

# IEO

Independent Evaluation Office  
of the International Monetary Fund

## BACKGROUND PAPER

---



BP/18-01/05

### **The IMF and Fragile States: Assessing Macroeconomic Outcomes**

Kevin Kuruc

**IEO Background Paper**  
Independent Evaluation Office  
of the International Monetary Fund

The IMF and Fragile States:  
Assessing Macroeconomic Outcomes

Prepared by Kevin Kuruc<sup>1</sup>

March 6, 2018

Abstract

This paper presents a quantitative analysis of the macroeconomic characteristics and performance of fragile states, especially in the context of their engagement with the International Monetary Fund. It finds, among other things: (i) fragility may be a more fluid state than previously documented; (ii) while in fragile states GDP growth is more volatile, it is only slightly slower, on average, than growth in non-fragile states; (iii) fragile states' GDP appeared to grow about 1 percentage point faster following approval of an IMF lending arrangement; and (v) foreign aid flows to fragile states increased by about 60 percent in the years following approval of IMF program engagement, with or without IMF financing (no such increase was observed for non-fragile states), illustrating the IMF's catalytic role. While this analysis provides a positive overall assessment of the IMF's role in fragile states, care must be exercised in interpreting the results, especially concerning causality.

The views expressed in this Background Paper are those of the author(s) and do not necessarily represent those of the IEO, the IMF or IMF policy. Background Papers report analyses related to the work of the IEO and are published to elicit comments and to further debate.

---

<sup>1</sup> Graduate student, University of Texas, Austin.

| Contents  | Page |
|---|------|
| Abbreviations .....   | iiv  |
| I. Introduction .....   | 1    |
| II. Quantifying the Characteristics of Fragile States .....   | 2    |
| A. Definition and Persistence of Fragility .....  | 2    |
| B. Macroeconomic Performance.....   | 4    |
| C. The IMF's Program Engagement with Fragile States.....  | 6    |
| III. Assessing the Economic Growth Characteristics of Fragile States When Data Quality Is Poor .... | 7    |
| A. Methodology to Use Satellites to Complement National Income Data on Activity .....               | 7    |
| B. Identifying the Economic Growth Characteristics of Fragile States .....                          | 9    |
| IV. Assessing the Impact of an IMF Arrangement on Fragile States.....                               | 10   |
| A. Event Study Methodology .....  | 11   |
| B. Quantifying the Impact of an IMF Arrangement on Economic Outcomes .....                          | 13   |
| V. Conclusion.....  | 17   |
| Figures   |      |
| 1. Persistence of Fragility, 2000–17 .....  | 4    |
| 2. Gross Financing Flows to Fragile States, 2000–15 .....   | 6    |
| 3. Examples of Luminosity Growth in Three African Countries, 1992 and 2013 .....                    | 8    |
| 4. Density of Augmented vs. <i>WEO</i> Growth Rates, 1993–2013 .....                                | 10   |
| 5. Volatility of Growth, 1992–2013.....   | 10   |
| 6. Evolution of GDP Surrounding IMF Financing, 2000–12 .....  | 14   |
| 7. Evolution of GDP for Fragile and Non-Fragile LICs, 2000–12.....                                  | 15   |
| 8. Evolution of ODA Surrounding IMF Arrangement, 2000–12 .....                                      | 16   |
| 9. Evolution of ODA for Fragile vs. Non-Fragile LICs, 2000–12 .....                                 | 17   |
| Tables  |      |
| 1. Number of Fragile States by Year, 2000–17 .....  | 3    |
| 2. Standards of Living in Fragile vs. Non-Fragile States Low Income Countries, 2014.....            | 4    |
| 3. Economic Performance of Fragile vs Non-Fragile States, 2000–16.....                              | 5    |
| 4. IMF Financial Engagement, 2000–16 .....  | 6    |
| Appendices  |      |
| 1. Fragile Classification .....   | 19   |
| 2. Mathematical Details of Satellite Use .....  | 20   |
| 3. Event Study Details .....  | 23   |
| References.....   | 25   |

**ABBREVIATIONS**

|            |  |
|------------|--|
| AU         | African Union  |
| CAR        | Central African Republic                                 |
| CPIA       | Country Policy and Institutional Assessment (World Bank) |
| EU         | European Union   |
| FCS        | countries in fragile and conflict-affected situations    |
| FGLS       | feasible generalized least squares                       |
| GDP        | gross domestic product                                   |
| IEO        | Independent Evaluation Office (IMF)                      |
| IMF        | International Monetary Fund                              |
| LICs       | low-income countries                                     |
| MICs       | middle-income countries                                  |
| NATO       | North Atlantic Treaty Organization                       |
| NOAA       | National Oceanic and Atmospheric Administration          |
| OAS        | Organization of American States                          |
| ODA        | official development assistance                          |
| OLS        | ordinary least squares                                   |
| PPP        | purchasing power parity                                  |
| PWT        | Penn World Tables  |
| SMP        | Staff-Monitored Program (IMF)                            |
| UN         | United Nations   |
| <i>WEO</i> | <i>World Economic Outlook</i> (IMF)                      |
| WDI        | World Development Indicators (World Bank)                |

## I. INTRODUCTION

1. This paper presents a quantitative analysis of the macroeconomic characteristics and performance of countries in fragile and conflict-affected situations (fragile states or FCS), especially in the context of their engagement with the International Monetary Fund (IMF). It provides supporting evidence for the IEO evaluation “The IMF and Fragile States.” It investigates, among other things: (i) the persistence and evolution of fragility in individual fragile states; (ii) the macroeconomic performance of fragile states; (iii) the trajectory of economic conditions in fragile states associated with IMF lending; and (iv) the responsiveness of foreign aid flows to IMF program engagement (with or without IMF financing) in fragile states. The study contributes to the ongoing, intense debate on the issue of state fragility, recognizing that 80 percent of the world’s needs for humanitarian assistance are driven by conflict and that, by 2030, nearly 50 percent of the world’s extremely poor are expected to live in countries characterized by fragility.<sup>1</sup>

2. At the outset, we must first explain how we identify fragile states for the purpose of this analysis. When identifying such countries, the IMF, along with many others, has broadly relied on the Harmonized List of Fragile Situations produced by the World Bank.<sup>2</sup> We too adopt this approach, but instead of using the yearly published lists, we use the World Bank’s stated criteria to recreate consistent lists going back to the year 2000. This is done for two reasons. First, the Harmonized List is only available from 2010 onwards and we would like to analyze a longer period. Second, some CPIA data were unavailable to us and some appear to have been updated since the creation of the yearly list. For consistency, we remake the FCS list for each year by including countries satisfying the stated criteria based on available data. Details on this procedure are provided in Section II.

3. Another aspect to keep in mind when analyzing the macroeconomic performance of fragile states is the quality of their national income data, which is generally considered to be questionable (Jerven, 2013). In an attempt to obtain a more reliable indicator of economic activity, we complement the national income data by employing a novel technique of utilizing satellite images of light visible from space as a proxy for economic activity, as recently pioneered in the academic literature (Chen and Nordhaus, 2011; Henderson and others, 2012). The relevant details, as well as the theoretical gains associated with this approach, are discussed in Section III.

4. The rest of the paper is organized as follows. Section II, after identifying a group of fragile states going back to 2000, analyzes the evolution of fragility in individual fragile states, the macroeconomic characteristics of fragile states as a group, and the extent to which the IMF has been engaged with fragile relative to non-fragile states. Section III explains the methodology of

---

<sup>1</sup> Fragility, Conflict, and Violence Group, World Bank. Conflict states can be thought of as a subset of fragile states.

<sup>2</sup> See <http://www.worldbank.org/en/topic/fragilityconflictviolence/brief/harmonized-list-of-fragile-situations>.

using satellite images to proxy economic activity, and applies this methodology to quantify the relationship between the variability of economic growth and a measure of fragility. Section IV uses event-study methodology to assess the impact of IMF lending on economic growth in fragile states and explore how foreign aid flows to fragile states respond to IMF program engagement. Section V presents conclusions. Finally, Appendices 1, 2, and 3, respectively, provide details on fragile state classification, the mathematics of the methodology to use satellite images to proxy economic activity, and the event-study methodology.

## II. QUANTIFYING THE CHARACTERISTICS OF FRAGILE STATES

### A. Definition and Persistence of Fragility

5. As a first step, a time-consistent means of identifying a fragile state is needed. Our definition broadly follows the criteria used to classify fragile states for the World Bank's harmonized list. We use a dynamic definition, that is, a classification of fragile states that changes from year to year, allowing countries to transition into and out of fragility. In contrast, a static list, which takes a list of fragile states in a given year (say 2015) and holds it fixed for the entire sample period, cannot be used to measure the persistence of fragility or even changes in economic performance over time. To see this, suppose that there is an economic surge associated with leaving fragility.<sup>3</sup> If the only countries classified as fragile over the previous 15 years are the countries that are still fragile at the end of the period, the true underlying distribution of growth outcomes will be severely underestimated, and, as the length of time from the year used to create the static list increases, the misrepresentation will become more severe. Countries that exited fragility must be included for a truly representative average.

6. This paper defines a country as fragile in a particular year if it meets one of the following two conditions. First, the country's average Country Policy and Institutional Assessment (CPIA) score (provided separately by the World Bank and the Asian Development Bank or the African Development Bank) is below a score of 3.2. The CPIA is the average of subjective rankings of 16 governance indicators intended to capture state capacity. We use the World Bank's CPIA scores if the scores are not available from the Asian Development Bank or the African Development Bank (see Appendix 1). Second, there has been a peacekeeping or peacebuilding operation in the country in the last three years.<sup>4</sup>

7. Using this definition, we produce a list of fragile states for each year from 2000 to 2017; 60 countries were classified as fragile at least at some point during these years, with the number fluctuating between 32 and 42 from year to year (Table 1).

---

<sup>3</sup> It has been observed that an end of conflict is often followed by a surge in economic activity.

<sup>4</sup> Missions by the United Nations (UN), African Union (AU), European Union (EU), Organization of American States (OAS), and North Atlantic Treaty Organization (NATO) are listed on the World Bank's website.

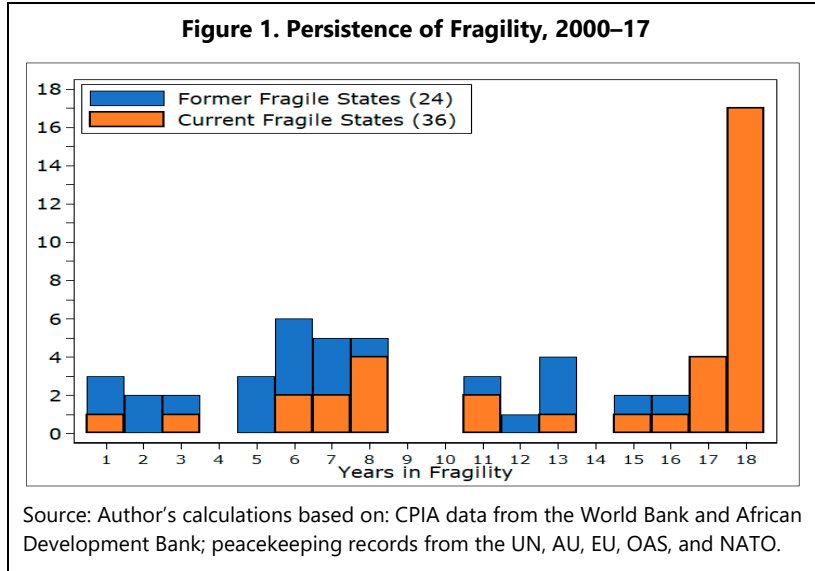
**Table 1. Number of Fragile States by Year, 2000–17**

| <b>Year</b> | <b>Number of Fragile States</b> |
|-------------|---------------------------------|
| 2000        | 41                              |
| 2001        | 39                              |
| 2002        | 39                              |
| 2003        | 36                              |
| 2004        | 40                              |
| 2005        | 40                              |
| 2006        | 37                              |
| 2007        | 34                              |
| 2008        | 34                              |
| 2009        | 32                              |
| 2010        | 34                              |
| 2011        | 37                              |
| 2012        | 37                              |
| 2013        | 42                              |
| 2014        | 40                              |
| 2015        | 39                              |
| 2016        | 35                              |
| 2017        | 36                              |

Source: Author's calculation.

8. Using the dynamic lists of fragile countries, we estimate a histogram of total years that a fragile state was labeled as fragile between 2000 and 2017 (Figure 1). Seventeen countries were labeled as fragile in each of the 18 years; for these countries fragility seems an almost permanent state. However, 24 previously fragile countries are not on the 2017 (or FY 2018) harmonized list. While these data cannot be used to calculate the average length of fragility, the changes highlight the importance of using a dynamic definition for analysis of fragility.<sup>5</sup> Several academic works conclude that the level of state capacity is a highly persistent variable (Collier and Chauvet, 2008; Pritchett and others, 2010). The fluid nature of fragility shown in Figure 1 is not meant to challenge these claims. Transitions will artificially look more common when using an arbitrary cutoff point (such as a CPIA of 3.2) derived from an underlying continuous input, especially if the input variable has high-frequency fluctuations. This is an unfortunate reality for any analysis that compares groups determined by a continuous underlying variable.

<sup>5</sup> Suppose that a country has been classified as fragile for the entire 18-year period. We cannot determine the "true" length of fragility until it ends. Therefore, the observed sample does not allow computation of the true average length.



## B. Macroeconomic Performance

9. Based on the dynamic lists of fragile states, we compare key macroeconomic variables across fragile and non-fragile states. First, Table 2 presents aspects of living standards in fragile vs. non-fragile low-income countries (LICs) for 2014.<sup>6</sup> Here, GDP per capita is taken from the IMF's *World Economic Outlook (WEO)* database, but the averages are qualitatively similar in the Penn World Tables, the leading source among academic researchers for comparisons of purchasing power parity (PPP)-adjusted GDPs. "Access to electricity" and "mortality rate" come from the World Bank's World Development Indicators (WDI). The uncertainty underlying these data will be discussed in more detail later in the paper; for now, these will be regarded as the best available estimates.

|                               | <i>Variable</i>           | <i>Fragile</i> | <i>Non-Fragile</i> |
|-------------------------------|---------------------------|----------------|--------------------|
| <i>Unweighted</i>             | GDP per capita (PPP \$)   | 2241           | 4535               |
|                               | Access to electricity (%) | 42.4           | 62.3               |
|                               | Mortality rate (per 1000) | 9.0            | 7.5                |
| <i>Weighted by Population</i> | GDP per capita (PPP \$)   | 2311           | 2632               |
|                               | Access to electricity (%) | 42.9           | 44.7               |
|                               | Mortality rate (per 1000) | 9.1            | 7.3                |
|                               | Observations              | 30             | 33                 |

Source: Author's estimates based on the *WEO* database; World Development Indicators.

<sup>6</sup> The year 2014 is the latest year for which the World Development Indicators provides consistent information on mortality and electricity.



10. Two contrasting observations can be made about these estimates for LICs. First, if we use a simple average, GDP per capita in fragile states is approximately half as large as that in non-fragile states, and the share of population with access to electricity is two-thirds as high. Second, however, when weighted by population, the gaps between fragile and non-fragile states diminish significantly. In fact, given the uncertainty surrounding these estimates, one cannot be confident that those living in a fragile state are poorer on average than those living in a non-fragile state.

11. Table 3 takes advantage of the *dynamic* lists of fragile states constructed to report the average macroeconomic performance of a fragile relative to a non-fragile state over the 2000–16 period.<sup>7</sup> Given the dynamic definition, the table captures the trajectory of macroeconomic performance of a yearly cohort of fragile states. Overall, GDP growth seems to be somewhat lower for fragile states, though subject to a large standard deviation.<sup>8</sup> Inflation is higher and external debt larger, while tax revenue is lower. However, fragile states experience smaller current account deficits, likely reflecting the fact that they have more limited access to foreign borrowing. These conclusions remain the same (except that external debt looks more comparable) if we use medians rather than means.

|   | Fragile | Non-Fragile |
|---|---------|-------------|
| GDP growth (in percent)                     | 3.7     | 4.5         |
| Inflation (in percent)                      | 9.7     | 6.3         |
| External debt (in percent of GDP)           | 74.6    | 55.6        |
| Tax revenue (in percent of GDP)             | 12.1    | 16.0        |
| Current account balance (in percent of GDP) | -6.1    | -9.3        |
| Observations                                | 427     | 606         |

Source: WEO database.

12. One omission from Table 3 is the volatility of growth. LICs, and fragile states in particular, are typically thought to suffer from less stable growth. However, this claim is statistically indistinguishable from these countries having “noisier” estimates of GDP. If fragile states, with their lower administrative capacity, are providing less stable estimates of GDP growth, then growth could appear more volatile regardless of the true underlying pattern. The statistical details will be dealt with later, but this paper shows below that fragile states do seem to experience greater growth volatility based on data that are independent of the reliability of measurement of national accounts.

<sup>7</sup> As an example, Cambodia, which is listed as fragile for 2000–07, is included in the “non-fragile” column for 2008–16.

<sup>8</sup> Assuming that the data are perfectly accurate, and that clustering should be done at the country level, this claim is not significant at the 5 percent level, but it is so at the 10 percent level. If either of these assumptions does not hold, there is even greater uncertainty about the validity of this claim.

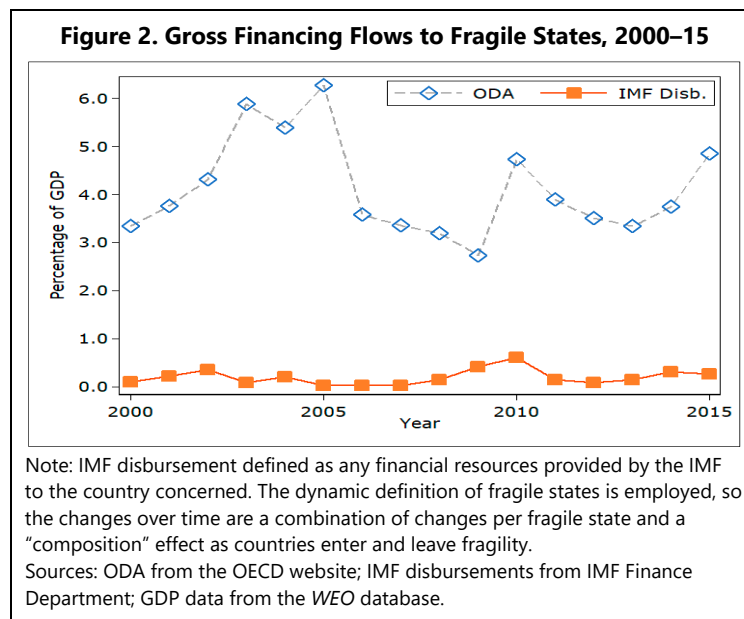
### C. The IMF's Program Engagement with Fragile States

13. We now examine the IMF's interaction with the donor community in fragile states. To begin with, it is relevant to know whether fragile states have received financing from the IMF proportionately more or less than non-fragile states. Table 4 reports the shares of country-years in which the IMF disbursed any funds, for fragile and non-fragile states, during 2000–16.

|                          | <b>Fragile<br/>(LICs &amp; MICs)</b> | <b>Non-fragile<br/>(LICs &amp; MICs)</b> | <b>Fragile<br/>(LICs only)</b> | <b>Non-fragile<br/>(LICs only)</b> |
|--------------------------|--------------------------------------|--|--------------------------------|------------------------------------|
| Share with IMF financing | 35                                   | 23                                       | 38                             | 44                                 |
| # Observations           | 646                                  | 2234                                     | 482                            | 688                                |

Note: "Share with financing" computed as the count of country-years with any IMF disbursement divided by total observations for each category.  
Source: Author's estimates based on data from IMF Finance Department.

14. Of the fragile state country-years, 35 percent involved some IMF financing, compared to 23 percent for all non-fragile LIC and MIC country years. However, focusing on LICs only, we find an opposite pattern: 38 percent of fragile LIC country-years involved some IMF financing, compared to 44 percent for non-fragile LICs. Their smaller share may represent the fragile states' greater difficulty in agreeing on a lending arrangement with the IMF or a greater propensity of FCS arrangements to go off-track, given the lack of administrative capacity and the more challenging political environments in these countries. It is also possible that fragile states, especially in a post-conflict situation, received generous official development assistance (ODA) from donors, minimizing the need for IMF financing. This last possibility is consistent with the relatively small financing role the IMF has played in fragile states (Figure 2).



### **III. ASSESSING THE ECONOMIC GROWTH CHARACTERISTICS OF FRAGILE STATES WHEN DATA QUALITY IS POOR**

#### **A. Methodology to Use Satellites to Complement National Income Data on Activity**

15. The poor quality of national income data in many developing countries is a constant concern of the international development community (Chen and Nordhaus, 2011; Henderson and others, 2012; Jerven, 2013; Johnson and others, 2013). In an analysis of the Penn World Tables by Johnson and others (2013), for instance, revisions between versions 6.1 and 6.2 created such a large variation that the same countries appeared on a list of top-10 and bottom-10 performers in Africa over the same 25-year period. Massive revisions are common in national income data. The most extreme example is for Equatorial Guinea, which was considered a bottom-10 performer in version 6.1 but was the second-highest performer in version 6.2.

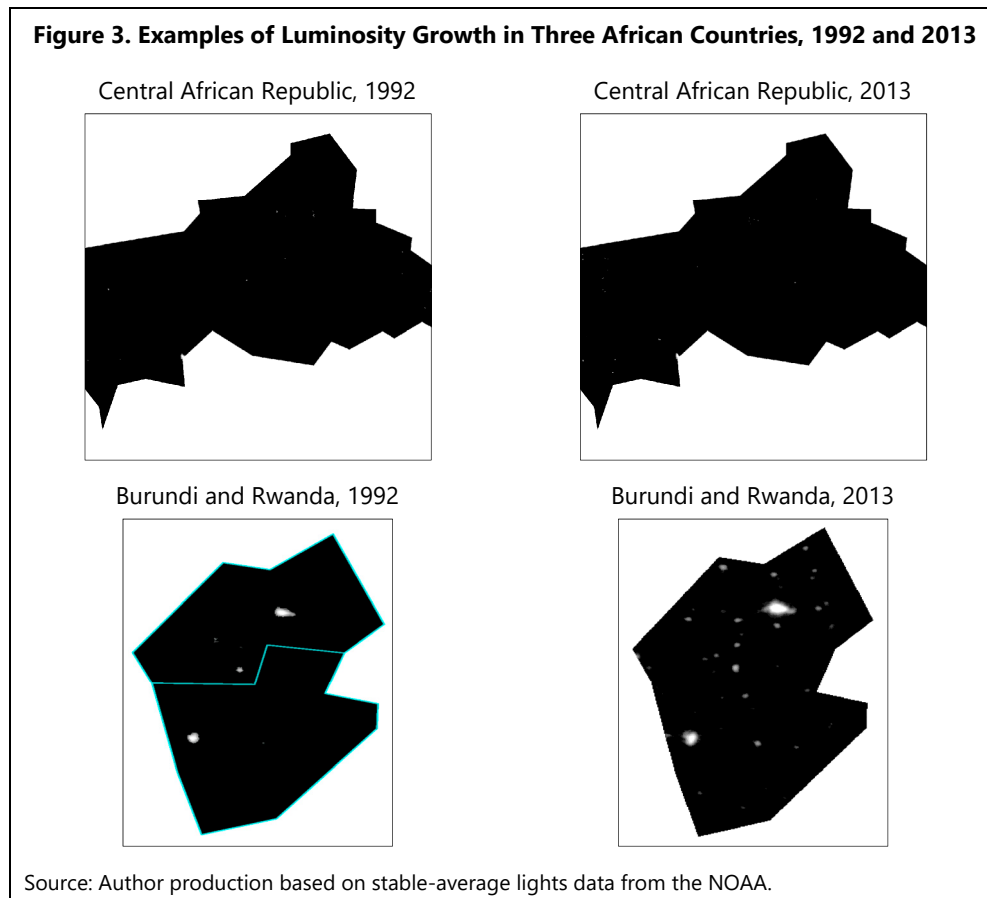
16. Poor data quality is particularly problematic when comparing growth rates. For example, it is substantially less likely that measurement error could lead us to erroneously conclude that the United States is wealthier than Haiti than to erroneously conclude that the United States grew faster than Haiti in a given year. Over a large enough sample, measurement error in growth rates should approximately cancel out, assuming that the mean measurement error is zero. However, when using smaller samples of fewer than 50, as will be done below when assessing the impact of IMF arrangements on fragile states, increasing the precision of the dependent variable can yield substantial improvements in statistical power.

17. In this paper, lights visible from space will be used as an independent source of economic activity estimates. It is now well known that there is a high correlation between how bright a country is in satellite imaging—what will be referred to as “luminosity”—and its level of economic activity (Chen and Nordhaus, 2011; Henderson and others, 2012). Not only is long-run development visible from space (e.g. more transport networks, broader access to electricity), but nearly all economic transactions at night require some light. Consider retail shops staying open later when shopping increases, or a manufacturing plant working overtime hours to fill an influx of orders. While this relationship is bound to have its own measurement error, if some new information is provided from a wholly independent source it can be leveraged to reduce total error (even if the new estimate has substantially more measurement error than the original).

18. Henderson and others (2012) and Chen and Nordhaus (2011) provide useful discussions of what exactly the satellite data are, and the discussion here will closely follow these studies. The data come from the Earth Observation Group in the United States National Oceanic and Atmospheric Administration (NOAA). Each data point is a small pixel with a luminosity score between 0 and 63. To appreciate how fine these pixels are, consider that the surface area of the United States yields more than 16 million pixels. For each country, the pixels are averaged to produce a country-specific luminosity score; this is an appropriate method because the pixel scores are designed to be comparable in proportion (i.e., a pixel with a luminosity score of 60 is approximately twice as bright as a pixel with a score of 30). The country-specific luminosity scores

will be the metric used in this paper. Given the obvious role that population density, technology, and culture play in relating luminosity to GDP across countries, only changes in luminosity will be considered. The raw data are yearly images, available from 1992 to 2013, created by the NOAA, which takes averages across all cloud-free images of that year. In some years, more than one satellite performs this operation; for the purposes of this paper their output is averaged.

19. Before constructing economic activity estimates from satellite images, two visual examples will be useful (Figure 3). It is difficult to visually observe small changes in luminosity, so these examples will be over the entire 1992–2013 horizon. Consider first the Central African Republic (CAR). This country is estimated to have had a cumulative GDP growth rate of negative 5 percent and 8 percent, respectively, by the *WEO* database and the Penn World Tables (PWT). Unfortunately, the CAR is extraordinarily dark to begin with (and end with, given the lack of growth), so the exact luminosity is difficult to see. Fortunately, not many countries experienced such weak growth over this horizon.



20. Burundi and Rwanda, by contrast, experienced much more robust GDP growth over this period. Burundi's GDP is estimated to have grown by 32 percent and 53 percent, respectively, by

the *WEO* database and by the PWT, and Rwanda's by 100 percent and 87 percent, respectively.<sup>9</sup> As can be seen in the contrasting images, much more light growth has taken place in these countries than in the CAR, and more in Rwanda than in Burundi—which is consistent with a positive relationship between lights and GDP growth.

### **B. Identifying the Economic Growth Characteristics of Fragile States**

21. To construct a proxy for growth in economic activity, by combining national income data on GDP growth with luminosity data (see Appendix 2 for the mathematical details), we:

- (i) Estimate the lights–GDP relationship;
- (ii) Generate a composite measure of light growth for each year to mitigate measurement error from satellites;
- (iii) Take these composite light-growth measurements to generate alternative GDP estimates strictly from satellite data, using the relationship from step (i);
- (iv) Use outside sources to estimate the error in national accounts to construct the optimal weight on predicted growth and reported growth for composite estimates;
- (v) Feed the predicted growth from satellite data and the growth estimates from national income data into this weighting for new, more accurate GDP growth estimates.

22. While this may seem like a convoluted process, the steps are grounded in statistical and econometric theory and are well known to increase statistical power. In a setting with highly uncertain estimates of GDP and a small sample, any gain in statistical power is critical. It turns out that the resulting revisions are fully consistent with the predictions of the statistical theory.

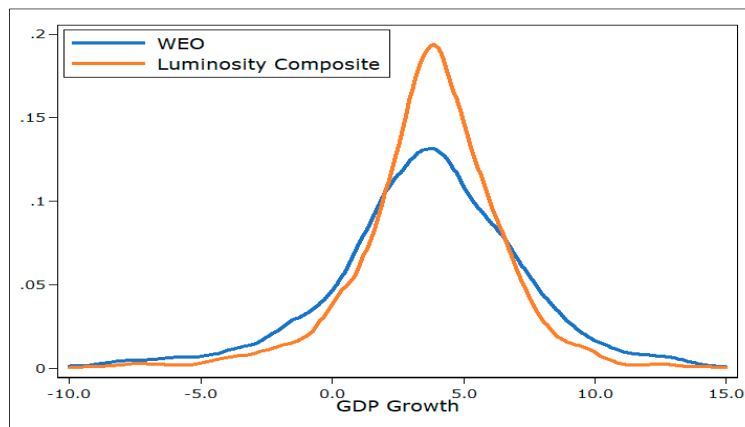
23. Figure 4 depicts the density functions of economic growth as reported by the *WEO* database and the resulting augmented growth when luminosity data supplements these estimates.<sup>10</sup> In the augmented estimates, there is a larger mass surrounding the mean (somewhere around 4 percent). With measurement error in reported GDP growth, reversion to the mean would be expected; countries reporting deep recessions, for example, are likely to have had a negative reporting error, and vice versa. Satellite data will help correct the cases where luminosity growth does not support that the country suffered a deep recession.<sup>11</sup>

---

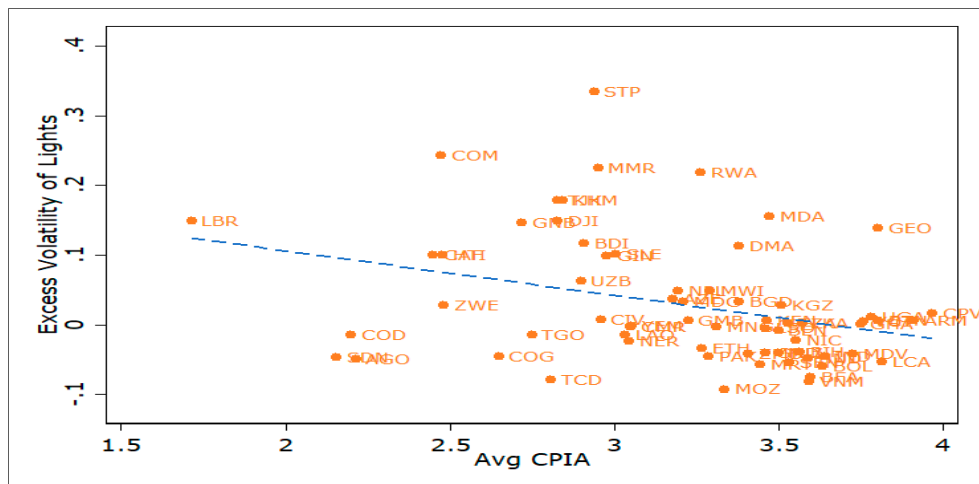
<sup>9</sup> To be clear, an estimate of 32 percent growth is still poor over a 20-year horizon.

<sup>10</sup> The tails have been cut off so the difference is visible near the mean. There are extreme growth spurts and recessions recorded in the data that appear less extreme when satellite data are used to augment reported growth data.

<sup>11</sup> To be clear, luminosity does not always revert to the mean. Some growth episodes are revised upwards due to exceptionally strong luminosity growth. Figure 4 does not imply the same ordering of countries, but shows that if satellite data do in fact correct measurement error, the expected true distribution should be “tighter” around the mean.

**Figure 4. Density of Augmented vs. WEO Growth Rates, 1993–2013**

Note: Sample period covers 1993–2013 when satellite estimates are available. Details of combining luminosity data and GDP data presented in Appendix 2.  
 Source: Author's calculation based on the WEO database and luminosity data from the NOAA.

**Figure 5. Volatility of Growth, 1992–2013**

Notes: Average CPIA over the sample against average “excess volatility.” Countries with higher growth rates will have a higher variance due to scaling; a more meaningful measure would be relative to average growth. “Excess volatility” here is computed as the residuals of volatility from a regression with volatility as the independent variable and average growth as the dependent variable. Sample covers all years with satellite data, 1992–2013; 3-letter country codes follow the IMF convention.

Sources: Stable-average lights from the NOAA; CPIA scores from the World Bank and African Development Bank.

24. One immediate result that can be confirmed without any econometrics is a difference between countries in the volatility of growth rates. Since there is little reason to believe that measurement error coming from satellite data is substantially worse for fragile states, if the volatility of luminosity growth is larger for fragile states this provides a good check for the results relying on national income data. Figure 5 plots the volatility of luminosity growth against CPIA scores (the continuous measure of state capacity), and confirms that a higher CPIA score tends to be associated with larger volatility of luminosity growth.

## IV. ASSESSING THE IMPACT OF AN IMF ARRANGEMENT ON FRAGILE STATES

### A. Event-Study Methodology

25. With revised estimates of GDP in hand, one can investigate the evolution of variables surrounding the approval of an IMF arrangement. The variables of interest will be GDP growth—to investigate the link between IMF financing and basic economic activity—and official developmental assistance—to see if donors respond to IMF program engagement (with or without IMF financing). To preview the results, there appears to be an increase in economic growth after an IMF lending arrangement is approved and a strong increase in ODA following the start of IMF program engagement.

26. The event-study methodology we use to obtain these results does not allow us to make strong causal statements.<sup>12</sup> To see why, consider that to justify a causal statement requires establishing the relevant counterfactual, such as “what would have been the path of GDP in a country had the IMF not provided financing?” If this could be established, it would be straightforward to calculate the average impact of IMF lending by subtracting the observed outcome from the counterfactual. The difficulty in making such a statement becomes immediately obvious: the IMF presumably chooses to lend to a country precisely because the country has a negative projected path that can hopefully be reversed.

27. Past attempts to get around the problem of assigning causality fall into two types. One is “selection correction,” which attempts to “control” for the likelihood of the IMF lending to a country with the same observable characteristics as the country of interest. The assumed counterfactual are countries that had similar economic trajectories but did not receive IMF financing. This approach essentially relies on observing two countries with identical projected outcomes, where the IMF has provided financing to only one of them for reasons uncorrelated with these economic projections.<sup>13</sup>

28. The second approach to “control” for the selection problem and construct a relevant counterfactual is an instrumental variable (IV) strategy (e.g., Barro and Lee, 2005). Deaton (2011) has an excellent discussion on why such strategies are flawed, though in a different context. Applying Deaton’s critique to our concerns, we see that the use of an IV strategy would require a variable that is correlated with IMF lending but uncorrelated with economic activity other than through IMF lending. Common instruments in this field are variables such as “political connections to the United States” because these countries are argued to be more likely to receive IMF

---

<sup>12</sup> A “causal” statement being: “IMF engagement caused growth to increase by 3 percent.”

<sup>13</sup> Technically, the mathematics does not require comparisons of two countries with identical projections because the functional forms imposed allow for extrapolative comparisons. This explanation is useful for thinking through the general assumptions, however.

financing.<sup>14</sup> This may satisfy the first condition, “correlated with IMF decisions,” but being tied politically to the U.S. may clearly impact economic activity through other channels.

29. Once the inherent flaws in these strategies are realized, it is not at all surprising that the statistical findings in the literature vary widely from “IMF financing significantly reduces growth” (Barro and Lee, 2005; Prezworski and Vreeland, 2000) to “IMF financing significantly increases growth” (Dicks-Mireaux and others, 2000). Applying an incorrect counterfactual leaves no promise of recovering anything close to the “true” impact of IMF lending. This paper will avoid making assumptions about the counterfactual path and instead just provide graphical evidence for what it may look like. This conservative approach cannot lead to strong causal claims, but this paper takes the stance that this is a problem with reality, not with any specific statistical technique.

30. The numerical technique in this methodology will be to compute a (weighted) average path of the change in some variable of interest surrounding the approval of an IMF arrangement for fragile states.<sup>15</sup> This technique normalizes the year of the “event,” that is, the approval of an IMF arrangement, to be year  $t = 0$  for all countries that have such an event. The average increase in GDP growth across countries in the year following the arrangement,  $t + 1$ , is then computed. This number can be interpreted, without a counterfactual assumption, as the growth increase that one would expect from an IMF arrangement outside of the sample used for estimation.<sup>16</sup> For this reason, the word “expected” will be used to represent the computed weighted-average path.

31. Event studies are conducted for economic growth and official aid inflows. What makes an event-study methodology telling is plotting the difference in the variable before and after the approval of an IMF arrangement. If GDP growth was trending upwards prior to an arrangement and continues trending up, it is difficult to argue that the IMF arrangement “caused” the observed increase. If there was no trend prior to an IMF arrangement, but the same increase is observed afterwards, a much stronger case can be made that this growth should not have been anticipated without the IMF arrangement. Hence, readers can judge the plausibility of the counterfactual for themselves by seeing the average pre-trend.

---

<sup>14</sup> In fact, this claim passes statistical tests based on conditional correlations.

<sup>15</sup> A weighted average is used to mimic feasible generalized least squares (FGLS). The details are presented in Appendix 3.

<sup>16</sup> This can be thought of as representing the increase in GDP growth rates that would be expected if the IMF engaged in an arrangement with a fragile state today (assuming no further information). It is argued that this is a more informative metric of past “success” than is the simple average of observed outcomes, for reasons presented in Appendix 3.



## B. Quantifying the Impact of an IMF Arrangement on Economic Outcomes

### Economic growth

32. The observations used for GDP growth results are all IMF lending arrangements with fragile states, as defined by the dynamic definition discussed in Section II, that were approved from 2000 to 2012. The observations cover only up to 2012, given the constraints on satellite data, as well as the need for a sufficient “post” period to observe any changes that may or may not have taken place. Only 38 IMF arrangements meet these criteria. The GDP series is the weighted average of growth as observed by satellite data and growth as reported in the Penn World Tables.<sup>17</sup>

33. Countries that received multiple successive lending arrangements are a complicating factor in the analysis. For instance, Burundi has three separate observations for 2004, 2008, and 2012. Statistically, these three lending arrangements are treated as three independent lending events and GDP outcomes. However, if the IMF is repeating programs for countries that continue to struggle, dropping these observations would increase the average reported change in growth. Likewise, countries with programs that have gone “off-track”—that is, whose loan disbursements stopped prior to their originally negotiated terms, because the country was not complying with the agreed program—are left in the analysis. In both cases, the lending arrangements are plausibly less likely to be successful than the average lending arrangement, and hence should not be the cause of a spuriously observed increase in growth outcomes. For the sake of a sufficiently large data series, and with the aim of being objective, these cases are included.

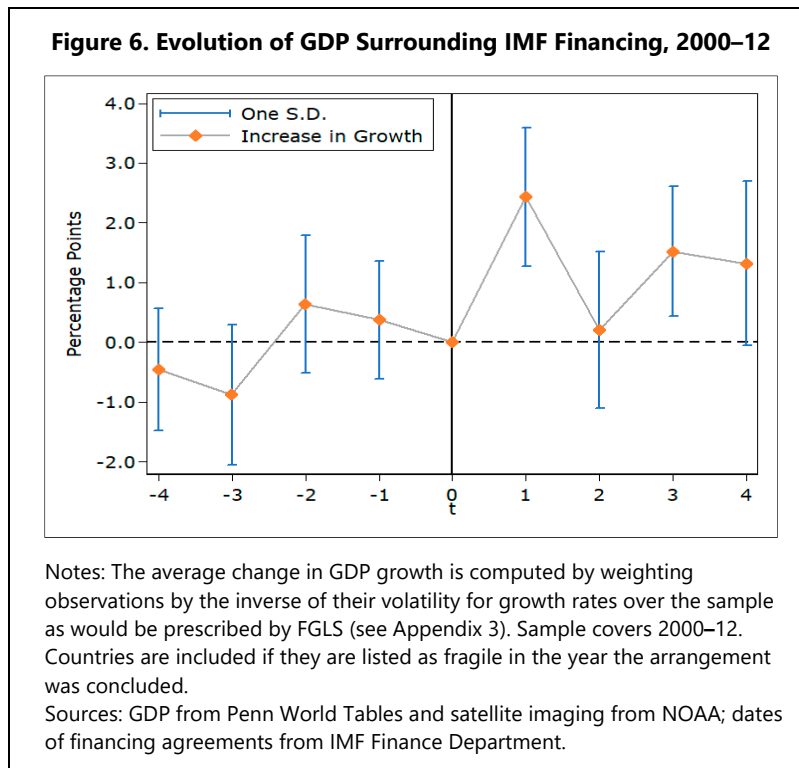
34. Figure 6 depicts the main results, showing the pattern of GDP growth rates before and after the approval of a lending arrangement for fragile states. Notice first that the value is zero at time zero ( $t = 0$ ). This is by construction; the values here are relative to growth at the time of Fund program approval. As shown in Figure 6, three years after approval, the expected growth rate is about 1.5 percentage points higher than it was in the year lending began. The blue extensions surrounding these points are standard error bands to provide a measure of statistical confidence.

35. While the technique does not tell exactly what the average change in growth is, because different countries had different experiences, it provides approximately 70 percent confidence that it is contained within these bands. Given the fact that all confidence intervals overlap with the dotted “zero” line in the “pre-period,” the weighted averages suggest that growth does not appear to be systematically lower or higher leading into an IMF arrangement. However, it seems

---

<sup>17</sup> We choose not to use the *WEO* data to avoid any possible implicit correlation between measurement error in *WEO* GDP and lending decisions for countries with limited national GDP data. Using an outside source for GDP helps to avoid this.

reasonable to believe that after the start of an arrangement the expected change in growth is positive.<sup>18</sup>

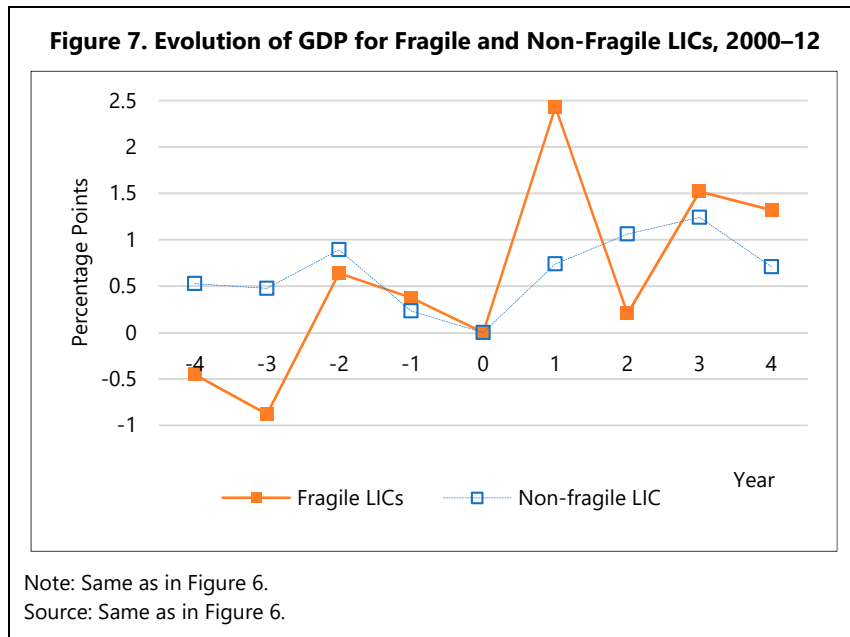


36. Next, Figure 7 compares the evolution of GDP surrounding IMF lending in fragile versus non-fragile low-income countries. It indicates that in fragile states, IMF financial support was associated with a clear turning point (from a period of decline or stagnation to a strong recovery), as noted above, but that in non-fragile LICs the growth acceleration following IMF lending was much less marked. These contrasting experiences may be because fragile states often approached the IMF following a prolonged and economically damaging conflict, while non-fragile states came to the Fund when they faced a less deep cyclical setback (such as a commodity-price downturn).

37. While these results are encouraging for the IMF's role in fragile states, they do not necessarily imply that IMF lending per se had a causal impact. For example, if the IMF begins lending at the close of civil conflict, it is likely that growth would increase regardless. This data point would contribute to the dynamic pattern above, without contributing any information about whether the IMF's efforts in fact helped. There are many such possibilities that are

<sup>18</sup> While a confidence interval of 70 percent is not typically what is used to assess "statistical significance," it is not necessarily the case that a test would reject a null hypothesis that "growth is not higher following IMF lending." The appropriate test for this claim is an F-test on all four coefficients being zero. Further, the claim "reasonable to believe" should not necessarily rely on a 5 percent level of significance; while the data constraints of course preclude strongly rejecting a null-hypothesis, this does not mean that the best guess given the data is that there is no effect.

consistent with these statistical results. However, the results weigh heavily against the claim that IMF lending to fragile states had a negative impact on economic activity. Taken in context, it is wise to interpret these results with cautious optimism in favor of the IMF's role.



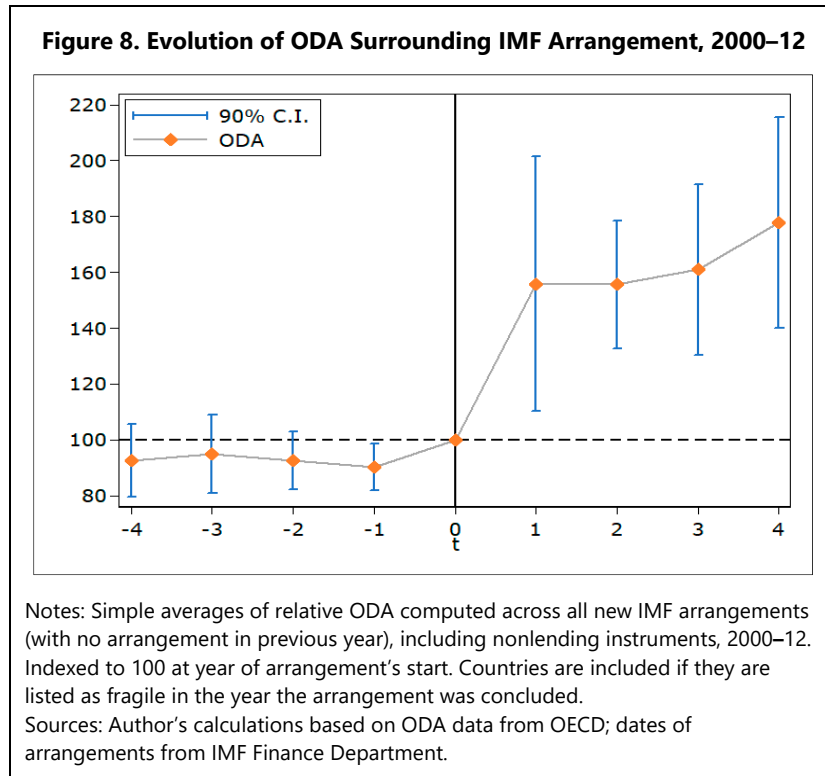
### Catalytic effect on aid inflows

38. A second question of importance for the IMF's role in fragile states concerns what has been termed the "catalytic effect" on donor assistance, the idea that the IMF's financial or program engagement catalyzes additional concessional financing or grants from donors. This effect could result from the signaling the IMF provides to the donor community that a government is sufficiently credible by committing to pursue sound economic policies under the tutelage of IMF monitoring and conditionality.

39. A similar event study tests the plausibility of this hypothesis by evaluating the evolution of official developmental assistance. In contrast to the analysis using GDP observations, countries that had an IMF arrangement in the prior period are removed. (The hypothesized mechanisms rely on the IMF beginning an arrangement with some country, so the omissions seem necessary in this case.) Likewise, it is not obvious that actual disbursements are necessary for the IMF to be catalytic, so the starts of Staff-Monitored Programs (SMPs), which involve no IMF financing, are included as well.

40. The results are plotted in Figure 8, where the normalization is in percentage terms rather than levels: 100 represents ODA flows at the approval of an IMF arrangement (with or without

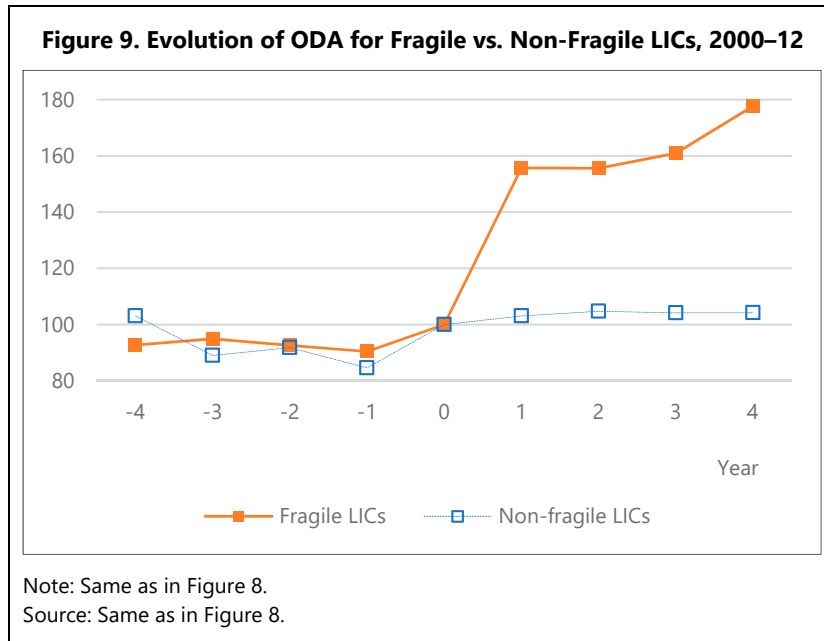
IMF financing), and 150 would represent 50 percent more ODA than at time zero.<sup>19</sup> The results from this exercise strongly support a “catalytic” effect. The confidence bands (expanded to 90 percent) do not come close to zero in the post-period, suggesting high levels of statistical significance in the results. The pre-period does not show any trend of improvement that could suggest that the observed 60 percent increase would have happened in the absence of an IMF arrangement.



41. Figure 9 compares the evolution of ODA inflows surrounding approval of an IMF arrangement in fragile versus non-fragile low-income countries. It is striking that a sharp pickup in ODA inflows observed for fragile states is hardly observable for non-fragile states.<sup>20</sup> In this case, a causal argument seems much more defensible, given the lagged jump.

<sup>19</sup> For ODA, simple averages are presented, rather than generalized least squares. Volatility in ODA essentially represents political choices. It seems less appropriate to interpret any correction by volatility in these numbers as simply an increase in statistical efficiency.

<sup>20</sup> Likewise, though not formally reported here, a similar strong catalytic role of an IMF arrangement was observed for fragile but not for non-fragile middle-income countries.



### Robustness of the event-study results

42. Before concluding this section, it is worth commenting on the robustness of the positive GDP growth and ODA results. For GDP growth, it turns out that using weighted averages rather than simple averages changes the results for the  $t + 4$  horizon (see the corresponding figures in Appendix 3). The same is true if we eliminate luminosity measures and use only the reported GDP growth from the Penn World Tables. In Figure 6, the  $t + 4$  coefficient is approximately 1, but changing either of these assumptions makes this number indistinguishable from zero. There are convincing theoretical reasons to think that one should weight and improve the data with luminosity information—which is why those results are presented and are assumed to be the “better” ones when the results conflict.

43. The results for ODA are more robust. Figure 2 makes clear that there is no upward trend of ODA flows to fragile states over the sample period, so there is no reason to believe that longer-term trends drive this result spuriously. Moreover, using the same sample that was used for the growth analysis (allowing for overlapping programs, and only considering lending arrangements), the results become weaker, but the overall pattern is the same. This is exactly what would be expected if only new arrangements matter for “catalyzing” aid.

## V. CONCLUSION

44. This paper has conducted a statistical analysis of the economic performance of fragile states, and of how performance relates to IMF lending or program engagement. Given the poor data quality and selection bias inherent in understanding how IMF engagement interacts with economic performance, much of the paper focused on utilizing improved methodologies for careful assessment.

45. The results suggest that compared with other states, fragile states grow at a somewhat slower pace, on average, and are more susceptible to growth volatility. Once an IMF arrangement has begun, however, fragile states appear to have significantly higher sustained economic growth rates. Moreover, following the start of IMF program engagement, with or without IMF financing, fragile states experience substantially larger inflows of foreign aid. Overall, these findings suggest that IMF involvement overall has had a positive impact on macroeconomic performance in fragile states.

## **APPENDIX 1. FRAGILE CLASSIFICATION**

While our methodology to classify fragile states is based on the approach taken by the World Bank, the list for any given year does not exactly match the harmonized list, for the following reasons.

First, the Asian Development Bank does not publish its own CPIA scores, although the World Bank publishes the ADB score for a country classified as fragile. This means that ADB's CPIA scores are available only for countries with sufficiently low CPIA scores, preventing us from creating a continuous series of CPIA scores for all low-income countries. Thus, for our analysis the World Bank's CPIA scores are averaged with the scores published by the African Development Bank (AfDB), and for countries not scored by AfDB, the World Bank's CPIA is the only input.

Second, we consistently use the CPIA scores currently available online irrespective of whether those were actually used in classifying countries as fragile on the harmonized list. This occasionally creates an inconsistency between our list and the harmonized list. For example, the World Bank's list for 2010 includes Cameroon, even though Cameroon's CPIA score for that year is 3.43 and the country had no peacekeeping or peacebuilding operation in place as far as we can verify. Our list would not include Cameroon, because the country does not meet either of the two criteria for inclusion as a fragile state.

Our methodology is to use the same criteria consistently to create a list of fragile states year by year, back to 2000, whereas the World Bank started publishing its harmonized list only from 2010. If the objective were only to look as far back as 2010, the harmonized list, as publicly available, could have been sufficient. It would be inconsistent, however, to utilize our approach until 2010 and then switch to the World Bank's published list thereafter. For a statistical analysis, it is important to use a consistent approach and avoid introducing systematic differences by using information selectively.

## APPENDIX 2. MATHEMATICAL DETAILS OF SATELLITE USE

The positive relationship between luminosity and GDP growth will be used in the following way.<sup>1</sup> First, start by assuming that there is a linear relationship between the percentage growth in true GDP growth,  $z$ , and the growth in how bright a country appears from space,  $l$ .

$$l_{i,t} = \tau_0 + \tau_1 z_{i,t} + \omega_{i,t}$$

Since true GDP growth is not known and the eventual goal is to feed luminosity growth in to predict GDP growth, an indirect estimate of  $\tau_1$  will be obtained by running the following inverted regression:

$$y_{i,t} = \beta_0 + \beta_L l_{i,t} + \mu_{i,t}$$

where  $y_{i,t}$  is observed GDP growth for country  $i$  in year  $t$ ,  $l_{i,t}$  is the percent growth in luminosity, and  $\mu_{i,t}$  is a regression error coming from measurement error in GDP growth as well as the imperfect fit this line has with true GDP growth,  $\omega$ .

This seemingly circular logic of using observed GDP estimates to derive “independent” estimates of GDP from lights is in fact grounded in econometric theory. Measurement error in the independent variable does not prevent the estimation of the best linear predictor. The measurement error in national income data will inform what weight to place on the data, but this paper will merely take a conservative estimate coming from the academic literature rather than re-estimating the measurement error. With these coefficients estimated, a new, independent series of noisy estimates can be obtained using only information coming from satellite data.

It is less straightforward to mathematically determine the optimal synthesis of these estimates. Let  $\rho$  be the weight placed on national income data. Then, the following equation is used to determine our new composite GDP growth estimate,  $\hat{z}$ :

$$\hat{z}_{i,t} = \rho y_{i,t} + (1 - \rho) \hat{y}_{i,t}$$

Here  $\hat{y}$  is the growth estimate coming from satellite data and the estimated  $\beta$ s. Letting  $z_{i,t}$  represent the true GDP growth of a given country in a given year, the objective is to find the  $\rho$  that minimizes the expected mean squared error of the composite from the true value.

$$\text{var}(\hat{z} - z) = \text{var}(\rho y_{i,t} + (1 - \rho) \hat{y}_{i,t} - z)$$

The equality here just results from substituting in the definition of our composite estimate. The resulting  $(1 - \rho)$ , optimal weight on satellite-data-derived growth, comes from Chen and Nordhaus (2011), and can be worked out with a few steps of calculus and algebra.

---

<sup>1</sup> The discussion presented here is a broad outline of the logic presented in Henderson and others (2012), who provide additional details including on mathematical assumptions.



$$(1 - \rho) = \frac{\tau_1^2 \sigma_\epsilon^2}{\tau_1^2 \sigma_\epsilon^2 + \sigma_\omega^2}$$

where  $\sigma_\omega^2, \sigma_\epsilon^2$  represent the variance in the errors for the true lights–growth linear relationship and the variance over measurement error in national income data, respectively.

It is important to note that the relevant error for lights is in the *true* relationship, not the observed one. The observed errors in the regression come from both measurement error and noise in the relationship. However, as this equation shows, if the only error came from errors in national income data (i.e.  $\sigma_\omega^2 = 0$ ), all weight should be put on satellite estimates ( $1 - \rho = 1$ ). If errors in national income data are extremely small ( $\sigma_\epsilon^2 \rightarrow 0$ ), there is no relationship between luminosity and GDP ( $\tau_1 = 0$ ) or if the lights–growth errors are very large ( $\sigma_\omega^2 \rightarrow \infty$ ), then  $1 - \rho \rightarrow 0$ . This equals zero only if it is believed that there is zero error in national income accounts, or there is zero correlation between observable lights and GDP. This is a strong conclusion: the correlation has been verified empirically, so unless national income data are perfect, precision can be increased by incorporating satellite data.<sup>2</sup>

In applying this methodology to actual data, there is one practical issue that has come up in the literature, namely, the non-trivial measurement error in luminosity due to weather or other events that cause variations in the satellite images. While not explicitly addressing this issue, Henderson and others (2012) nevertheless conclude that this relationship is not as strong at yearly frequencies. Kuruc (2017) shows that this weaker relationship is caused by the measurement error and can be corrected by properly accounting for it. In this connection, it should be mentioned that Chen and Nordhaus (2011) are more skeptical about the weight to place on satellite data, given the measurement error in luminosity.

For this paper we re-estimate the luminosity–GDP growth relationship at yearly frequencies, accounting for measurement error in satellite data, utilizing the technique introduced in Kuruc (2017), who demonstrates that the measurement error is substantial and provides a strategy for generating superior estimates of yearly luminosity growth for estimating yearly GDP growth. The key insight of this technique is that if measurement error is throwing substantial “noise” into the short-run data, an instrumental-variables technique should recover a coefficient that looks like the result when one uses the long-run data.<sup>3</sup> This is confirmed using an internally constructed instrument inspired by Griliches and Hausman (1986). It is then shown that supplementing

---

<sup>2</sup> In situations where national income data are trustworthy, lights will optimally be weighted close to zero; this in turn implies that the practical gains will be close to zero. So, while it is theoretically correct that some gain exists in every context, it is not practically relevant for all countries.

<sup>3</sup> This is a well-known property of instrumental variables. A regression with a noisy independent variable will bias downward the resulting coefficient. A second noisy measure of this independent variable can be used as an instrument to recover the “true” coefficient.

observed light growth by an average that includes some fraction of the observed cumulative growth rate over the period provides superior estimates.<sup>4</sup>

As there is substantial uncertainty in the literature surrounding these optimal combinations, the weights placed on the supplementing variables (outer growth rates in luminosity, followed by the weight on satellites itself) will be conservative. For nested growth rates, Kuruc (2017) suggests more than half should be placed on the outer difference optimally. This paper uses 55 percent. In the context of statistics produced by low-capacity agencies, Chen and Nordhaus (2011) suggest assigning around one-third weight to luminosity versus GDP, while Henderson and others (2012) suggest assigning about half; this paper uses 35 percent. It is important to note that any combination will produce consistent estimates of the true values of interest; using the “optimal” combinations is in hopes of improving accuracy in finite samples. Not believing that these numbers are truly the best combination does not imply that the results are in some way biased.

---

<sup>4</sup> While this technique introduces future and past growth information into the current estimate, this merely results in serial correlation in the errors. If it improves the information on period  $t$  growth (which it should), the results are improved, albeit non-independent, estimates of the path of average changes.

### APPENDIX 3. EVENT STUDY DETAILS

For the event studies, the following regression equation is used:

$$W_{i,t} = C_i + \sum_{j \in J} \alpha_j IMF_{t-j} + u_{i,t}$$

where  $W_{i,t}$  is an outcome that varies from its value at the time an IMF arrangement is approved,  $C_i$ , non-parametrically depending on how many periods the country is from the start of an arrangement. The set  $J$  runs from  $\{-4,4\}$  excluding 0.  $IMF_t$  is an indicator for whether an IMF arrangement began in time  $t$ . Finally,  $\alpha_j$  is the estimated coefficient that allows this equation to trace out the average path. Suppose, as an example,  $\alpha_1 = A$ . The prediction of this model is that one year following ( $t + 1$ ) the start of arrangement  $i$ , the outcome is  $W_{i,t+1} = C_i + A$ . When looking at one year following an arrangement, only  $IMF_{i,t-1}$  will be non-zero, so only  $\alpha_1$  survives.<sup>1</sup>

On a technical note, it is well-known that ordinary least squares (OLS) is inefficient when the error,  $u_{i,t}$ , has different variance across observations (i.e. heteroskedasticity). As mentioned when discussing use of satellite data (see Appendix 2), efficiency gains are extraordinarily important, given the data constraints for this problem. Therefore, feasible generalized least squares (FGLS) will be employed. This technique essentially posits that for observations with more “noise” thrown in (large variance for  $u_{i,t}$ ) less can be learned from their observed outcomes. For instance, if a country with very stable GDP growth had an increase following an IMF arrangement, statistically there would be more confidence that the growth increase was not driven by randomness than if the same increase were observed for a country with widely fluctuating GDP growth. To put this idea into practice, the observations will be weighted by the inverse of their variance of their yearly growth rates from 1992 to 2013 (the sample for which satellite data are available) to generate the  $\alpha$  coefficients.

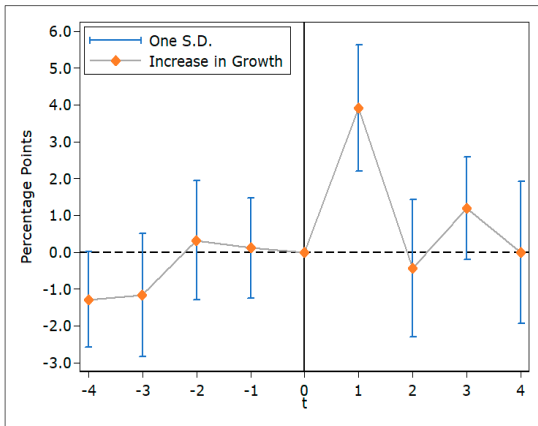
How robust are the patterns of GDP growth in Figure 6 in the text? There are two obvious ways to deviate from the reported results. First, ignoring the luminosity-implied growth and strictly using Penn World Table growth rates (small changes to the weights make very little difference). Second, not using the GLS technique and treating every observation with identical weight. The two variations are reported in Figures A.3.1a and A.3.1b; the results are not quite as strong, but have a similar pattern. Both modifications increase efficiency, so the results being “less strong” is not surprising or particularly concerning, though it would have been reassuring if the results held regardless of the removal of some statistical power. Both panels imply a weaker response four years out, as noted in the text. This is the most troubling disagreement and should give the reader some pause when interpreting the sustained growth implied by the main results.

---

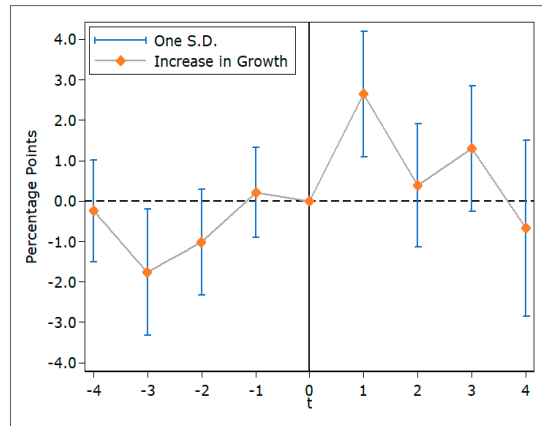
<sup>1</sup> Technically, some lending arrangements have windows that overlap. The example presented assumes that this is not the case.

**Figure A.3.1. Robustness of GDP Growth to Two Alternative Specifications, 2000–12**

(a) Growth obtained solely from Penn World Tables Data



(b) Growth obtained from simple averages of Penn World Tables and Luminosity-Implied Data



Source: Author's calculations based on: GDP from Penn World Tables and satellite imaging from NOAA; dates of financing agreements from IMF Finance Department.

## REFERENCES

- Barro, R., and J.-W. Lee, 2005, "IMF Programs: Who is Chosen and What are the Effects?" *Journal of Monetary Economics*, 52(7): 1245–1296.
- Chen, X., and W. Nordhaus, 2011 "Using Luminosity Data as a Proxy For Economic Statistics," *Proceedings of the National Academy of Sciences* 2(60): 8589–8594.
- Collier, P., and L. Chauvet, 2008, "What Are the Preconditions for Turnarounds in Failed States?" *Conflict Management and Peace Sciences* 25(4): 332–348.
- Deaton, A., 2011, "Instruments, Randomization, and Learning About Development," NBER Working Paper.
- Dicks-Mireaux, L., M. Mecagni, and S. Schadler, 2000, "Evaluating the effect of IMF lending on low-income countries," *Journal of Developmental Economics* 61(2): 495–526.
- Dreher, A., 2006, "IMF and economic growth: The effects of programs, loans, and compliance with conditionality," *World Development* 34(5): 769–788.
- Griliches, Z., and J. Hausman, 1986, "Errors in Variables in Panel Data," *Journal of Econometrics* 31: 93–118.
- Henderson, J. V., A. Storeygard, and D. Weil, 2012, "Measuring Economic Growth From Outer Space," *American Economic Review* 2(102): 994–1028.
- Jerven, M., 2013, *Poor Numbers: How We Are Misled By African Development Statistics and What to Do About It* (Ithaca, NY: Cornell University Press).
- Johnson, S., W. Larson, C. Papageorgiou, and A. Subraamanian, 2013, "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates," *Journal of Monetary Economics* 2(60): 255–274.
- Kuruc, K., 2017, "The Truth Is Out There: Measuring Yearly Growth From Outer Space," University of Texas: Working Paper.
- Prezowski, A., and J. R. Vreeland, 2000, "The effect of IMF programs on economic growth," *Journal of Developmental Economics* 62(2): 385–421.
- Pritchett, L., M. Woolcock, and M. Andrews, 2010. *Capability Traps? The Mechanisms of Persistent Implementation Failure* (Washington D.C.: Center for Global Development).