity and Cronbach’s alpha, which evaluate how well different measures seem to be measuring the same thing, are both taken from the broader literature on construct validity. In Ott’s analysis, lopping off the size of government yields a far improved evaluation of economic freedom indices as measuring a single thing. If it were true that we ought to lop off size of government, it would lead to a considerable shift in evaluating which countries are the most economically free.

This paper criticizes Ott’s suggestion. First, the value of Cronbach’s alpha that Ott finds is based on an older version of economic freedom data² and is based on very specific years of data. Second, the construction of a metric of economic freedom is a conceptual matter, and it is unclear why a purely statistical criterion such as Cronbach’s alpha should dictate how we select variables; the diagnostic and approach Ott uses are applicable to what are known as reflective constructs, not formative constructs. Measures of economic freedom are formative constructs.

1. Bridwell Institute for Economic Freedom, Cox School of Business, Southern Methodist University, Dallas, TX 75275.

2. While Ott (2018) considers both the Heritage *Index of Economic Freedom* and the Fraser Institute’s *Economic Freedom of the World* (Gwartney et al. 2021), we will focus on the Fraser index.

Third, if we were to apply Ott’s methodology to other sets of institutions and other indices in use, we would quickly reach untenable conclusions.

Ott’s article was published in the journal *Social Indicators Research*. I submitted a comment along the lines of the present article to that journal, but it was summarily turned away.

While Ott (2018) will be the focus of this analysis, it is in a long line of a family of criticisms and re-framings of data on economic freedom. Some research has analyzed the economic freedom data and suggested that the size of government should be treated differently because of its weak correlation with the rest of the index (Heckelman and Stroup 2005; Rode and Coll 2012; Bjørnskov 2016) and others have raised similar issues (Heckelman and Stroup 2000; Carlsson and Lundstrom 2002; Justesen 2008). Summarizing this position, Christian Bjørnskov (2016, 15) states that “spending and revenue components tend to only be weakly associated with other elements and are therefore a separable dimension” of institutions. Even if there is more than one dimension underlying economic freedom, it is inappropriate to remove a dimension from a formative construct, as we will see.

Others have suggested that weighting or manipulating the data differently will improve its in-sample explanatory power; these concerns often focus on the size of government.³ While not all of this literature presents itself as opposed to the current choices in methodology for economic freedom indices, taken together, that literature constitutes perhaps the most substantial academic criticism of the EFW index. That literature has also begun leaking into the public policy world as a criticism (see, e.g., Hammond 2018, 7ff.). What will be explored therefore speaks to the literature on economic freedom indices generally, not only to Ott (2018).

Convergent validity and economic freedom

The *Economic Freedom of the World* (EFW) index (Gwartney et al. 2021)

3. Some of the emphasis found in Ott (2018) is also on simple correlations between institutional measures and outcomes (especially happiness). Reconstructing the index using such correlations have previously been attempted by Heckelman and Stroup (2000), who create weights (including negative weights) from a hedonic regression. Applying weights in this manner is the primary methodology used in the subnational index, *Freedom in the 50 States* (Ruger and Sorens 2018), although they do not apply any negative weights. Huskinson and Lawson (2014) apply clustering analysis to the data, and found that countries with high degrees of economic freedom are bifurcated between those with small governments and large governments. On the surface, countries with high economic freedom and large governments perform better in measures of economic development. While there may be some use in performing these exercises, the analysis found here will suggest limits on what they can tell us. Diagma and Vallee (2021) use principal component analysis for suggesting a different way of weighting the Heritage index.

measures economic freedom for up to 165 countries, going back as far as the year 1970 in the main dataset. The data provides scorings once every five years from 1970–2000 and yearly from 2000–2019. There are five “areas” of economic freedom: the (limited) size of government, the quality of the legal system and property rights, sound money, the freedom to trade internationally, and (limited) regulation. Underlying each area are numerous components and subcomponents. All data is placed on a zero to ten scale with ten always corresponding to “more freedom.” The index has been shown to be robustly related to growth (De Haan et al. 2006) and generally correlates with positive social outcomes (Hall and Lawson 2014). The size of government area, in its present form, is composed of government consumption as a percentage of total consumption, transfers and subsidies as a percentage of GDP, government investment as a percentage of total investment, the top marginal tax rate and the level at which it applies, and government ownership of the economy.

Ott (2018) has a trio of research questions that he hopes to answer at the intersection of happiness research and economic freedom. The second of these questions is what is of interest—whether economic freedom indices can be improved, primarily by using the concept of convergent validity as its method. More broadly, validity in the statistical sense is concerned with whether a statistical construct is measuring what it purports to be measuring. While Ott assesses both EFW and another measure, that from the Heritage Foundation, I focus on EFW, which is more significant in the academic sphere and with which I am associated. Ott’s ultimate conclusion is that measures of the size of government should be omitted from economic freedom indices because they greatly reduce Cronbach’s alpha, an assessment of convergent validity: lopping it off would yield a value of over 0.8 (which Ott classifies as “good”), whereas the value for the index as-is, as measured by Ott, is below 0.7 (the cutoff Ott gives as being “acceptable”).

For all facets of his analyses, Ott (2018, 482) uses the years 2010–2012 in considering his hypotheses. Ending in year 2012 implies that the version of the data report published in 2014 was used. Additionally, Ott apparently omits about⁴ 25 countries from the analyses, as he applies a consistent set of countries across the different exercises he performs in the paper. The country list has not been reproduced, but results with the full list of countries differs immaterially from what Ott observes. These findings are reproduced in Table 1.

My first concern about Ott’s findings is that if the sample is not restricted to 2010–2012 and instead the sample from 1970–2012 is used, Cronbach’s alpha for EFW immediately rises to 0.73, clearly crossing the first threshold given by Ott.

4. Under Table 3 in Ott (2018) it says the analyses in the table use “125–127 nations.” In this version of the Fraser data, 152 countries were assigned economic freedom ratings.

Restricting the sample to more recent years is a defensible position, but it is not explicated or justified by Ott (besides it simply matching the sample of countries and years from his other analyses in the paper). Truncating the sample may purge a large number of data points where large governments were observed to coincide with bad institutional quality, particularly those countries pursuing socialism in the period before the fall of the Soviet Union. Because of data issues, there were only so many data points before the fall of the Soviet Union in this version of the data set, so if anything, the historical instances of large governments coinciding with other measures of bad institutions remain underemphasized in the data.

TABLE 1. Ott (2018) result for Cronbach's alpha, as reported and as reproduced

	EFW	EFW, no size of government	EFW, size of government reversed
Ott (2018)	0.66	0.85	0.76
As reproduced	0.67	0.83	0.73

Since the 2014 publication of the 2012 data, there have been several methodological changes to the data. An adjustment was put in place for the legal system and property rights area, which adjusts countries downwards when there is disparate treatment of women and their economic rights (Fike 2017). Another important update was the inclusion of data on government ownership of the economy, which addresses particularly conspicuous interventions in otherwise economically free Hong Kong and Singapore. The variable is slotted into the size of government area, with the data's source originating in the *Varieties of Democracy* dataset (Coppedge et al. 2021). The legal system and property rights area has been buttressed with historical data from *Varieties of Democracy* as well as data on the rule of law from Drew Linzer and Jeffrey Staton (2015), making ratings for this area before 2000 far more robust. These and further changes have improved the precision of the index, especially for years prior to 2000.

If we simply calculate Cronbach's alpha for all years 1970–2019 using the current index, as is found in Table 2, it is now 0.77, comfortably higher than Ott's "acceptable" threshold of 0.7 and not that distant from the "good" threshold of 0.8. As for Ott's narrow period of 2010–2012, with the revised data it is 0.75. If we look at the full period 2000–2019 (with 2000 as a cut-off when data had strongly improved in quality), Cronbach's alpha holds steady at 0.76. If we look at just the data for the year 2019, which is the most recent at the time this writing, it is also 0.76.

TABLE 2. Further recalculations of Cronbach's alpha, under various assumptions

	EFW	EFW, no size of government	EFW, size of government reversed
Current EFW data, all years	0.77	0.85	0.72
Current EFW data, 2010–2012	0.75	0.86	0.74
Current EFW data, 2000–2019	0.76	0.87	0.75
Current EFW data, 2019	0.76	0.87	0.78

Ott (2018) asks whether the size-of-government data could be reversed because doing so also would improve Cronbach's alpha, though he doesn't conclude this change should be implemented. With the current version of the data, reversing the sign of size of government no longer improves Cronbach's alpha, and actually reduces it for all of the periods in question except for the analysis using only the data for the year 2019. It remains true that eliminating the size of government causes Cronbach's alpha to creep higher still, but when the full data set is examined, EFW is above the first threshold and approaching the second. One would only conclude to exclude the size of government if one prefers improved convergent validity lexically above any other concern we may have in constructing an index.

Validity and economic freedom

'Validity' asks whether a statistical construct is measuring what it claims to be measuring. 'Reliability' asks how well indicators are zeroing in on a single concept and about the internal consistency among those indicators (Coaley 2010, 128–129). Cronbach's alpha is primarily about measuring reliability, with convergent validity being one method among many for assessing validity. "Although [Cronbach's alpha] is a useful tool for summarizing the internal consistency of items on a scale as a measure of reliability, reliability is necessary but not sufficient evidence of validity" (Flake et al. 2017, 6).

Jessica Flake, Jolynn Pek, and Eric Hehman have criticized the use of Cronbach's alpha as a means of choosing which variables to include as one of the main misuses of the tool in the literature: "The heavy reliance on [Cronbach's alpha] also suggests that researchers are using it as a criterion for scale use and even item selection. Indeed, we noted numerous instances in which [Cronbach's alpha] was reported to justify item removal. Reliability is important to consider in construct validation, but *it should not be maximized at the expense of other evidence*" (Flake et al. 2017, 6, emphasis added). There are several other dimensions of validity to consider besides convergent validity, among them content validity, which importantly asks whether every dimension of a theory is being captured by the construct.

In any case, Ott misapplies these methodologies because he implicitly assumes that *Economic Freedom of the World* is a reflective construct. EFW is a formative (causal) construct. In a reflective construct, it is assumed that there is some concept out there which has a set of observable effects. The reflective construct aggregates these effects into a single construct. In a formative construct, this relationship is reversed. Each of the individual components of the concept *define* or *cause* the concept being measured. In the case of EFW, those components are the size of government, the quality of the legal system and property rights, sound money, the freedom to trade internationally, and regulation. (See Bollen and Diamantopoulos 2017a for a lengthy discussion of formative models.)

Indicators in a reflective model are all expected to be rather strongly correlated with one another. They are all effects of the same latent cause. In contrast, “causal indicators of the same concept can have positive, negative, or no correlation” (Bollen and Lennox 1991, 307). In a reflective model, one can think of different items to include as being a sample of the causal indicators. But one *cannot* do this in a formative construct. The list of enunciated concepts defining the construct is better thought of as a census. “Omitting an indicator is omitting part of the construct” (ibid., 308). To use other terminology, convergent validity is irrelevant and content validity is essential for formative constructs. What Ott is asking us to do is analogous to applying convergent validity to the human body, finding that the human head is poorly correlated with the rest of the body, and concluding we should chop it off because it doesn’t fit with the rest of the body.

Ott at one point implies that there is in fact no theory underlying the concept of economic freedom. “It is however debatable to use taxation and government spending as positive indicators for economic freedom. There is no theoretical justification; government spending, consumption and transfers and subsidies included, can be directed at very different policies, liberal or less liberal!” (2018, 485). Ott’s claim that there is “no theoretical justification” is a dismissal of the entire classical liberal tradition of social philosophy (see Rothbard 1973; Friedman 1962, 177–195; White 2012, 332–359, 382–412). And if a certain level of government spending is needed to support the liberal society and free markets, we can note that the index does not punish government consumption below six percent of total consumption, which seems to be a sufficient level for the provision of basic law, courts, and national defense (cf. Gwartney et al. 1998).

Economic freedom, the size of government, and other measures of institutions

We should also think about what Ott's method implies, and ask: What if we apply the strictures of convergent validity to other measures of institutions? For one, we could apply it to the five components of the size of government itself. Cronbach's alpha for the five size-of-government EFW components is 0.220. Does this mean we should say that certain kinds of government spending, taxation, or ownership do not count as the size of government because convergent validity indicates that they are not measuring the same concept?

As another example, consider the *Varieties of Democracy* dataset. Its baseline definition of democracy is "electoral democracy," which it then combines with other democratic or democracy-adjacent concepts to create measures of liberal democracy, participatory democracy, deliberative democracy, and egalitarian democracy. Electoral democracy is in turn composed of five mid-level indices which are meant to capture the core of the idea of democracy—freedom of expression, freedom of association, universal suffrage, free and fair elections, and elected officials. In the 2020 version of the data, the Cronbach's alpha for these five variables is 0.827. If we include all their historical data, it rises to 0.885. But if we remove the variables that are less tightly correlated with the other democracy variables, universal suffrage and elected officials, Cronbach's alpha rises to 0.940. Have we improved the measurement of democracy by increasing Cronbach's alpha by ignoring whether there is universal suffrage in a country?

The answer to these questions is no. The reason why the answer is no is because arbitrarily removing certain kinds of spending from the size of government or universal suffrage from a measure of democracy means that you are picking apart necessary, definitional dimensions of a formative construct and no longer are adequately addressing content validity of what we have defined theoretically at the outset of examining the question. The same is true for the size of government and economic freedom.

Discussion and conclusion

Ironically, a careful consideration of *Economic Freedom of the World's* nature as a formative construct actually demands that we do *not* remove size of government from the index, because doing so would leave us a set of measures that is missing an

important aspect of classical liberal economic institutions—limited government. The point applies not only to Ott (2018), but to others who have suggested that the size of government should be removed from economic freedom because it behaves somewhat differently than the rest of the index (e.g., Heckelman and Stroup 2005; Rode and Coll 2012; Bjørnskov 2016; cf. Heckelman and Stroup 2000; Carlsson and Lundstrom 2002; Justesen 2008). It also raises questions concerning what relationship the results of applications of principal component analysis, factor analysis, or clustering analysis even mean in relation to economic freedom (as in Caudill et al. 2000; Sturm et al. 2002; Rode and Coll 2012; Huskinson and Lawson 2014; Diagla and Vallee 2021).

That is not to say that one cannot disaggregate and restructure the data; Scott Sumner (2010) previously adjusted economic freedom data in a similar manner to Ott (2018), but chose to call economic freedom data without the size of government “neoliberalism” and economic freedom data with government spending inverted “egalitarian neoliberalism.” Whether such simple adjustments to the index conceptually capture all the dimensions of (i.e., perform a ‘census’ for) “neoliberalism” and “egalitarian neoliberalism” would then be a further question to pose. But if it is simply argued that neoliberal institutions are better for human welfare than classical liberal economic institutions (economic freedom), that is a hypothesis to be investigated and debated. It isn’t a reason to use convergent validity in order to cease testing of economic freedom as it has been conventionally defined.

In fact, if we are using these methods as bases for improving *Economic Freedom of the World*, content validity would suggest that several elements of regulation data are currently lacking.⁵ There is no data in the index on environmental or antitrust regulation; and “credit market regulation,” the closest variables to financial regulation in the index, mostly features rather weak proxies for such kinds of regulation. One could plausibly argue that these kinds of regulations, unlike other elements of intervention, have clearer pathways for having positive impacts on economic performance. However, that is not a reasonable rationale for excluding them (though that is not the reason why they are excluded; there simply is not good data available that is sufficiently dense in its coverage).

However, as alluded to above, disaggregating may be desirable for an entirely separate reason—to investigate the origins of economic freedom. The existing literature on the origins of economic freedom is principally concerned with the

5. While these are areas for improvement in the index, Bollen and Diamantopoulos (2017b, 606–607) do caution not to take “census” too literally. The issue is whether lacking a dimension will bias the construct in one direction or another. Leaving out the size of government clearly changes the results and therefore its exclusion would lead to bias.

effects on the overall index (Lawson et al. 2020). The origins of property rights, a high-quality legal system, sound money, free trade, and light regulation may be the same as general measures of institutional quality. The origins of the size of government may differ. We may well be concerned with the net, overall effects on economic freedom, but we may add a certain amount of nuance for this specific question. The main point is to not remove the size of government when one believes the hypothesis being tested concerns economic freedom.

Ultimately, the relationships among the size of government, other dimensions of economic freedom, and other variables remain opaque and understudied. There was a point in time that simply looking at the size of government was thought of as a reasonable first approximation of how much economic freedom there was in a country (e.g., Peltzman 1980), and it wasn't understood this was not the case until the projects to measure economic freedom were well underway (Gwartney 2009). Yet it is wholly inappropriate to remove the size of government from our definition of economic freedom merely because the world is complex.

Data and code

Data and code used in this research is available from the journal website ([link](#)).

References

- Bjørnskov, Christian.** 2016. Economic Freedom and Economic Crises. *European Journal of Political Economy* 45: 11–23.
- Bollen, Kenneth A., and Adamantios Diamantopolus.** 2017a. In Defense of Causal-Formative Indicators: A Minority Report. *Psychological Methods* 22(3): 581–596.
- Bollen, Kenneth A., and Adamantios Diamantopolus.** 2017b. Notes on Measurement Theory for Causal-Formative Indicators: A Reply to Hardin. *Psychological Methods* 22(3): 605–608.
- Bollen, Kenneth A., and Richard Lennox.** 1991. Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin* 110(2): 305–314.
- Carlsson, Fredrik, and Susanna Lundstrom.** 2002. Economic Freedom and Growth: Decomposing the Effects. *Public Choice* 112: 335–344.
- Caudill, Steven B., Fenderando C. Zanella, and Franklin G. Mixon.** 2000. Is Economic Freedom One Dimensional? A Factor Analysis of Some Common Measures of Economic Freedom. *Journal of Economic Development* 25(1): 17–40.
- Coaley, Keith.** 2010. *An Introduction to Psychological Assessment and Psychometrics*. London:

SAGE.

- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Anna Lührmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt.** 2021. *V-Dem Codebook v11.1*. Gothenburg, Sweden: V-Dem Institute.
- De Haan, Jakob, Susanna Lundstrom, and Jan-Egbert Sturm.** 2006. Market-Oriented Institutions and Policies and Economic Growth: A Critical Survey. *Journal of Economic Surveys* 20(2): 157–191.
- Diagla, Isaaka, and Thomas Vallee.** 2021. The Index of Economic Freedom: Methodological Matters. *Studies in Economics and Finance* 38(3): 529–561.
- Fike, Rosemarie.** 2017. Adjusting for Gender Disparity in Economic Freedom and Why It Matters. In *Economic Freedom of the World*, ed. James Gwartney, Robert Lawson, and Joshua Hall. Vancouver: Fraser Institute.
- Flake, Jessica K., Jolynn Pek, and Eric Hehman.** 2017. Construct Validation in Social and Personality Research: Current Practice and Recommendations. *Social Psychology and Personality Science* 8(4): 370–378.
- Friedman, Milton.** 1962. *Capitalism and Freedom*. Chicago: University of Chicago Press.
- Gwartney, James.** 2009. Institutions, Economic Freedom, and Cross-Country Differences in Performance. *Southern Economic Journal* 75: 937–956.
- Gwartney, James, Randall Holcombe, and Robert Lawson.** 1998. The Size of Government and the Wealth of Nations. *Cato Journal* 18(2): 163–190.
- Gwartney, James, Robert Lawson, Joshua Hall, and Ryan H. Murphy.** 2021. *Economic Freedom of the World*. Vancouver: Fraser Institute.
- Hall, Joshua, and Robert Lawson.** 2014. Economic Freedom of the World: An Accounting of the Literature. *Contemporary Economic Policy* 32(1): 1–19.
- Hammond, Samuel.** 2018. The Free-Market Welfare State: Preserving Dynamism in a Volatile World. May. Niskanen Center (Washington, D.C.). [Link](#)
- Heckelman, Jac C., and Michael D. Stroup.** 2000. Which Economic Freedoms Contribute to Growth? *Kyklos* 53(4): 527–544.
- Heckelman, Jac C., and Michael D. Stroup.** 2005. A Comparison of Aggregation Methods for Measures of Economic Freedom. *European Journal of Political Economy* 21(4): 953–966.
- Huskinson, Toby, and Robert Lawson.** 2014. Clusters of Economic Freedom. *Applied Economics Letters* 21(15): 1070–1074.
- Justesen, Morgens K.** 2008. The Effect of Economic Freedom on Growth Revisited: New Evidence on Causality from a Panel of Countries 1970–1999. *European Journal of Political Economy* 24(3): 642–660.
- Lawson, Robert, Ryan H. Murphy, and Benjamin Powell.** 2020. The Determinants of Economic Freedom: A Survey. *Contemporary Economic Policy* 38(4): 622–642.
- Linzer, Drew A., and Jeffrey K. Staton.** 2015. A Global Measure of Judicial Independence,

- 1948–2012. *Journal of Law and Courts* 3(2): 223–256.
- Ott, Jan.** 2018. Measuring Economic Freedom: Better Without Size of Government. *Social Indicators Research* 135: 479–498.
- Peltzman, Sam.** 1980. The Growth of Government. *Journal of Law and Economics* 23(2): 209–287.
- Rode, Martin, and Sebastian Coll.** 2012. Economic Freedom and Growth: Which Policies Matter the Most? *Constitutional Political Economy* 23: 95–133.
- Rothbard, Murray.** 1973. *For a New Liberty: The Libertarian Manifesto*. New York: Macmillan.
- Ruger, William P., and Jason Sorens.** 2018. *Freedom in the 50 States*, 5th ed. Washington, D.C.: Cato Institute.
- Sturm, Jan-Egbert, Erik Leertouwer, and Jakob de Haan.** 2002. Which Economic Freedoms Contribute to Growth? A Comment. *Kyklos* 55(3): 403–416.
- Sumner, Scott.** 2010. The Great Danes: Cultural Values and Neoliberal Reforms. Working paper, June 24. [Link](#)
- White, Lawrence H.** 2012. *The Clash of Economic Ideas: The Great Policy Debates and Experiments of the Last Hundred Years*. New York: Cambridge University Press.

About the Author



Ryan Murphy is a research associate professor at the Bridwell Institute for Economic Freedom at Southern Methodist University. He received a Ph.D. in economics from Suffolk University in September 2013. His research interests are diverse and include institutional economics, public policy, and macroeconomics. His email address is rhmurphy@smu.edu.

[Jan Ott's reply to this article](#)
[Go to archive of Comments section](#)
[Go to March 2022 issue](#)



Discuss this article at Journaltalk:
<https://journaltalk.net/articles/6044>