Southern Economic Journal 2020
DOI: 10.1002/soej.12428

Does Balance Among Areas of Institutional Quality Matter for Economic Growth?

J. Brandon Bolen* and Russell S. Sobel†

Countries with better institutions, as measured by the *Economic Freedom of the World* index, have greater prosperity, growth, and well-being. However, the previous literature has ignored interrelationships among the subareas, instead implicitly assuming substitutability and independence. We hypothesize that when some areas are weak, better scores in other areas are not a simple substitute. We find that including a simple measure of the within-country standard deviation among coexisting area scores results in a meaningfully significant improvement in the empirical growth specifications. Based on our findings, this should become routine practice in future empirical work. We also explore interactions among areas and find ones with legal system and property rights most important to include. These findings suggest a need to reconsider the implicit assumption of independence among area scores, as well as a balanced approach to institutional reform in practice. Improving the worst areas is more beneficial to growth.

JEL Classification: H10, O10, P0

1. Introduction

The economic prosperity of a nation is a complex function of many different factors, ranging from the country's colonial or legal origins to the levels of ethnolinguistic fractionalization, resource endowments, influential leaders, and religious homogeneity.[1] Countries even influence the prosperity of other nations both directly and indirectly through geographic contagion effects, international trade, tourism, immigration, and foreign aid.[2] Despite the roles played by these many factors individually, the most robust finding in the empirical literature is that economic prosperity is a function of the quality of the country's institutions under which an economy's resources are put to productive use.[3]

In the empirical literature, economists have measured economic institutions using country-level indexes, with the most widely employed being *The Economic Freedom of the World* (EFW) index by Gwartney et al. (2018).[4] The index scores the overall institutional quality of countries on a

* School of Business, Mississippi College, P.O. Box 4014, Clinton, MS, 39762, USA; E-mail: bbolen@mc.edu; corresponding author.
† Baker School of Business, The Citadel, 171 Moultrie Street, Charleston, SC, 29409, USA.
Received October 2018; accepted February 2020.

[1] See, for examples, Glaeser and Shleifer (2002), La Porta, Lopez-de-Silanes, and Shleifer (2008), Acemoglu, Johnson, and Robinson (2001), and Robinson, Acemoglu, and Johnson (2005), Easterly (2001), Glaeser et al. (2004), Easterly and Levine (1997), Sachs and Warner (1999), Grier and Maynard (2016), and Collier (2010).

[2] See Leeson, Sobel, and Dean (2012), Sobel and Leeson (2007), Powell and Ryan (2006), Powell, Clark, and Nowrasteh (2017), and Heckelman and Knack (2009).

[3] The literature defines institutions as the formal and informal rules, and customs governing human interactions (see North 1990, 1991). The link between these institutions and economic growth has been recognized since Smith (1976 [1776]), and this literature has grown based on the contributions of many scholars including Hayek (1944 [1994], 1945), North (1990, 1991), Easterly (2001), Olsen (1982), and Baumol (1990).

[4] Hall and Lawson (2014) survey the over 400 articles that use this EFW institutional measure.

scale of 0–10 based on scores in five major subareas (government size, legal system and property rights, sound money, freedom to trade, and regulations). Most studies using the EFW index employ the overall score of the country, which is the simple average of the index's five component area scores. Based on the robustness of the overall finding that better institutions foster growth and prosperity, studies have now begun to examine the underlying data in more detail to discern which areas may be the most important when a country is attempting to reform.[5] Studies have also begun to examine the effect of volatility in institutional quality over time as measured by the variance or unpredictability of institutional indexes and their subcomponents.[6] While these address important and often overlooked aspects of institutional structure, each treats the area subcomponents as independent and linearly substitutable, which we find seriously problematic.

Our article explores an aspect that has been completely overlooked in the prior empirical literature—how these different areas are interdependent and whether dispersion (or "balance") among the areas matters.[7] The problem with the prior empirical literature is it forces the empirical specifications to treat countries with the same overall score identically because they have the same average score, even when they have wildly different distributions of area scores around that average. In a simplified nutshell, the prior literature makes an implicit empirical assumption that two area scores that are both equal to five produce the same growth outcome as one area being 0 and the other a 10—essentially saying it is okay to have no property rights as long as you have a good tax code, and that having imbalanced scores is roughly the same for growth as having balanced scores in both areas. We find this is clearly *not* the case empirically, and this omission is easy to control for in future empirical work with our specific adjustment of adding a variable measuring the within-country standard deviation (SD) across coexisting areas scores.

The anecdotal evidence is clear that these interactions are present and is typified by the work of Hernando de Soto (2000, 2006). His research shows that when private property rights are not well defined and enforced, the quality of the country's other economic policies do not matter very much.[8] That is, even with good tax policies or regulatory codes, without proper titles, deeds, and identification systems, capital simply cannot be collateralized and the economy cannot function well. Similarly, when the former Soviet-bloc countries were in transition, Friedman (2002) stressed the complexity of the interrelation among policy areas when widespread privatization was unable to produce growth without the proper rule of law in place.

Some prior literature has used an analogy of a "three-legged bar-stool" to discuss the importance of institutional interaction (see, for example, Boettke 2001 and Boettke et al. 2007). Boettke (2001) invokes this analogy to explain the intricate institutional balance necessary for economic development among economic institutions, political/legal institutions, and social/cultural institutions and argues that imbalances, particularly in the case of Russia, were a major issue curtailing

---

[5] See Carlsson and Lundström (2002), Sobel (2017), Bjørnskov and Foss (2008), and Berggren and Jordahl (2005) for examples.

[6] See Asteriou and Price (2001), Berggren, Bergh, and Bjørnskov (2012), Bolen and Williamson (2019), Campos, Karanasos, and Tan (2012) and Dawson (2015).

[7] Note that the interdependency among EFW areas is distinct from the volatility of the EFW index over time. In fact, the measures of interdependence that we introduce later in the article have only a small correlation with the volatility of the EFW index over time.

[8] Institutions are generally viewed as broader constraints on action, while policies are usually viewed as the decisions or choices that emerge from within those constraints. Policies and institutions are closer to synonymous when discussing the economic institutions measured in these indices, such as strong property rights, sound money, and free trade; and we sometimes use these terms synonymously in this article due to specifically referencing their implementation and measurement as an index of different areas of each country's economic policies and institutions.

progress. As Boettke (2001, p. 5) puts it, "[u]nless all three legs are equally strong, the bar-stool will not be able to stand when we sit on it." Rajan (2019) describes this institutional interaction in terms of "three pillars." Rajan argues, "What then is the source of today's problems? In one word, *imbalance*! When the three pillars of society are appropriately balanced, society has the best chance of providing for the well-being of its people," (p. xvii).

The idea that institutional subareas are intertwined, while it has escaped introspection in the prior empirical growth literature, is something that is directly pointed out by the authors of the EFW index itself. In the publication describing the index, the authors explicitly state "there is reason to question whether areas (and components) are independent or work together like the wheels, motor, transmission, driveshaft, and frame of a car."[9] Therefore, while even the authors of the index readily acknowledge the presence of these possible interrelations, thus far this question remains a clear gap unanswered in the current empirical growth literature.

Interdependence among the institutional dimensions measured by the EFW index has two primary implications for the way empirical studies typically use the EFW index. First, interdependence among institutional dimensions implies that studies using only the *overall* EFW index ignore the potentially confounding effects of the interrelations among EFW areas. Therefore, empirical models may be mis-specified, omitting important information, which may affect the remaining coefficient estimates. Second, interdependence implies that studies including the separate sub-area scores without interaction terms measure only the weighted average of the conditional effects of these areas and ignore the conditional effects. Therefore, the estimated effects of the index's components may also suffer from the misspecification of the model.

These implications inform the outline of this article. First, we test whether imbalance among institutional dimensions affects economic growth independently of the overall average score. We indeed find a negative relation between institutional imbalance and economic growth rates, and we find that this result is robust to the choice of sample period, panel and cross-sectional data, income levels, and controls for endogeneity. Second, we explore some of the specific interrelations among EFW areas that underlie the importance of institutional imbalance, using interactions among the EFW areas in growth models. We find that, for example, the growth effects of the size of government vary with the quality of legal system and property rights which may help explain some of the contradictory results in the prior literature for this area's impact on growth.

In the end, we find that the inclusion of a simple measure of the SD among the within-country area scores for each country results in a *significant* and meaningful improvement in the empirical estimation, while retaining an intuitive interpretation of the model coefficients. This easy to implement and important inclusion should become standard practice in future empirical work using the overall index and area scores. We also highlight the need for future research to explore, both theoretically and empirically, the specific interrelations among EFW areas more exhaustively.

This is the first in-depth exploration of how standard institutional growth models are affected by the inclusion of measures of interdependence and dispersion (or balance) among the area scores. The implications of our findings for economic policy are also clear—balanced policies and reforms create more growth, and reforms should focus on bringing up the weakest institutional areas first.

---

[9] See Gwartney et al. (2018), pp. 5–6.

2. Why and How Area Interdependence Matters

Having the best tax system in the world may not matter much if a country has no rule of law or property rights, just as having good labor and business regulations may not matter much if a country severely restricts free trade or has rampant hyperinflation and an unstable currency. Without placing too much emphasis on the specifics of the analogy, consider the many different forms a neo-classical production function may take to model output (*Y*) based on labor (*L*) and capital (*K*). A linear production function of the form $Y = \gamma + \alpha \cdot L + \beta \cdot K$ specifies all inputs as perfect substitutes (unit adjusted). Even if, say, labor (*L*) goes to zero, output (*Y*) can still be positive as long as capital (*K*) is present. The idea of perfect substitutability in this relationship does not mean the inputs substitute at a 1 to 1 ratio, just that some linear combination of one input can substitute for units of the other.

If this is how institutions are related in terms of their ability to "produce" growth, it is possible for a country to "make up for" having bad scores in one area by simply having better scores in another area. That is, a highly stable currency could make up for having a weak rule of law. This is the assumption being made implicitly in all prior empirical literature using the index. Empirical studies that use the overall index implicitly assume that countries with identical average levels of the EFW index will have similar outcomes, or in other words, that the underlying distribution of institutional quality among coexisting institutional dimensions is irrelevant for economic outcomes. Empirical studies that include the area scores separately, but in a linear regression, still assume independence among the linear combination of area scores. Our main hypothesis is that these implicit assumptions are incorrect.

In contrast, a Cobb–Douglas production function $Y = \gamma \cdot L^{\alpha} \cdot K^{\beta}$ specifies inputs as dependent upon each other. If any single input, say labor (*L*), goes to zero, output (*Y*) goes to zero regardless of the quantity of capital (*K*) input. Again, without placing too much emphasis on the specific analogy but instead on the overall concept, our hypothesis is that economic growth is a process dependent on many different institutional areas. If one area is very weak, it may completely inhibit the process of economic growth even if other areas are strong.

If our hypothesis is correct, prior empirical models of growth across countries have been misspecified by not including measures of the relationships among the scores in the different subareas. The prior literature assumes independence among coexisting institutional structures, which imposes a perfect substitute relationship among the areas. That is, the implicit assumption of the empirical models used in the prior growth literature is that a weaker score in property rights could simply be made up for with a stronger score in another area such as business regulation. Therefore, the impact of a change in one area is independent on the levels of the scores in other areas. So, in prior work, a one-unit change in the regulation score is assumed to have the same impact regardless of whether the property rights area score is zero or nine.[10]

[10] Interestingly, one reason area score differentials may be negatively correlated with economic outcomes, ceteris paribus, is that it is a reflection that some areas of policy are being designed to benefit narrow interest groups. Policy based on overall ideology would tend to be implemented in a broad-based pattern, leading to more or less government intervention and regulation across all areas of policy (see, for example, Holcombe 2013). As a narrow example, a country simply trying to generate revenue may levy high tax rates on all forms of economic activity or all goods. When there are vast differentials with some industries or goods receiving much more favorable treatment than others, it is a sign of political influence and unproductive entrepreneurship (see Baumol 1990, Sobel 2008, and Sobel, Clark, and Lee 2007). All policies adopted by a country's government are the result of a collective-action political process that is influenced by voters, interest groups, and government agents and is best understood using the models in the public choice literature (see, for examples, Ekelund and Tollison 2001, Holcombe 1998, Shughart 1997, and Calcagno and López 2012).

We believe there is good reason to think that the productivity of resources would and do suffer from these types of imbalance in institutions, particularly with regard to property rights and the rule of law based on the work of Hernando de Soto (2000, 2006) and Friedman (2002). With weak property rights, it may be hard for resources to be put to productive use regardless of the stability of a nation's currency, for example. In de Soto's work, he illustrates, for example, how the inability to collateralize capital due to weak property structures inhibits growth even in places where other policy areas are strong as it lowers the productivity of capital investment.

This article attempts to empirically examine whether these area imbalances impact economic prosperity in addition to the overall quality of institutions. One of the possible relations that is the most obvious to start with is the within-country dispersion of the area scores. Somewhat similar to the analogy of institutions as a "three-legged bar stool" or "three pillars", it seems reasonable to ask whether two countries with the same overall average score, but with different dispersions of the area scores, should expect similar outcomes. A high dispersion, or variance, of area scores around the mean score reflects a very unbalanced portfolio of economic policies and/or institutions. Is a balanced portfolio of area scores with less dispersion and less variance better? In the case of a bar stool, obviously three equal legs work better than having one leg an inch too short and another an inch too long—is the same true for country institutions?

As an example, consider the countries of Ghana and Haiti. In 2015, these two countries had an identical EFW overall score of 6.53, so the current empirical growth literature would predict similar outcomes.[11] However, Ghana's five area scores all fit within a narrow range of 1.53 units from a low of 5.44 in Area 2 (legal system and property rights) to 6.97 in Area 1 (size of government), while Haiti's five area scores have a wide range of 5.30 units from a low of 2.48 in Area 2 (legal system and property rights) to 7.78 in Area 4 (freedom to trade). Haiti's legal system and property rights institutions are the third worst in the sample of all 159 countries scored in 2015—should we really expect similar economic growth outcomes in these two countries as the current empirical literature econometrically imposes?

Our assertion is that this imbalance in area scores would indeed matter—it would be harmful to prosperity and implies that Haiti should have worse economic outcomes than Ghana despite having the same overall institutional score. The raw data would seem to support our hypothesis. According to World Bank data, real GDP per capita (in constant 2010 U.S. dollars) in Ghana grew from $969 in 2000 to $1686 in 2015, while in Haiti it declined over the same period from $767 to $728. Alternatively stated, despite an *identical* overall EFW score, Ghana achieved 3.6% annual average real GDP per capita growth, while Haiti suffered −0.3% annual average real GDP per capita decline. This type of within-country dispersion of area scores, and its impact on economic performance, has been *completely overlooked* in the prior literature that uses the overall EFW index to explain growth.

Let us continue by examining more country pairs for consideration using data on their EFW area scores and economic growth. We also use this opportunity to introduce one of the empirical measures of that we employ in our later regression analysis. Table 1 shows data for five pairs of countries, including the example just discussed of Ghana and Haiti. Each pair is chosen so that the two countries have a virtually identical overall EFW score in 2015 but very different levels of area dispersion, or balance. For each pair, the country with a more balanced portfolio of area scores (less dispersion among areas) is listed first in italics, while the country with less balance (more dispersion among areas) is listed second in normal font.

---

[11] EFW data in these examples are from Gwartney, Lawson, and Hall (2017).

Table 1. Examples of Country Pairs with Different EFW Area Imbalance in 2015

Country	EFW Index Overall 2015	Area 1 Size of Government	Area 2 Legal System and Property Rights	Area 3 Sound Money	Area 4 Freedom to Trade Internationally	Area 5 Regulation	SD of Area Scores 2015
Barbados	*6.47*	*6.30*	*5.76*	*6.72*	*7.09*	*6.49*	*0.49*
Madagascar	6.44	8.80	2.88	8.05	6.61	5.86	*2.28*
Ghana	*6.53*	*6.97*	*5.44*	*6.87*	*6.51*	*6.85*	*0.61*
Haiti	6.53	7.63	2.48	7.76	7.78	7.01	*2.27*
Bhutan	*7.11*	*7.75*	*6.86*	*7.03*	*6.27*	*7.62*	*0.53*
El Salvador	7.13	8.41	3.70	9.58	7.66	6.28	*2.21*
Qatar	*7.43*	*6.54*	*6.30*	*8.37*	*7.56*	*8.40*	*0.84*
Belgium	7.44	3.93	7.25	9.66	8.26	8.09	*2.11*
Chile	*7.77*	*7.91*	*6.45*	*9.31*	*8.21*	*6.97*	*1.02*
Denmark	7.77	4.07	8.25	9.61	8.42	8.49	*2.10*
Average for "balanced" countries with low SD of area scores (countries in italic)							*0.70*
Average for "imbalanced" countries with high SD of area scores (countries in regular font)							2.19
Difference (balanced – imbalanced)							−1.50

Notes: EFW data are from Gwartney, Lawson, and Hall (2017). Bhutan has EFW observations only for 2013–2015, and Qatar has no Real Per Capita GDP growth observations prior to 2000 and EFW observations only for 2010–2015.

The final column of Table 1 shows one of the measures that we employ extensively in our more formal empirical analysis: the within-country SD of the coexisting area scores. The bottom of the table shows the averages for each group of countries (the balanced and imbalanced) and the differential between the group averages. For the five balanced countries (those with low dispersion), the average SD in each country's area scores is 0.70, while it is 2.19 for the imbalanced countries, a difference of −1.50.

Like the example discussed above of Ghana and Haiti, these other pairs differ in that one of the countries has at least one area out of line with the others. Some are striking. For example, while El Salvador has an extremely low score on legal system and property rights (3.70), it has a very high score for sound money (9.58). Similarly, Belgium and Denmark score poorly on size of government (3.93 and 4.07, respectively) while scoring very high on sound money (9.66 and 9.61, respectively).

Is it true that Madagascar and Haiti, who both have legal system and property rights scores below three points, can compensate by improving their scores in another area? And, does El Salvador's adoption of the U.S. dollar and Belgium's use of the Euro, resulting in high scores in sound money, compensate for poor quality institutions in other areas? Using an equally weighted average to aggregate the area scores into one overall EFW score assumes the answer to these questions is yes. Based on this assumption, Madagascar and Barbados have identical quality institutions despite Madagascar's poor legal system and property rights because it has some combination of lower rates of inflation, less government spending, and/or lower marginal tax rates. The same logic is also implied for the institutional comparisons between El Salvador and Bhutan, Belgium and Qatar, and Denmark and Chile. The comparisons presented in Table 1 clearly show the implicit assumption of the index—an assumption that, in the very least, warrants further empirical testing.

Now let us consider how differentials such as these have affected the economic performance of these countries. To align with the period of time considered in our later empirical models, Table 2 shows the average from 2000 to 2015 of each country's overall EFW score, the average area

Table 2. Economic Performance Comparison of Country Pairs, 2000–2015

Country	Avg. EFW Index Overall (2000–2015)	Avg. SD of Area Scores (2000–2015)	Avg. Real Per Capita GDP Growth (2000–2015)
Barbados	*6.43*	*0.85*	*0.69%*
Madagascar	6.08	2.03	−0.09%
Ghana	*6.53*	*0.80*	*3.64%*
Haiti	6.46	2.43	−0.34%
Bhutan	*7.01*	*0.56*	*5.31%*
El Salvador	7.36	2.03	1.56%
Qatar	*7.44*	*2.03*	*1.47%*
Belgium	7.42	2.06	0.95%
Chile	*7.75*	*1.12*	*2.75%*
Denmark	7.83	2.21	0.68%
Average balanced countries	*5.86*	*1.07*	*2.31%*
Average imbalanced countries	5.86	2.15	0.46%
Difference (balanced – imbalanced)	0.00	−1.08	1.85%

Notes: EFW data are from Gwartney, Lawson, and Hall (2017). Bhutan has EFW observations only for 2013–2015, and Qatar has no Real Per Capita GDP growth observations prior to 2000 and EFW observations only for 2010–2015.

SD, and the average growth rate of real GDP per capita (US$). The differentials are striking and confirm the prediction of our hypothesis—imbalance matters, significantly. Despite having virtually identical overall levels of EFW index, in each and every pair the country with more balance (less dispersion) grew faster. Across all pairs the differential is large, a 1.85 percentage point per year lower growth rate for countries with more imbalance in their area scores. The five with less imbalance experienced real per capita GDP growth of 2.31%. Alternatively, those with more imbalance grew at a meager 0.46%, and two of them experienced negative growth rates.

While these are simple averages of pairs selected based on the similarity of their overall scores, the difference in the performance of the countries is striking. By itself, we believe the differential presented above warrants a deeper analysis and calls into question the precision of the empirical results from the dozens, if not hundreds, of empirical studies that have used the overall EFW score in cross-country growth regression without controlling for this imbalance among the area scores. The prior empirical literature has completely overlooked this important issue of dispersion of area scores around the average value causing an impact on growth. Quite simply, failure to take this into account may bias the results as it forces the data to explain both the good and bad performances with the same overall score. It is worth explicitly mentioning again that the studies that also include the individual area scores in a regression do not solve this problem as they all enter linearly—with the analogy earlier to a linear production function applying in which it forces them to be perfect linear substitutes. Our comparisons suggest you simply cannot add to one area to make up for a low score in another area.

We now turn to a more sophisticated empirical analysis to see if controlling for other factors in a regression framework matters, and to see how the standard empirical model previously estimated in the literature is affected by including measures of imbalance as well as other measures of interactions among the area scores. In the end, we conclude that the basic differences found here in the raw data comparison show up in magnitude almost precisely in the more sophisticated empirical techniques and are not only robust, but are easy to correct for empirically by simply including the within-country SD of the area scores.

3. Measures of Imbalance

There are many alternative ways to measure the relationships and interactions among the institutional areas. We begin with simple measures of imbalance before highlighting a few of the specific interdependencies that explain the relation between growth and imbalance. Our simple measures are:

1. The EFW Area SD—which is calculated as the within-country SD of the five area scores in a single period.
2. The EFW Area Range (Max ‑ Min)—which is calculated as the difference between the highest and lowest area scores within each country in a single period.

Let us briefly discuss these measures and what they may reflect, as well as our expectations. The *EFW Area SD* and *EFW Area Range* are both measures of the variance or dispersion of the area scores around the mean value. The mean value is what is used as each country's overall EFW score in the index. Because the range measure is based only on the highest and lowest area values, it uses information from only two of the five area scores. Holding constant the overall mean EFW score, we hypothesize that increasing the dispersion would hurt economic performance, and this is what we examined in the raw data presented earlier. We do not have a prior of which measure may be better, but the SD measure includes more information. These two measures get at the idea of institutional balance and the "three-legged bar stool" analogy.

The *EFW Area SD* was the measure we employed in our illustrative tables of country pairs earlier. Let us consider the following example of how it is calculated. The EFW score for the United States in 2016 is 8.03 according to the EFW 2018 annual report. This is the equally weighted average of five area components scores of 6.43, 7.40, 9.85, 7.65, and 8.83. The SD of the five areas is 1.33 points. The straightforward interpretation of this measure is based on the SD essentially measuring the amount by which an average one of the subareas differs from the mean. The second measure, the *EFW Area Range* would be 3.42 based on the differential between the highest area score (9.85) and the lowest (6.43). In our subsequent models that are based on panel data, the implementation is straightforward for these variables. Whenever we estimate cross-sectional models using long-term growth data, however, it is worth mentioning that we use the average of the annual values for these dispersion measures over the period.

These measures get at the main underlying problem that plagues the prior empirical literature, namely that empirical studies using the overall index to quantify institutions do not control for the dispersion of the area scores around the overall mean value. However, a second approach to this issue which may have merit (even independently) is examining interactions among EFW areas to explore the interdependencies underlying the growth effects of each area.

The interaction terms get at a slightly more complicated relationship, which is whether the growth effects of one area depends on the level of another area. That is, for example, does the growth effect of improving the area score for size of government from a 3 to a 4 differ if the country's score in some other area, say regulation, is a 5 or a 9. A positive value for a pairwise interaction would suggest the effects of a change in one area grow with the value in the other area. In contrast, a negative interaction would suggest the growth effects of a change in one area lessen with higher values in the other area.

Picking pairs from the five areas yields 10 possible pairwise interactions. However, there is no reason to expect the interactions to be simply pairwise. A production function with five inputs, for example, would produce marginal effects that are multiplicative in all inputs. In this analysis, we focus on a few of the interactions that we believe are most important to consider based on the previous findings in the empirical literature. In earlier, more lengthy drafts of this article we showed that none of these interaction approaches, even including all 26 possible combinations of two, three, four, or even five areas interacted can come close to the improvement from simply including our SD measure.

The interaction terms have two drawbacks relative to the measures of dispersion that cause them to underperform and not really address the main issue we uncover. First, the interaction terms allow the impact of a one-unit change in, say, area 1 to depend on the other area scores, but it still forces the effect to be the same for all values of area 1. For example, say area 1 goes from 3 to 4. The interactions allow the effect to differ depending on, say, whether area 2 is 4 or 8. But the balance issue implies that the effect of changing area 1 from a 3 to a 4 may be different than area 1 going from 7 to 8, and this, then, also depends on the other area scores. In other words, the simple interaction forces the partial derivative of growth with respect to a given area to depend on the values of other areas, but not on the value of the original area itself and how it compares to the other areas.

In addition, with interaction terms in a model, the partial derivatives are functions of multiple coefficients and variables requiring additional steps for interpretation. With a measure such as SD, the interpretation is much simpler. For example, a model including both each country's overall EFW score, and the SD of the area scores can be interpreted as follows. The coefficient on the overall EFW score is interpreted as an increase in economic freedom holding constant the dispersion of area values around the mean. In other words, it is a balanced increase in freedom among the subareas. The coefficient on the SD measure is interpreted as an increase in the imbalance holding constant the mean value of the area scores, so simply an increase in spread around a given mean value.

In the empirical analysis that follows, we revisit a standard model from the prior literature and add to it both of the measures of imbalance described above. We assess the growth effects of imbalance and the impact that the addition of these measures has on the goodness of fit of the models, using the adjusted *R*-squared. After presenting these benchmark results, we then see if the results are robust to different samples of data, and to different empirical methods such as system GMM estimation, panel data analysis, and more discrete measures of single area institutional outliers. Lastly, we explore some of the interactions among EFW areas that help explain why the inclusion of an imbalance measure is an improvement to the literature's growth models.

4. Data, Model, and Benchmark Results

Because our interest is in exploring how and whether these measures should be included in the existing literature, we start by adopting the most heavily cited institutional growth model first estimated by Gwartney, Holcombe, and Lawson (2006), hereinafter referred to as the Gwartney, Holcombe, and Lawson (GHL) model.[12] Based on subsequent improvements to that model in the literature, we add additional controls for malaria ecology and regional fixed effects to the authors'

[12] We acknowledge that there was some debate surrounding the interpretation of the coefficients on $EFW_{i,\text{ initial}}$ and ΔEFW_i in this model by Cole and Lawson (2007) and De Haan and Sturm (2006). The interpretation of these coefficients is irrelevant for the discussion herein, and the results of this article are robust to excluding either/both of these variables from the model. Also, in their original article, Gwartney, Holcombe, and Lawson (2006) also try to account for the endogeneity of private investment, which we do not do here as it is not important for our current purposes, but would result in higher coefficient estimates on the EFW institutional variables.

original cross-sectional model, and we include the more recent data. We estimate the model without our measures. Then, we add measures of imbalance one at a time to compare them to one another and to the baseline model. Using Z_i to represent our measure(s) of imbalance for country i, we estimate:

$$\overline{\Delta \text{GDPPC}}_i = \beta_1 + \beta_2 \text{EFW}_{i,\text{initial}} + \beta_3 \Delta\text{EFW}_i + \beta_4 \overline{Z}_i + \beta_5 \ln(\text{GDPPC})_{i,\text{initial}} + \beta_6 \overline{\text{Priv}}_i + \beta_7 \overline{\text{Pub}}_i + \beta_8 \Delta\text{HC}_i + \beta_9 \text{Trop}_i + \beta_{10} \text{Coast}_i + \beta_{11} \text{Malaria}_i + \gamma_i + e_i$$

where $EFW_{i,\text{ initial}}$ and ΔEFW_i measure the initial quality of institutions and the change in institutions respectively from Gwartney et al. (2018). $\overline{\text{Priv}}_i$ and $\overline{\text{Pub}}_i$ are the average private and public investment rates as percentages of GDP. Additionally, we control for initial GDP per capita, the change in human capital, the percent of the population in a tropical climate, the percent of the population within 100 km of a coast, malaria ecology, and region fixed effects (γ_i).[13] Following GHL, the dependent variable in our models is real GDP per capita growth, collected from The World Development Indicators (WDI, 2018). We will slightly vary the overall specification as we check our results for robustness and will describe these changes as we proceed. Summary statistics for the data used in this article are provided in Table A1 and Table A2 in Appendix A.[14]

Our benchmark results are presented in Table 3. Column 1 basically shows the foundational model from the previous literature following GHL. Columns 2 and 3 show the two measures of institutional imbalance: the SD and the range of the EFW's five co-existing areas. Starting with the results in column 2, when we include the SD of the area scores in the regression, the variable has a coefficient estimate that is negative and statistically significant. The coefficient estimate of −1.380 is also economically meaningful. A country whose area scores, on average, are one unit away from the mean value (thus a SD of one unit) will experience an average annual growth rate of real GDP 1.38 percentage points lower than a country whose area scores are identical and all equal to the mean value.

As an illustrative example this means a country with area scores of 6, 4, 6, 5, and 4 would have significantly lower growth (−1.38 percentage points) than a country with area scores of 5, 5, 5, 5, and 5 even though both have an identical overall EFW index mean score of 5.0. More succinctly, imbalance in the area scores matters, and it hurts economic performance. Interestingly, this 1.38 coefficient estimate is almost identical to what the raw data in Tables 1 and 2 showed. In those tables, the average difference in the SD among the groups was 1.5 in 2015, and it produced a growth differential of 1.85, the ratio of which is 1.23, quite similar to the 1.38 coefficient estimate. This confirms our first main hypothesis that the within-country dispersion of EFW area scores matters significantly, and independently, in addition to any effect from the average EFW level in explaining the cross-country growth data.

This does not mean, however, that a country with imbalance among the areas can gain by simply cutting the area most above the average. The reason is that this coefficient must be interpreted as an increase or decrease in the dispersion (SD) *holding constant* the overall average of the area

[13] Private and public sector gross fixed capital formation from WDI, which is calculated as a percentage of GDP, serve as measures of private and public investment. Where WDI investment data are unavailable, OECD measures of private and public investment as a percentage of GDP substitute for WDI data. Penn World Table human capital data, which combines measures of average schooling and returns to schooling, serves as the human capital measure in this article. Geography controls include the percentage of a nation's population located in a tropical climate and the portion of the population located within 100 km from an ocean coastline from Gallup, Sachs, and Mellinger (1999) and malaria ecology from Acemoglu, Johnson, and Robinson (2002).

[14] In the Appendix, we provide summary statistics separately for the cross-section and panel models.

values (the overall EFW index score). For example, consider a country with area scores of 10, 6, 6, 6, and 2. If it were to cut the 10 score to a 6, it would indeed lower the dispersion, which would have a positive effect on growth. But the offsetting factor is that doing so would also lower the overall average, which based on the ΔEFW variable coefficient's significant positive value of 0.508, would more than offset the positive effect. On the other hand, if that same country were to raise the 2 to a 6, there would be two positive effects, one from lowering dispersion and the other from increasing the overall average EFW score. Using this example if a country were to have made those changes in 2012 for example, average real GDP growth would have declined by 0.048 in the one cutting the 10 to a 6, while it would increase by 0.765 in the one raising the 2 to a 6. The other implication is that a one unit increase in the lowest area's score will improve growth more than a one unit increase in the highest area's score for a country. While both result in the same increase to the overall average, one reduces dispersion while the other increases it. Both have overall positive growth effects, but the effect is greater for improving the lowest area. The larger implication for policy is that balanced increases in economic freedom cause more growth, and improving the weakest areas should be a priority.

Not only is the coefficient on the SD measure individually and economically significant, there is a marked improvement in the explanatory power of the models. Adding this one variable causes the adjusted R-squared to rise from 0.490 to 0.561, a meaningful and large 15% increase in the explanatory power of the model even after inflicting a punishment for adding another variable. It is worth noting that not only does the coefficient on the SD measure have an easy interpretation, but its impact on the main variables of interest in the prior literature (the ΔEFW variable) is also easy and intuitive. The proper interpretation of its coefficient is now an increase in the average level of freedom, holding constant the dispersion of the area scores—a rising tide that lifts all boats.

Column 3 presents an alternative measure to control for this dispersion, the range of the area scores. It too is statistically and economically significant and also shows that dispersion hurts growth independent of overall levels of economic freedom. While the interpretation using this measure is similarly intuitive, the impact it has on the explanatory power of the model is much smaller—which is reasonable based on it containing less information than the SD measure. Based on the goodness-of-fit criterion, we think it is obvious to prefer the SD measure to the range measure, although either is an improvement on the model in the existing empirical literature.

Before drawing final conclusions, it is worthwhile to see whether this pattern of results is robust. We begin by simply taking the model from its standard cross-sectional form to its panel form, and these results are shown in Table 4. For the panel data models, we include both country and year fixed effects (in place of the regional fixed effects in the cross-sectional model). We also extend the sample period to include data from 1980 to 2015 and use five-year observations. This five-year specification is more common for panel models in the prior empirical literature so that the coefficients in the model explain longer term trends in growth rather than annual fluctuations.[15] Due to data limitations, we include total investment as a percent of GDP rather than separating it into private and public investment.

The results in Table 4 basically confirm our findings in Table 3. The SD and range measures both show that dispersion significantly hurts growth, and they perform similarly in terms of explanatory power. Again, the explanatory power of the empirical growth specification is vastly improved (by roughly 60%) by including these measures. Therefore, our main findings are clearly robust to either estimation, and clearly show the value of including a dispersion measure in future empirical growth models.

[15] The negative growth effects of imbalance found in the cross-sectional models presented in Table 3 are also robust to the 1980–2015 sample period, but we lose 17 observations due to a lack of available data (see Appendix B).

Table 3. Cross-sectional Economic Growth Regressions: Growth Effects of Institutional Imbalance (2000–2015)

Dependent Variable: Per Capita Real GDP Growth	(1)	(2)	(3)
Independent variables			
$EFW_{initial}$	0.098	0.217	0.785**
	(0.246)	(0.244)	(0.383)
ΔEFW	0.539*	0.508*	0.833**
	(0.291)	(0.264)	(0.325)
EFW area SD		−1.380***	
		(0.347)	
EFW area range (max - min)			−0.709**
			(0.325)
$\ln GDP_{initial}$	−0.835***	−0.920***	−0.942***
	(0.188)	(0.192)	(0.206)
Tropical population	0.590	0.374	0.269
	(0.744)	(0.603)	(0.741)
Coastal population	−0.112	−0.002	−0.109
	(0.404)	(0.412)	(0.406)
Malaria ecology	−2.697**	−1.964*	−2.732**
	(1.301)	(1.142)	(1.314)
Private investment	0.121***	0.103***	0.111***
	(0.036)	(0.036)	(0.037)
Public investment	0.111***	0.093***	0.101***
	(0.038)	(0.031)	(0.038)
Human capital	1.185	1.068	1.113
	(0.948)	(0.951)	(0.977)
Constant	7.380***	9.377***	9.985***
	(2.476)	(2.679)	(2.961)
Region fixed effects	Yes	Yes	Yes
Observations	97	97	97
R-squared	0.570	0.634	0.589
Adjusted R-squared	0.490	0.561	0.507

Notes: The SD and range are computed as the overall averages of annual values. Robust standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provides GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).
***$p < 0.01$,
**$p < 0.05$,
*$p < 0.1$.

Some studies, rather than using the overall average value of the EFW index, include the five area scores independently. We made the case earlier that this separation still imposes independence restrictions (linear substitutability among the areas) that we think are unfounded. To see, we re-estimate our baseline models from Table 3 but this time separate the EFW score (both the initial value and the changes) into the five area scores. The results of doing so are provided in Table 5.

To conserve space, as there would be 10 additional variables, we put "YES" in place of the coefficients and show the asterisks for the test of joint significance of the variables. What is of interest here for our purposes is to examine the significance of our new measures of the imbalance among the areas and the increase in explanatory power from including them.

Confirming our earlier statements regarding the prior literature, it is clear from the results in Table 5 that the inclusion of individual area scores does not fix the underlying problem we are considering here in this article. The explanatory power of these models again jumps with the inclusion

Table 4. Panel Economic Growth Regressions: Growth Effects of Institutional Imbalance (1980–2015)

Dependent Variable: Per Capita Real GDP Growth	(1)	(2)	(3)
Independent variables			
$EFW_{initial}$	0.320**	1.048***	1.527***
	(0.157)	(0.224)	(0.315)
ΔEFW	0.557**	0.938***	1.393***
	(0.262)	(0.207)	(0.269)
EFW area SD		−0.552**	
		(0.253)	
EFW area range (max ‐ min)			−0.475***
			(0.165)
Lagged per capita real GDP	−5.039***	−5.544***	−5.449***
	(0.708)	(0.778)	(0.786)
Total investment	−0.042**	0.108**	0.107**
	(0.019)	(0.045)	(0.045)
Human capital	1.751*	1.817	1.809
	(0.915)	(1.405)	(1.348)
Constant	38.179***	37.172***	36.416***
	(5.521)	(6.604)	(6.609)
Country fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	681	681	681
Number of countries	125	125	125
R-squared	0.264	0.407	0.410
Adjusted *R*-squared	0.252	0.396	0.400

Notes: Robust standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provide GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).
***$p < 0.01$,
**$p < 0.05$,
*$p < 0.1$.

of our new measures. The SD measure is again both statistically and economically significant, and provides more than a 12% boost in the explanatory power of the model. The actual coefficient estimate on the dispersion measure is −1.696 which is also roughly similar to the earlier results from the cross-sectional model as well as the raw data analysis.

Based on the results presented above, we conclude that indeed the empirical literature is missing something important by not including an adjustment for the imbalance or dispersion among the area scores. Simply computing the SD of the area scores as a separate variable for each country significantly improves the fit of these models. Perhaps more importantly for policy, however, is our finding that institutional imbalance hurts growth. A country simply cannot make up for having a weak score in one area by having better scores in other areas. Mean preserving increases in dispersion hurt growth, and mean preserving decreases in dispersion help growth. Again, to be clear, simply cutting the high areas does not increase growth, though, because it has the offsetting negative impact of lowering the mean which more than offsets it. Improving the worst areas, however, has a twofold beneficial impact of increasing the mean as well as lowering the dispersion. Reforms should focus on bringing up the weakest areas to have the most impact on growth.

Table 5. Cross-sectional Economic Growth Regressions with Individual EFW Areas: Growth Effects of Institutional Imbalance (2000–2015)

Dependent Variable: Per Capita Real GDP Growth	(1)	(2)	(3)
Independent variables:			
EFW $Areas_{initial}$ (for all five areas separately)	Yes***	Yes***	Yes***
ΔEFW areas (for all five areas separately)	Yes***	Yes***	Yes***
EFW area SD		−1.696***	
		(0.480)	
EFW area range (max - min)			−0.477
			(0.442)
$lnGDP_{initial}$	−1.087***	−1.064***	−1.124***
	(0.189)	(0.183)	(0.200)
Tropical population	0.697	0.576	0.523
	(0.741)	(0.555)	(0.739)
Coastal population	−0.287	−0.067	−0.236
	(0.433)	(0.451)	(0.429)
Malaria ecology	−2.173	−1.736*	−2.279*
	(1.342)	(1.040)	(1.341)
Private investment	0.130***	0.103***	0.123***
	(0.033)	(0.033)	(0.035)
Public investment	0.128***	0.106***	0.122***
	(0.036)	(0.034)	(0.037)
Human capital	1.432	1.03	1.331
	(0.986)	(1.004)	(1.007)
Constant	9.700***	12.555***	11.418***
	(2.488)	(2.704)	(3.359)
Region fixed effects	Yes	Yes	Yes
Observations	96	96	96
R-squared	0.656	0.711	0.661
Adjusted *R*-squared	0.547	0.613	0.547

Notes: The SD and range are computed as the overall averages of annual values. Robust standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provides GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).
***$p < 0.01$,
**$p < 0.05$,
*$p < 0.1$.

5. Robustness Tests of the Growth Effects of Institutional Imbalance

In this section, we test the robustness of our results for the possibility of omitted variable bias and endogeneity. Because we have made our strong case for using the SD measure, it is the method that we focus on it here in the robustness section. We begin with the Oster test using the stability of coefficients and movement in the *R*-squared values to determine if a result is robust to omitted variable bias (Oster 2019).

Table 6 presents the results of the Oster test, which specifically tests the robustness of the EFW area SD presented in column two of Table 3. We calculate the degree of selection of unobserved variables relative to the observed variables, δ, where $\delta = 1$ if unobserved variables and observed variables are equal in their ability to explain average growth rates and $\delta > 1$ if unobserved variables explain more of the variation in average growth rates than the observed variables. Altonji et al. (2011) and Oster (2019) recommend using $\delta = 1$ as the critical value above which the result is

Table 6. Selection on Unobservables (Oster) Test for Growth Effects of EFW Area SD

	Coefficient, Uncontrolled Model	Coefficient, Controlled Model	δ for $\beta = 0$; $\max(R^2) = 1$	Oster bounds of β; $\max(R^2) = 1$
Institutional imbalance	−1.54	−1.38	2.86	(−1.38 to −1.20)
R-squared	0.11	0.63		

Notes: Selection on unobservables is conducted using the "psacalc" command in Stata by Oster (2019). δ is the proportional selection of unobservables, where $\delta = 1$ means that unobservables are equally as important as observables in the model. The bounds of β are calculated using $\delta = 0$ and $\delta = 1$. Test conducted on the OLS model presented in column 2 of Table 4, where the average GDP per capita growth rate from 2000 to 2015 is the dependent variable and controls include average private and public investment rates, initial institutional quality, the change in institutional quality, changes to human capital, geography and initial GDP.

robust to omitted variable bias. We find $\delta = 2.86$, which suggests that the negative growth effects of institutional imbalance or dispersion, as measured by the mean annual SD of EFW areas, is robust to omitted variable bias.[16]

The fourth column of Table 6 presents the bounds of the estimated coefficient by forcing $\delta = 0$ and $\delta = 1$ separately. Again, the results suggest that the negative growth effects of the SD of EFW areas is robust to omitted variable bias because the bounds do not include zero.

Turning to other possible sources of endogeneity, Table 7 presents the re-estimation of the columns 1 and 2 in our panel model (Table 4) using system generalized method of moments (GMM) estimation, as described by Arrellano and Bond (1991), Blundell and Bond (2000), and Roodman (2015). System GMM is appropriate for small T, large N panels and instruments endogenous variables with their own lags. We use two-step system GMM with Windmeijer-corrected standard errors and report the instrument count and the results of the Arrellano-Bond test for AR (2) autocorrelation and the Hansen test for over identification for each model. The model with the variable of interest satisfies the rules of thumb for determining the appropriate fit of the model (Roodman 2015). The results support the original findings that the negative and significant growth effects of the SD of EFW area scores are robust to instrumenting endogenous variables. If anything, the magnitude of the coefficient grows in size in system GMM estimation. Thus, our main conclusions and results are robust to either of these issues.

In the Appendix (Appendix B, Table B1) we also show the results from robustness tests of our baseline model across subsamples of the data based on income levels (e.g., either excluding low- or high-income countries) and across sample periods. The conclusion is that in all subsamples, the SD measure is economically and statistically significant, which is consistent with our main findings.

As a final robustness check, we consider the possibility that what we are finding is simply a relic of a country having one bad area score. To do so, we create a dummy variable equal to one if the mean of any EFW area falls below a threshold, and we consider the 10th, 15th, 20th, and 25th percentile of each individual EFW area separately. We include this variable along with our SD measure and the control variables used throughout the article. We provide the results in the appendix (Appendix B, Table B2) showing the models where we add a dummy variable equal to one if the mean of any of the EFW components falls below the threshold. Regardless of the threshold used, the growth effects of the SD measure remain negative and statistically significant, with the

[16] In these tests, we employ the strictest assumption that the maximum R-squared value equals one.

Table 7. System GMM Estimation of EFW Area SD (Five-year Panel Data 1980–2015)

Dependent Variable: Per Capita Real GDP Growth	(1)	(2)
Independent variables		
$EFW_{initial}$	1.570	2.817***
	(1.153)	(0.922)
ΔEFW	2.218	3.907***
	(1.421)	(1.117)
EFW area SD		−4.788***
		(1.648)
Lagged per capita real GDP growth	0.056	0.147
	(0.269)	(0.202)
Lagged per capita real GDP	−0.986	−0.138
	(1.789)	(1.168)
Total investment	0.099	0.000
	(0.107)	(0.109)
Human capital	3.567	2.050
	(2.749)	(2.033)
% tropical population	0.877	2.088
	(2.045)	(1.541)
% coastal population	−1.319	−2.264*
	(1.676)	(1.256)
Malaria ecology	1.241	1.125
	(2.622)	(2.249)
Constant	0.000	0.000
	(0.000)	(0.000)
Year fixed effects	Yes	Yes
Observations	558	558
Number of countries	115	115
Number of instruments	32	37
p value A–B test (AR2)	0.593	0.182
p- value Hansen (overid.)	0.006	0.112

Notes: Windmeijer-corrected standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provides GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).
***$p < 0.01$,
**$p < 0.05$,
*$p < 0.1$.

magnitude roughly unchanged, even after controlling for a country having one low area score. Meanwhile, the growth effect of at least one EFW area falling below the threshold is statistically indistinguishable from zero in each model. These results clearly show that the imbalance measure across the areas matters independently of simply having one low area score—and it matters more, statistically and economically speaking.

6. A Limited Examination of the Underlying Interdependencies among EFW Areas

Up until now, we have only discussed general measures of institutional imbalance and provided evidence of a negative relation between institutional imbalance and economic growth. We now turn to examining the interdependencies among institutional areas that underlie the growth effects of institutional imbalance. While we believe the inclusion of the SD measure of imbalance, by itself, is a necessary and significant improvement to the specifications of the prior empirical literature, the inclusion of interaction terms allows for a more detailed look at the specific interdependencies among EFW areas.

Carlsson and Lundström (2002) is the earliest and most heavily cited attempt to measure the growth effects of the EFW's components. The authors find that the index's components are positively related to growth except size of government and freedom to trade, which are negatively related to growth rates. This finding has persisted as the empirical literature expanded, particularly with regard to the size of government, provoking scholars to offer explanations for the seemingly counterintuitive finding. Some, like Dawson (2003), argue that causality runs the other direction. That is, prosperous countries choose larger governments. Bergh (2020) argues that larger government is associated with higher growth rates because size of government indicators only capture fiscal policies and not the extent to which economic choices are centralized.

Regardless of which explanation(s) is (are) correct, the literature has yet to acknowledge that the growth effect in question is the weighted average of the conditional growth effects of the size of government because of the functional form taken by the statistical models. As we have already demonstrated in the previous section, the growth effects of the EFW's areas depend on the levels of the other EFW areas. Therefore, the growth effects of the size of government likely depends on the level of the scores in the other four areas.

The first panel of Table 8 presents the marginal growth effects of a one unit increase in the size of government conditional on the quality of the legal system and property rights. The second panel presents the marginal growth effects of a one unit increase in the legal system and property rights conditional on the quality of the size of government measure. Figure 1 depicts these relations graphically. The model is identical to the cross-sectional model presented in Table 8, but instead of including the initial value of and the change in the overall EFW index, we include the initial values of, the changes in, and the interactions between the measures of size of government and the legal system and property rights.[17]

As depicted in the first panel of Table 8 and the left-hand graph in Figure 1, the marginal growth effect of increasing the size of government by one point varies with the quality of the legal system and property rights. At low legal system and property rights scores, the marginal effect of the size of government is negative and statistically significant at the 10% level. However, as the score for legal system and property rights increases, the marginal effect of size of government becomes statistically insignificant, and at very high score for legal system and property rights, the marginal effect is positive and nearly statistically significant (*p* value of 0.108 where legal system and property rights equal 10).

Thus, the finding in the prior literature that improvements in the size of government score (i.e., smaller government sectors) harm growth is present only for countries with very low scores for their legal system and property rights. This is suggestive that the work of de Soto mentioned earlier may be the likely explanation for this counterintuitive result in the prior literature with regard to the size of government measure.

The second panel of Table 8 and the right-hand graph in Figure 1 show the corresponding conditional growth effects of increasing the score for legal system and property rights on the score

[17] These interactions are levels on levels. For example, we interact the initial level of size of government with the initial level of legal system and property rights.

Table 8. Conditional Marginal Effects of the Size of Government and Legal System and Property Rights from Cross-sectional Growth Model

Conditional Marginal Effects of the Size of Government		
Legal System and Property Rights	Marginal Effect	*p* Value
0	−0.678	0.055
1	−0.570	0.065
2	−0.463	0.082
3	−0.356	0.117
4	−0.249	0.198
5	−0.142	0.399
6	−0.035	0.825
7	0.073	0.650
8	0.180	0.315
9	0.287	0.169
10	0.394	0.108
Conditional Marginal Effects of the Legal System and Property Rights		
Size of Government	Marginal Effect	*p* Value
0	−0.175	0.573
1	−0.068	0.799
2	0.039	0.859
3	0.147	0.427
4	0.254	0.098
5	0.361	0.007
6	0.468	0.000
7	0.575	0.000
8	0.682	0.000
9	0.790	0.000
10	0.897	0.001

for size of government. At low scores for the size of government, the growth effects of legal system and property rights are statistically insignificant, but at higher score for size of government, the growth effects become positive and highly significant, both economically and statistically.

This result highlights the interdependency of the size of government and the quality of the legal system and property rights. It is unsurprising based on our cursory analysis of raw data in the beginning of the article, as well as the literature by de Soto (2000, 2006) and Friedman (2002), that the property rights and the legal system/rule of law measure is an important component of this interdependency.

As shown in Figure 2, the growth effect of the size of government also increases with the quality of the three other EFW areas (sound money, freedom to trade, and regulations), but the effects are statistically insignificant. Note that the conditional growth effects of the size of government are consistent with our measure of institutional imbalance. The EFW's score for size of government is more positively associated with growth rates when the quality of other institutional dimensions is high. Thus, increasing the size of government score (i.e., shrinking government size) may increase growth if balance is increasing among the areas (i.e., other institutional dimensions are also high-quality). Though, increasing the size of government score may decrease growth if balance is decreasing among the areas (i.e., other institutional dimensions are low-quality).

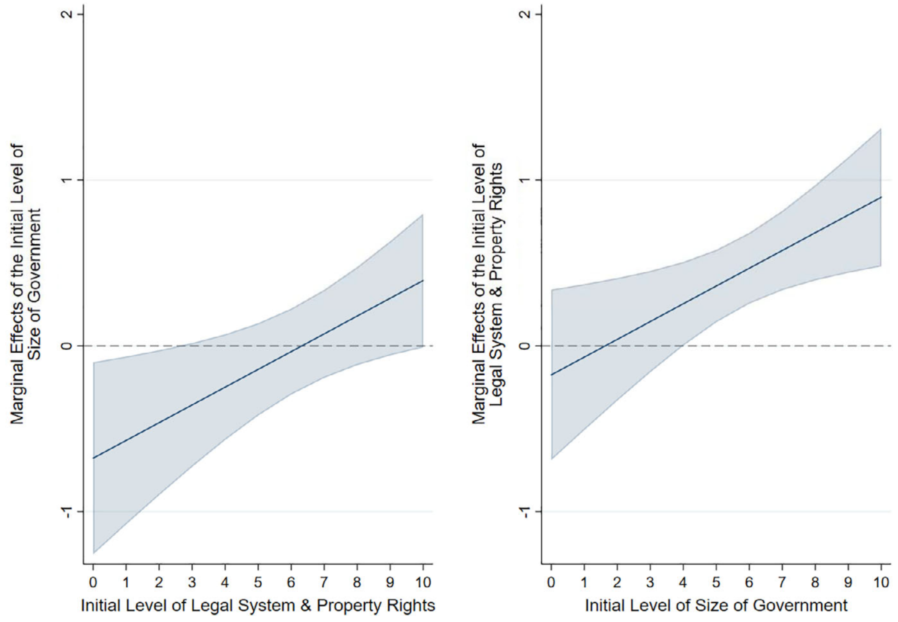

Figure 1. Conditional Marginal Effects of Size of Government and Legal System and Property Rights. [Color figure can be viewed at wileyonlinelibrary.com]

Table 9 and Figure 3 present a similar exercise except freedom to trade is substituted for size of government. The growth effects of increasing the score on freedom to trade range from statistically insignificant at low levels of legal system and property rights to negative and statistically significant at high levels of legal system and property rights. Thus, the counterintuitive result from Carlsson and Lundström (2002) of the negative impact for trade freedom occurs only for countries with very high levels of legal system and property rights. In addition, the growth effects of improving legal system and property rights by one point become less positive, and even insignificant, as levels of freedom to trade increase. Again, the quality of the legal system and property rights is an important component of the interdependency.

For completeness with regard to the score for legal system and property rights, we also provide the conditional growth effects of sound money and regulations on the quality of the legal system and property rights, presented in Figure 4. The growth effect of the EFW score for regulations is statistically insignificant at low scores for legal system and property rights. But, it is negative and statistically significant at scores at or above five for legal system and property rights. Similarly, the growth effect of sound money also decreases as the quality of the legal system and property rights increases, though the growth effects remain statistically insignificant regardless.

Thus, the growth effects of three EFW areas (sound money, freedom to trade, and regulations) decrease as the quality of legal system and property rights increases, and the growth effects of one EFW area (size of government) increase as the quality of legal system and property rights increases. Moreover, three out of the four EFW areas are statistically significant at some levels of legal system and property rights but not at other levels. These findings highlight the complex interdependencies among the EFW's areas and the need for a more exhaustive investigation of these interdependencies in future research.

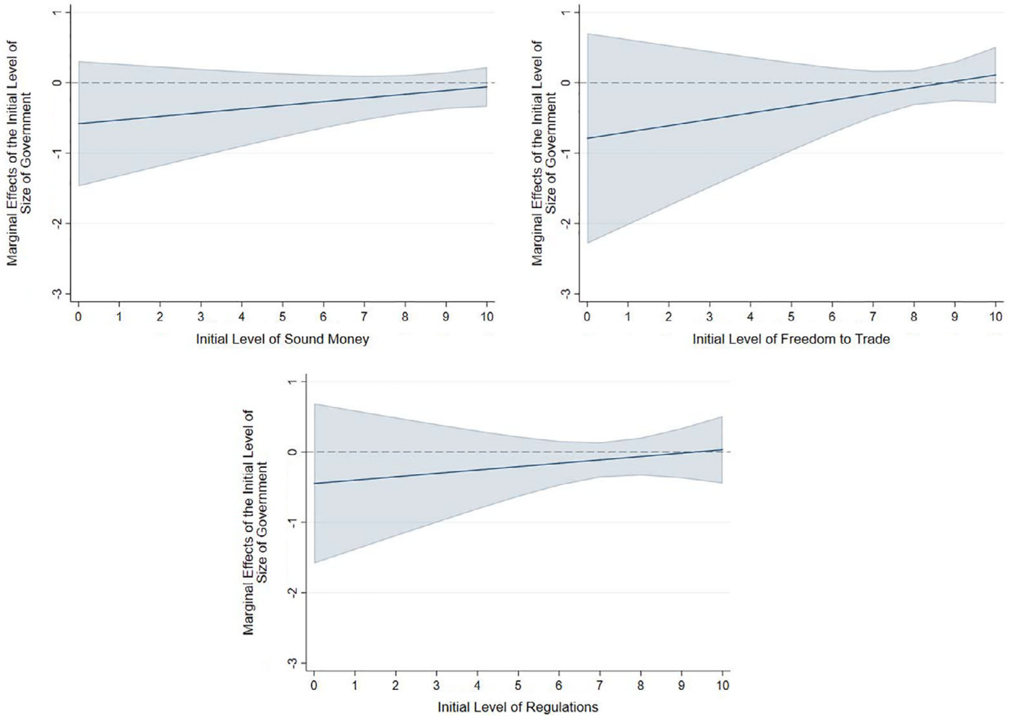

Figure 2. Conditional Marginal Effects of Size of Government and Other Institutional Dimensions. [Color figure can be viewed at wileyonlinelibrary.com]

While an exhaustive investigation of the conditional marginal effects of the EFW's components is beyond the scope of this article, this exercise accomplishes two things: (i) it provides evidence of the interdependencies underlying the imbalance measure developed earlier in the article and (ii) it suggests that some of the findings of the prior empirical literature may be misleading. We leave an exhaustive exploration of the conditional marginal effects of the EFW's areas to future researchers and suggest that an exhaustive exploration *may* require measuring the conditional effects of levels and/or changes of the EFW areas on levels and/or changes of other EFW areas in two-way, three-way, four-way, and/or five-way interactions.

Although the results are not presented here, we conducted a cursory analysis where we model growth as a function of all possible interactions among the changes in the EFW areas using cross-sectional data like that presented in Table 3. We test whether the interactions of the areas have an effect on growth and find that the interaction terms of all five areas have some level of joint significance in different specifications of the model. This suggests that it may be difficult to identify a simple model of the interactions among the EFW's areas. Though, modeling the relation among institutional dimensions theoretically may help determine which interactions might be most salient and why.

Despite the many possible interactions that could be explored in future empirical work, we conclude at a minimum that future growth models can (and should) be significantly improved by including a simple measure of imbalance (i.e., the SD of EFW's coexisting areas). And, researchers should give attention to the conditional marginal effects of the EFW's areas both theoretically and empirically prior to drawing conclusions about the empirical effects of changes in one area score.

Table 9. Conditional Marginal Effects of the Freedom to Trade and Legal System and Property Rights from Cross-sectional Growth Model

Conditional Marginal Effects of the Freedom to Trade		
Legal System and Property Rights	Marginal Effect	*p* Value
0	−0.036	0.953
1	−0.128	0.804
2	−0.221	0.617
3	−0.314	0.402
4	−0.407	0.206
5	−0.499	0.087
6	−0.592	0.042
7	−0.685	0.032
8	−0.778	0.037
9	−0.871	0.048
10	−0.963	0.062
Conditional Marginal Effects of the Legal System and Property Rights		
Freedom to Trade	Marginal Effect	*p* Value
0	1.298	0.077
1	1.205	0.060
2	1.113	0.043
3	1.020	0.026
4	0.927	0.012
5	0.834	0.004
6	0.741	0.000
7	0.649	0.000
8	0.556	0.001
9	0.463	0.028
10	0.370	0.191

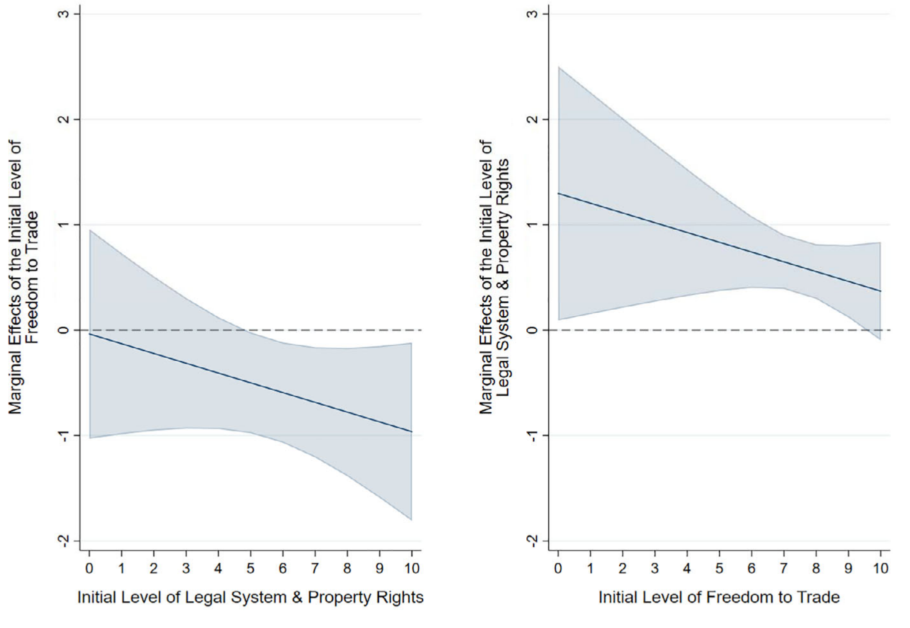

Figure 3. Conditional Marginal Effects of Freedom to Trade and Legal System and Property Rights. [Color figure can be viewed at wileyonlinelibrary.com]

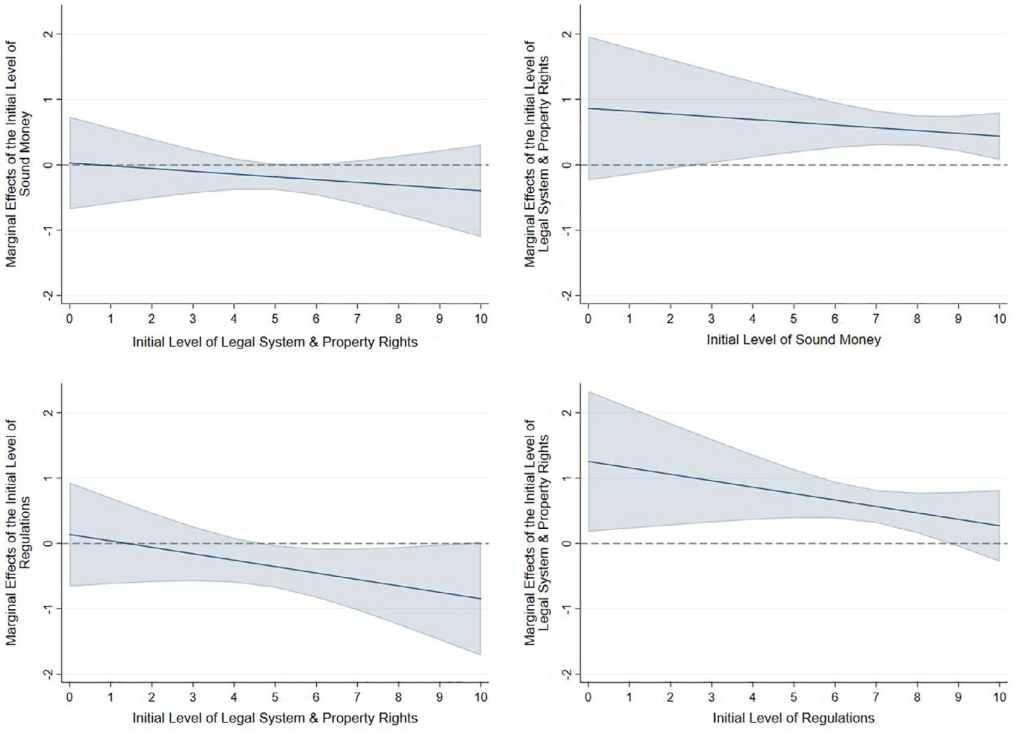

Figure 4. Conditional Marginal Effects of Sound Money and Regulations and Legal System and Property Rights. [Color figure can be viewed at wileyonlinelibrary.com]

7. Conclusion

Hundreds of articles have employed the EFW index to measure institutional quality without questioning the implicit assumption of the indices' construction—the perfect substitutability and independence of the five institutional dimensions. The idea that there could be important interrelationships is even clearly stated by the authors of the index itself in the publication. Simply put, prior empirical studies force an empirical assumption on their models that countries with the same overall average have similar economic outcomes. Therefore, a country with five area scores of 2, 4, 6, 8, and 10, which has an overall average of 6, would have similar economic outcomes to a country with all five area scores equal to 6. In practice, it forces an assumption that a country can make up for poor property rights by having better tax policy, for example. We find this is simply incorrect, and it matters a lot for explaining economic growth rate differences.

We show that the institutional dimensions within these indices are not independent of one another, but rather, they are interrelated. Thus, institutional weakness in one area cannot be compensated by improving other areas. Institutional areas seem to function like a "three-legged bar stool" in that weakness in any of the stool's legs (i.e., legal, monetary, regulatory, trade dimensions) cannot be remedied by strengthening the other legs—they must be balanced to work well. Therefore, the institutional production function for growth is not linear in the component areas.

We construct measures of these interrelations using the underlying components of the EFW index and see how they perform using one of the most cited growth models from the institutional literature as our baseline. We find these interrelations to be both statistically and economically significant; and their inclusion results in a rather dramatic increase in the explanatory power of the growth models on international data.

We find that including the SD of the within-country area scores for each country as an independent variable is perhaps the easiest and key method that future researchers can use to incorporate our contribution and properly specify their models in the future. Including multiple interaction terms may also address other interrelations among institutional dimensions and allow for exploration of the specific interrelations that exist. However, the inclusion of interaction terms necessitates the use of the disaggregated EFW index rather than the commonly used overall index, and it requires more careful interpretation of the coefficients on the institutional components to compute the partial derivatives properly. It also does not by itself control for the main issue we find regarding dispersion (or balance) of the area scores mattering. We however highlight a few of these interdependencies, including an important interaction between the area measure for the size of government and the area measure for the legal system and property rights that may help to solve a puzzle from the prior empirical literature about the effects of government size on economic growth.

Our results clearly show that the impact of an area is dependent on the other areas and how they compare to each other. The area scores are not linearly independent, nor linearly related. There are important multiplicative interactions between the areas. The impact of scoring one point higher in one area depends critically on the pre-existing scores the country has in the other areas. So, for example, improving the size of government score (shrinking government) may decrease growth unless the legal system and property rights are strong. Importantly our results confirm this is not simply corrected by including individual area scores, and our results are not explained by an alternative hypothesis that having one bad area score drives the results. Our results are robust to both panel and cross-sectional methods, as well as tests for omitted variable bias and use system GMM estimation to control for endogeneity.

The implications of these results for policy makers and academics are straightforward. Policy makers should pursue a balanced institutional structure, especially in countries where at least one dimension is particularly poor. Broad reforms lifting all areas a little will generally produce more growth than a large reform to only one area, even if it results in the same change to the overall average score. Improving the weakest area scores will also contribute more to growth than improving already strong areas even if they have the same impact on the overall average score. Additionally, academics should always include a correction for the imbalance among the five area scores when using the overall EFW index (or even the areas); and our strong suggestion is to, at least, employ a measure of the within-country SD of the area scores as an added independent variable in future empirical models.

Appendix A

Table A1. Summary Statistics (Cross-sectional Data 2000–2015)

Variable	Observations	Mean	SD	Min.	Max.
EFW (2000)	97	6.611	0.998	4.170	8.630
Area 1 (2000)	97	6.141	1.225	3.250	8.930
Area 2 (2000)	97	5.382	1.881	1.900	9.010
Area 3 (2000)	97	7.731	1.802	2.240	9.840
Area 4 (2000)	97	7.329	1.339	3.220	9.370
Area 5 (2000)	97	6.469	1.081	4.180	8.560

(Continues)

Table A1. Continued

Variable	Observations	Mean	SD	Min.	Max.
ΔEFW	97	0.241	0.598	−2.770	2.120
ΔArea 1	97	0.169	1.015	−3.100	2.540
ΔArea 2	97	0.058	0.784	−1.870	3.890
ΔArea 3	96	0.673	1.506	−3.620	6.690
ΔArea 4	97	−0.188	0.846	−4.420	2.460
ΔArea 5	97	0.528	0.787	−2.930	2.490
EFW area SD (mean)	97	1.475	0.391	0.643	2.443
EFW area range (mean)	97	8.421	1.039	5.574	9.694
ln(per capita real GDP) (2000)	97	8.521	1.635	5.431	11.445
ΔHuman capital (2000–2014)	97	0.273	0.142	−0.057	0.703
Per capita real GDP growth (mean)	97	2.216	1.854	−2.196	9.653
Private investment (% of GDP) (mean)	97	16.744	5.164	0.000	36.467
Public investment (% of GDP) (mean)	97	6.159	3.987	1.884	26.365
% Tropical population	97	0.455	0.482	0.000	1.000
% Coastal population	97	0.452	0.360	0.000	1.000
Malaria ecology	97	0.247	0.383	0.000	1.000

Notes: Change variables, denoted by Δ, are differences from 2000 to 2015. Variables with the notation "(2000)" are initial values from the beginning of the period. The SD and range of the EFW index are annual measures of dispersion in a country's EFW areas, which are then averaged.

Table A2. Summary Statistics (Panel Data 1980–2015)

Variable	Observations	Mean	SD	Min.	Max.
EFW	624	6.335	1.212	1.810	9.190
ΔEFW	624	0.297	0.593	−1.990	3.050
EFW area SD	624	1.507	0.509	0.279	3.239
EFW area range	624	8.012	1.250	3.210	9.850
ln(per capita real GDP)	624	8.548	1.541	5.391	11.561
Human capital	624	2.307	0.689	1.031	3.672
Per capita real GDP growth	624	1.837	2.684	−8.938	10.947
Total investment (% of GDP)	624	23.137	6.755	0.000	71.281
% tropical population	558	0.483	0.483	0.000	1.000
% coastal population	558	0.472	0.361	0.000	1.000
Malaria ecology	558	0.251	0.382	0.000	1.000

Note: Change variables, denoted by Δ, in the panel data are five-year changes. The SD and range of the EFW index are annual measures of dispersion in a country's EFW areas. Panel data do not include geography variables because fixed effects in the models eliminates the need to control for geography. System GMM models do include geography variables and, thus, have few observations due to data limitations.

Appendix B

Table B1. Robustness Tests of EFW Area SD to Income Levels and Sample Period

Dependent Variable: Per Capita Real GDP Growth	(1)	(2)	(3)	(4)	(5)
Independent variables	No high income	No low income	1980–2015	1980–2015 no low income	1980–2015 no high income
$EFW_{initial}$	0.162	0.030	0.090	−0.184	0.204
	(0.332)	(0.228)	(0.213)	(0.197)	(0.236)
ΔEFW	0.445	0.275	0.149	−0.053	0.180
	(0.346)	(0.247)	(0.148)	(0.147)	(0.152)
EFW area SD	−1.558***	−0.963***	−0.948***	−1.250***	−0.729**
	(0.536)	(0.330)	(0.291)	(0.337)	(0.319)
$\ln GDP_{initial}$	−0.888***	−1.049***	−0.843***	−1.185***	−0.924***
	(0.306)	(0.212)	(0.148)	(0.195)	(0.155)
Tropical Population	0.121	−0.126	−0.760*	−0.792	−1.128***
	(0.810)	(0.533)	(0.451)	(0.666)	(0.389)
Coastal population	−0.169	0.072	0.218	0.330	0.289
	(0.636)	(0.424)	(0.378)	(0.451)	(0.412)
Malaria ecology	−2.240**	−1.280	−1.824**	−1.896***	−1.049
	(0.989)	(1.180)	(0.809)	(0.636)	(1.041)
Private investment	0.097**	0.103***	0.084**	0.082**	0.088**
	(0.038)	(0.036)	(0.034)	(0.032)	(0.038)
Public investment	0.110***	0.090***	0.043	0.056	0.060
	(0.039)	(0.031)	(0.047)	(0.058)	(0.052)
Human capital	−0.147	0.200	1.142**	1.960***	0.983*
	(1.427)	(0.972)	(0.479)	(0.614)	(0.528)
Constant	11.065***	11.614***	8.643***	12.596***	8.187***
	(3.990)	(3.049)	(2.299)	(2.158)	(2.650)
Region fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	58	81	80	51	68
R-squared	0.672	0.696	0.745	0.850	0.742
Adjusted R-squared	0.555	0.620	0.681	0.786	0.661

Notes: The SD is computed as the overall averages of annual values. Robust standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provides GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.1$.

Table B2. Growth Effects of EFW Area SD and Components Scoring Below the Threshold

	(1)	(2)	(3)	(4)
Variables	Avg. GDP Growth	Avg. GDP Growth	Avg. GDP Growth	Avg. GDP Growth
EFW (2000)	0.395	0.375	0.103	0.140
	(0.287)	(0.284)	(0.284)	(0.287)
ΔEFW (2000–2015)	0.613**	0.635**	0.424	0.452
	(0.271)	(0.290)	(0.285)	(0.284)
EFW area SD	−1.614***	−1.613***	−1.218***	−1.260***
	(0.424)	(0.422)	(0.395)	(0.404)
EFW area threshold	0.387	0.372	−0.267	−0.201
dummy	(0.362)	(0.348)	(0.352)	(0.377)
Constant	8.898***	8.906***	9.746***	9.674***
	(2.607)	(2.626)	(2.724)	(2.758)
Threshold (percentile)	10th	15th	20th	25th
Observations	97	97	97	97
R-squared	0.638	0.638	0.636	0.636
Adjusted R-squared	0.560	0.560	0.558	0.557

Notes: Each model controls for initial income, malaria ecology, tropical population, proximity to a coast, average private and public investment rates and human capital. The SD is computed as the overall averages of annual values. Robust standard errors shown in parenthesis. The EFW data source is the Economic Freedom of the World: 2018 Annual Report by Gwartney et al. The World Bank's World Development Indicators (WDI) provides GDP data and investment data. OECD investment data substitutes for WDI data when WDI investment data are unavailable. Penn World Table 9.0 provides the annual measure of human capital. Geography data are from Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).
***$p < 0.01$,
**$p< 0.05$,
*$p < 0.1$.

Acknowledgments

The authors would like to thank Claudia Williamson, Randy Campbell, Cheng Li, Kathleen Thomas, Travis Wiseman, conference participants at the 2018 Southern Economic Association and Association of Private Enterprise Education Meetings, seminar participants at Texas Tech University's Free Market Institute, and two anonymous referees of this journal for helpful comments and suggestions.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. The colonial origins of comparative development. *American Economic Review* 91:1369–401.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics* 117(4):1231–94.

Altonji, Joeseph, Timothy Conley, Todd Elder, and Christopher Taber. 2011. *Methods for using selection on observed variables to address selection on unobserved variables*. Mimeo: Yale University.

Arellano, Manuel, and Stephen Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58(2):277–97.

Asteriou, Dimitrios, and Simon Price. 2001. Political instability and economic growth: UK time series evidence. *Scottish Journal of Political Economy* 48(4):383–99.

Baumol, William. 1990. Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy* 98(5): 893–921.

Berggren, Niclas, Andreas Bergh, and Christian Bjørnskov. 2012. The growth effects of institutional instability. *Journal of Institutional Economics* 8(2):187–224.

Berggren, Niclas, and Henrik Jordahl. 2005. Does free trade really reduce growth? Further testing using the economic freedom index. *Public Choice* 122(1–2):99–114.

Bergh, Andreas. 2020. Hayekian welfare states: Explaining the coexistence of economic freedom and big government. *Journal of Institutional Economics*: 16(1), 1–12.

Bjørnskov, Christian, and Nicolai J. Foss. 2008. Economic freedom and entrepreneurial activity: Some cross-country evidence. *Public Choice* 134(3–4):307–28.

Blundell, Richard, and Stephen Bond. 2000. GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews* 19(3):321–40.

Boettke, Peter J. 2001. *Calculation and coordination: Essays on socialism and transitional political economy.* New York, NY: Routledge.

Boettke, Peter, Emily Chamlee-Wright, Peter Gordon, Sanford Ikeda, Peter Leeson, and Russell Sobel. 2007. The political, economic, and social aspects of Katrina. *Southern Economic Journal* 74(2):363–76.

Bolen, J. Brandon, and Claudia R. Williamson. 2019. The path of reform: The consequences of institutional volatility. *Kyklos* 72(3):400–25.

Calcagno, Peter T., and Edward J. López. 2012. Divided we vote. *Public Choice* 151:517–36.

Campos, Nauro F., Menelaos G. Karanasos, and Bin Tan. 2012. Two to tangle: Financial development, political instability and economic growth in Argentina. *Journal of Banking & Finance* 36(1):290–304.

Carlsson, Fredrik, and Susanna Lundström. 2002. Economic freedom and growth: Decomposing the effects. *Public Choice* 112(3–4):335–44.

Cole, Julio H., and Robert A. Lawson. 2007. Handling economic freedom in growth regressions: Suggestions for clarification. *Econ Journal Watch* 4(1):71.

Collier, Paul. 2010. The political economy of natural resources. *Social Research* 77(4):1105–32.

Dawson, John W. 2003. Causality in the freedom–growth relationship. *European Journal of Political Economy* 19(3):479–95.

Dawson, John W. 2015. The empirical volatility growth relationship: Is economic freedom the missing link? *The Journal of Private Enterprise* 30(2):61–82.

de Soto, Hernando. 2000. *The mystery of capital: Why capitalism triumphs in the west and fails everywhere else*. New York, NY: Basic Books.

de Soto, Hernando. 2006. The challenge of connecting informal and formal property systems some reflections based on the case of Tanzania. In *Realizing Property Rights*, edited by H. de Soto and F. Cheneval. Zürich: Rüffer & Rub.

De Haan, Jakob, and Jan-Egbert Sturm. 2006. How to handle economic freedom: Reply to Lawson. *Econ Journal Watch* 3(3): 407–10.

Easterly, William. 2001. *The elusive quest for growth*. Cambridge: The MIT Press.

Easterly, William, and Ross Levine. 1997. Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics* 112(4):1203–50.

Ekelund, Robert B., and Robert D. Tollison. 2001. The interest-group theory of government, Chapter 17. In *The Elgar companion to public choice*, edited by W. F. Shughart II and L. Razzolini. Cheltenham: Edward Elgar, pp. 357–78.

Friedman, Milton. 2002. Economic freedom behind the scenes. Preface. In *Economic freedom of the world 2002 annual report*, edited by J. D. Gwartney, R. Lawson and C. Edwards. Vancouver, B.C.: Fraser Institute, pp. xvii–xxi.

Gallup, John L., Jeffrey D. Sachs, and Andrew D. Mellinger. 1999. Geography and economic development. *International Regional Science Review* 22(2):179–232.

Glaeser, Edward L., and Andrei Shleifer. 2002. Legal origins. *Quarterly Journal of Economics* 117(4):1193–229.

Glaeser, Edward L., Rafael Laporta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. Do institutions cause growth? *Journal of Economic Growth* 9(3):271–303.

Grier, Kevin, and Norman Maynard. 2016. The economic consequences of Hugo Chavez: A synthetic control analysis. *Journal of Economic Behavior & Organization* 125:1–21.

Gwartney, James D., Randall G. Holcombe, and Robert A. Lawson. 2006. Institutions and the impact of investment on growth. *Kyklos* 59(2):255–73.

Gwartney, James D., Robert A. Lawson, and Joshua C. Hall. 2017. *Economic freedom of the World 2018 annual report*. Vancouver, B.C.: Fraser Institute.

Gwartney, James D., Robert A. Lawson, Joshua C. Hall, and Ryan H. Murphy. 2018. *Economic freedom of the World 2018 annual report*. Vancouver, BC: Fraser Institute.

Hall, Joshua C., and Robert A. Lawson. 2014. Economic freedom of the world: An accounting of the literature. *Contemporary Economic Policy* 32(1):1–19.

Hayek, Friedrich A. 1944 [1994]. *The road to serfdom*. Chicago, IL: The University of Chicago Press.

Hayek, Friedrich A. 1945. The use of knowledge in society. *American Economic Review* 35(4):519–30.

Heckelman, Jac C., and Stephen Knack. 2009. Aid, economic freedom, and growth. *Contemporary Economic Policy* 27(1): 46–53.

Holcombe, Randall G. 1998. Tax policy from a public choice perspective. *National Tax Journal* 51(2):359–71.

Holcombe, Randall G. 2013. South Korea's economic future: Industrial policy, or economic democracy? *Journal of Economic Behavior & Organization* 88:3–13.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. The economic consequences of legal origins. *Journal of Economic Literature* 46(2):285–332.

Leeson, Peter T., Russell S. Sobel, and Andrea M. Dean. 2012. Comparing the spread of capitalism and democracy. *Economics Letters* 114(1):139–41.

North, Douglass C. 1990. *Institutions, institutional change and economic performance*. New York: Cambridge University Press.

North, Douglass C. 1991. Institutions. *Journal of Economic Perspectives* 5(1):97–112.

Olsen, Mancur. 1982. *The rise and decline of nations*. New Haven, CT: Yale University Press.

Oster, Emily. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37(2):187–204.

Powell, Benjamin, and Matt E. Ryan. 2006. Does development aid lead to economic freedom? *Journal of Private Enterprise* 22(1):1–21.

Powell, Benjamin, J. R. Clark, and Alex Nowrasteh. 2017. Does mass immigration destroy institutions? 1990s Israel as a natural experiment. *Journal of Economic Behavior & Organization* 141:83–95.

Rajan, Raghuram. 2019. *The third pillar: How markets and the state leave the community behind*. New York, NY: Penguin Press.

Robinson, James A., Daron Acemoglu, and Simon Johnson. 2005. Institutions as a fundamental cause of long-run growth. *Handbook of Economic Growth* 1A:386–472.

Roodman, David. (2015). xtabond2: Stata module to extend xtabond dynamic panel data estimator.

Sachs, Jeffrey D., and Andrew M. Warner. 1999. The Big Rush, natural resource booms and growth. *Journal of Development Economics* 59(1):43–76.

Shughart, W. F. II, ed. 1997. *Taxing choice: The predatory politics of fiscal discrimination*. New Brunswick: Transaction Publishers.

Smith, Adam. 1976 [1776]. *An inquiry into the nature and causes of the wealth of nations*. New York: Oxford University Press.

Sobel, Russell S. 2008. Testing Baumol: Institutional quality and the productivity of entrepreneurship. *Journal of Business Venturing* 23(6):641–55.

Sobel, Russell S., and Peter T. Leeson. 2007. The spread of global freedom. In *Economic freedom of the world: 2007 annual report*, edited by J. D. Gwartney and R. A. Lawson. Vancouver: The Fraser Institute.

Sobel, Russell S., J. R. Clark, and Dwight R. Lee. 2007. Freedom, barriers to entry, entrepreneurship, and economic progress. *Review of Austrian Economics* 20:221–36.

Sobel, Russell S. 2017. The rise and decline of nations: The dynamic properties of institutional reform. *Journal of Institutional Economics* 13(3):549–74.

World development indicators. 2018. *Washington*. DC: The World Bank.