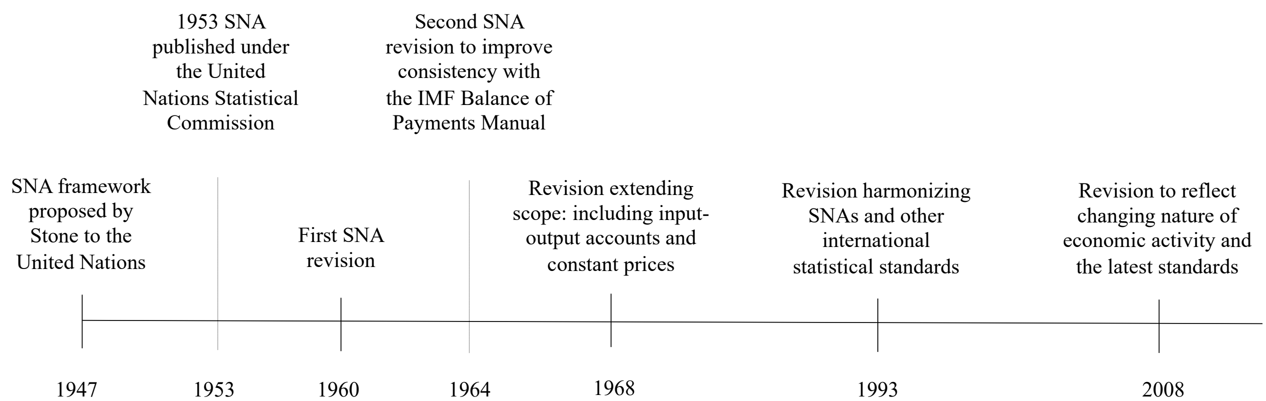
ONLINE APPENDIX

A. Additional Figures and Tables

Figure A1

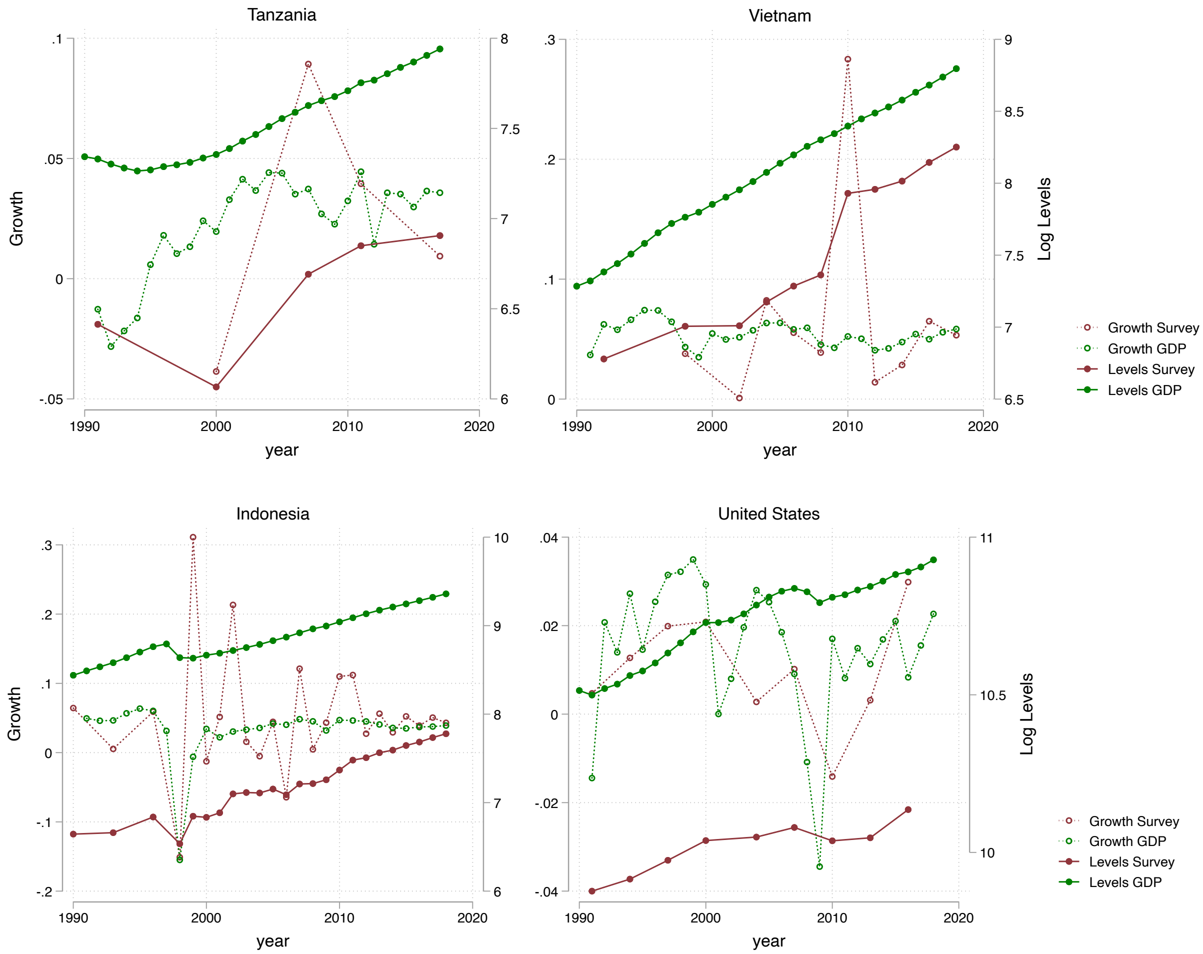
Timeline of System of National Accounts (SNA) Establishment and Revisions



Source: United Nations Statistics Division, unstats.un.org/unsd/nationalaccount/hsna.asp
Notes: The European System of National and Regional Accounts (ESA) included a revision in 2010 for Member States to Eurostat, which is consistent with the 2008 SNA.

Figure A2

Example Countries Growth and Income Levels in GDP and Surveys



Notes: Each graph maps year-on-year levels and growth rates for survey data as well as national accounts GDP data. We depict connected plots for each year where data is available. For growth, we calculate growth in terms of annualized log first differences.

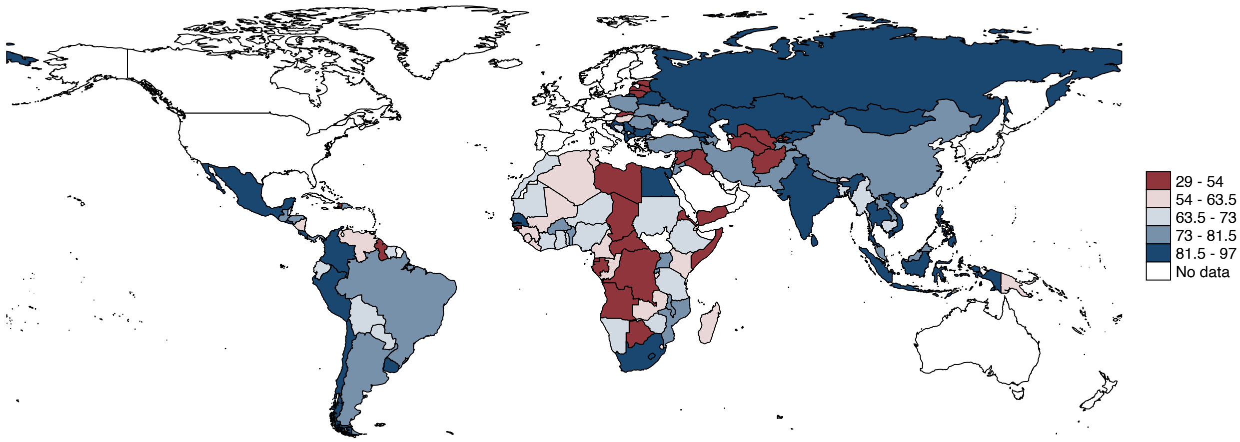
To get a sense of what growth measures look like for countries at different levels of national income, we look at some examples. We use the World Bank's country classification system to pick one country from each income category: low-income (Tanzania), lower-middle income (Vietnam), upper-middle income (Indonesia), and high-income (United States). In Figure A2, we show over-time measures of growth and income levels for each of these countries. The GDP measure is GDP per capita based on SNA, while the survey measure is per capita consumption or per capita income (depending on what is available) based on household surveys.

Three patterns emerge clearly. First, year-on-year growth, which receives enormous attention, is highly volatile in all four countries. Moreover, the correlation of year-on-year growth rates based on SNA versus survey data seems low, again for all four countries. Second, in contrast to the pattern documented for year-on-year growth, long-run growth patterns and trajectories emerge clearly. The long-run growth trajectory based on survey data is broadly consistent (though not identical) with the one emerging from SNA data, revealing the potential utility of alternative sources of data to adjudicate GDP estimates. Again, this pattern holds across all four examples, although to different degrees. Finally, the limited number of survey data points for Tanzania is striking.[1] This suggests that survey data availability may be a binding constraint in some developing countries, limiting the ability to adjudicate GDP estimates with this data source alone. This is an additional reason that we triangulate reliability of growth estimates with alternative sources of data such as satellite-based night-time lights.

[1] The panel in Figure 2 for the United States indicates that our survey-based estimates are available biennially. The official data used by the Government of the United States for estimating the prevalence of poverty comes from the Current Population Survey (March Supplement) and is available annually, but the data we use (drawn from PovcalNet) comes originally from the Luxembourg Income Study which creates an income vector that is harmonized across countries in their archives.

Figure A3

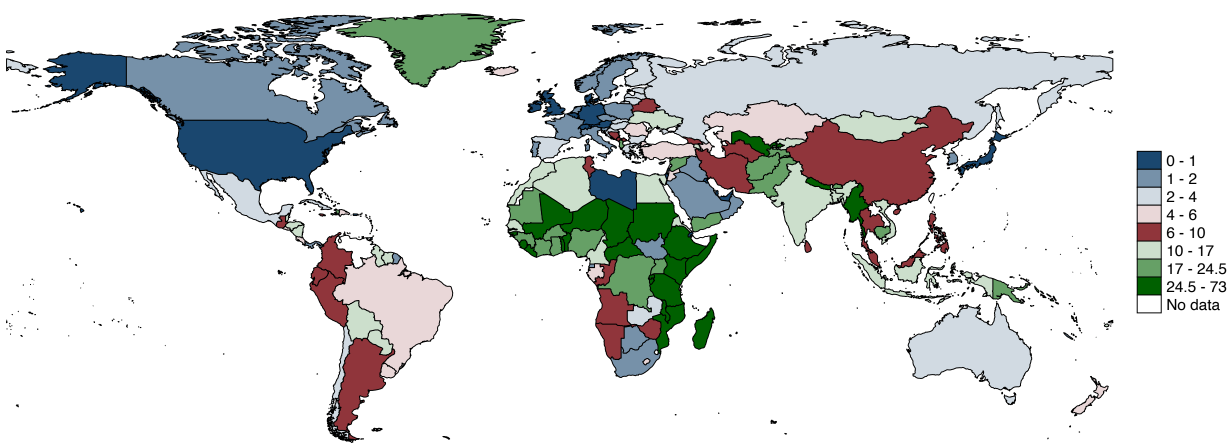
Statistical Capacity Index Latest Available Year



Source: World Bank Statistical Capacity Index. We use data from the latest year available.

Figure A4

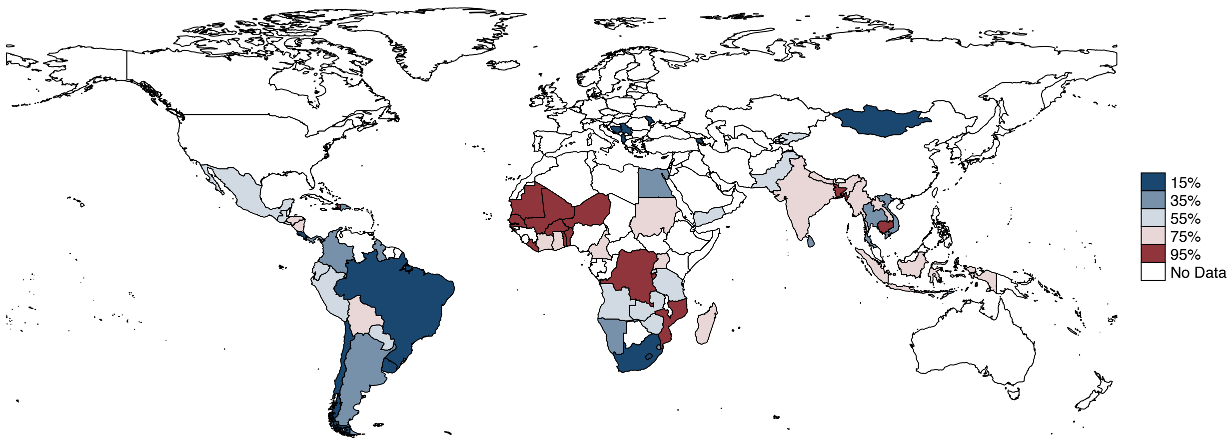
Share Agricultural Value-Added (% GDP) Latest Year



Source: World Bank and United Nations national accounts data.

Figure A5

Share Non-Ag Informal Employment from the ILO for the Latest Available Year

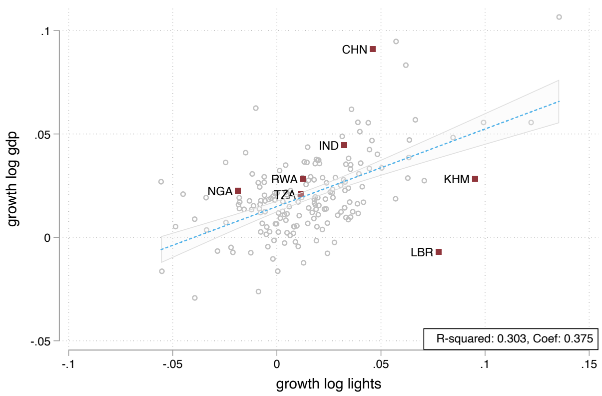


Source: Harmonized data series from the ILO (2018) drawing from direct household surveys for 69 countries.

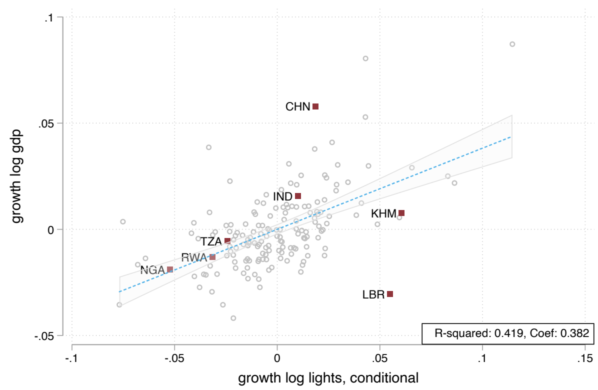
Figure A6

Growth in GDP vs Night-time Lights

A: Without Controls



B: Controlling for Measurement Challenges



Notes: Figure A6 includes average growth for 164 countries from 1992-2012. Growth in lights and GDP is the average annual growth rate between 1992 and 2012. Growth is calculated as first differences in logs.

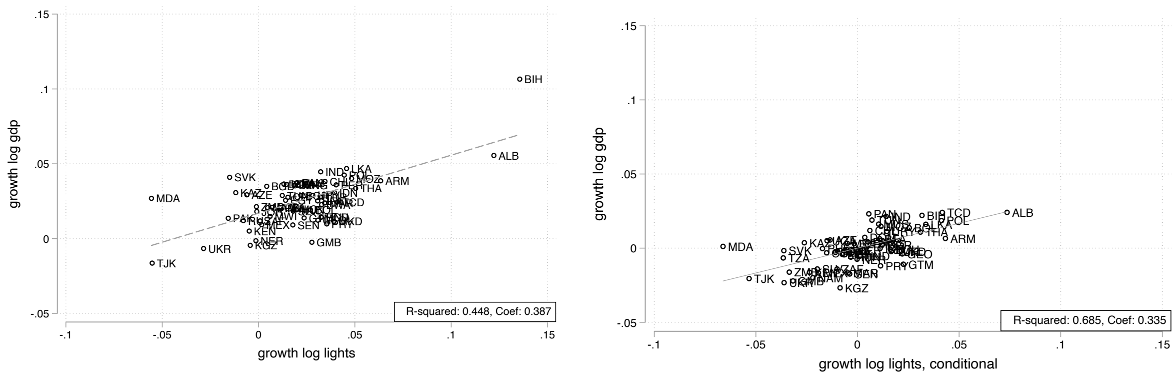
In Figure A6, we first examine whether specific countries buck the trend, and then repeat this exercise after conditioning on those factors we suspect may be responsible for growth mismeasurement. Specifically, we control for the indicators on statistical capacity from the World Bank and GDP compilation practices as well as price measurement practices data collected by the IMF, for agricultural value-added in national accounts. We examine whether after controlling for these factors, the divergence from the average elasticity is reduced and whether the R-squared of the associated regression is larger.

In Panel B, we control for the indicators on statistical capacity, price measurement practices, and agricultural value-added in national accounts across 164 countries. The inclusion of controls results in a tighter concentration around the fitted line as revealed by the substantially improved R-squared, which increases from .303 to .419. Nigeria and Rwanda, for example, converge fully to the line of best bit. While some outliers remain, such as China and Liberia, among others, the distance to the best-fit line has been reduced even for these outliers.

Figure A7

Growth in GDP vs Night-time Lights, with Quality, Capacity and Integrity Controls

A: No Controls *B: Conditional on Quality, Capacity and Integrity*



Notes: Figure A7 includes average growth for 60 countries from 1992-2012 for lights and GDP. We condition on quality, statistical capacity and integrity using the novel IMF data, in addition to the controls included in Figure 9 (i.e., price measurement practices, agricultural value-added in national accounts, and the share of GDP attributed to natural resources).

We use the novel data from the IMF based on the Reports on the Observance of Standards and Codes (see Table 2) to assess the role of not only quality and statistical capacity, but also integrity and transparency. The IMF audit data covers 83 countries total. The integrity indicator is missing for some countries, e.g., China. When we combine all controls from Figure A6 and the IMF data we end up with 60 countries in our sample.

Figure A7 shows results with no controls versus results with controls, and reveals a jump in R-squared from .448 to .685. Bosnia and Herzegovina sees a striking convergence to the average elasticity and we no longer observe any outlier countries. This suggests that in this case, the divergence from the average elasticity prior to conditioning was plausibly due to low-quality and/or manipulation of GDP data, and it virtually disappears when we control for those factors. Moreover, we no longer observe any outliers among the 60 countries for which we have the data to run this regression.

Additional Tables

Table A1

Over Time Within-country Correlations

	(1) Low income	(2) Lower middle income	(3) Upper middle income	(4) High income
Panel A: All Data				
Correlation: Survey and GDP Growth	0.16	0.23	0.19	0.33
	0.83	0.54	0.51	0.32
Correlation: Survey and Light Growth	-0.17	-0.04	0.27	-0.00
	0.82	0.57	0.44	0.38
Correlation: GDP and Light Growth	0.09	0.02	0.07	-0.02
	0.34	0.21	0.24	0.16
Panel B: Panel B: Only if High Data Availability (N > 3)				
Correlation: Survey and GDP Growth	0.33	0.33	0.27	0.33
	0.64	0.37	0.42	0.32
Correlation: Survey and Light Growth	-0.36	0.10	0.22	-0.00
	0.62	0.47	0.41	0.38
Correlation: GDP and Light Growth	-0.04	0.06	0.10	-0.02
	0.19	0.20	0.24	0.16

Notes: Correlations are computer using year-on-year data from 1992-2012 in pair-wise comparisons among data sources. These correlations are then averaged by income category. We also compute standard deviations *across* countries within a given income group of the *within-country* correlations.

In Table A1, we find that when we compare within-country correlations of survey and national accounts, we see higher correlations at higher income levels, varying from .16 to .33. However, if we restrict the sample to countries with survey data for more than 3 time periods in Panel B, this pattern virtually disappears. Hence it seems that the lower year-to-year correlations in low-income countries are driven by limited data. For example, Rwanda has only three survey data points. This means Rwanda has growth rates estimated at two points in time: in 2005 and 2010. Any correlation over-time is thus derived from the difference in growth from 2005 to 2010 in survey estimates relative to the difference in national accounts. As it turns out in Rwanda, based on two data points, the national accounts data is in close correspondence with survey data. But of course, it would be unwise to assume that this relationship always holds in year-to-year data. In some other developing countries, the correlation between growth rates estimated from household and survey data is -1. This reveals the importance of enhancing data availability for reliable measurement in developing countries.

Table A2

Statistical Capacity by Income Category and Region

	Statistical Capacity		
	Mean	SD	Observations
Low income	59.6	16.4	33
Lower middle income	69.3	13.7	47
Upper middle income	69.6	16.6	53
East Asia & Pacific	66.4	17.0	21
Europe & Central Asia	77.8	15.9	21
Latin America & Caribbean	71.1	12.9	23
Middle East & North Africa	59.9	18.0	13
South Asia	72.1	14.3	8
Sub-Saharan Africa	61.6	14.1	47

Notes: Data comes from the World Bank Statistical Capacity Index. We use data from the latest year available. SD is the cross-country standard deviation by income category, and observations is the total number of countries. We omit high-income countries from the income group disaggregation.

Table A3

Descriptive Characteristics of Systems of National Accounts

	Low income		Lower middle income		Upper middle income		High income	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Vintage of SNA	1996	6.20	1999	9.97	2000	7.72	2006	6.81
Benchmark Year	2005	4.97	2005	6.01	2006	5.27	2009	2.66
Availability of Quarterly GDP	0.38	0.48	0.59	0.49	0.77	0.42	0.91	0.28
Availability of Annual GDP	0.97	0.17	1.00	0.00	1.00	0.00	1.00	0.00
Annual GDP -Production constant	0.94	0.24	0.98	0.15	1.00	0.00	1.00	0.00
Annual GDP - Expenditure - constant	0.94	0.24	0.85	0.36	0.74	0.44	0.84	0.36
Annual GDP -Production - current	0.94	0.24	1.00	0.00	1.00	0.00	1.00	0.00
Annual GDP -Expenditure - current	0.97	0.17	0.89	0.31	0.91	0.29	0.93	0.25
Quarterly GDP - Production - costant	0.34	0.48	0.57	0.50	0.75	0.43	0.86	0.34
Quarterly GDP -Expenditure - constant	0.12	0.33	0.41	0.49	0.51	0.50	0.76	0.43
Quarterly GDP -Production - current	0.22	0.41	0.52	0.50	0.70	0.46	0.86	0.34
Quarterly GDP -Expenditure - current	0.12	0.33	0.43	0.50	0.51	0.50	0.81	0.39
Income Approach	0.44	0.50	0.48	0.50	0.66	0.47	0.83	0.38
Independent Compilation	0.12	0.33	0.30	0.46	0.40	0.49	0.76	0.43
Timely release of Quarterly GDP	0.16	0.36	0.52	0.50	0.53	0.50	0.81	0.39
Timely release of Annual GDP	0.62	0.48	0.61	0.49	0.75	0.43	0.84	0.36
Advance Release Calendars	0.28	0.45	0.54	0.50	0.57	0.50	0.76	0.43

Notes: The data presented here is compiled and structured from text responses to periodic IMF surveys conducted with 189 countries globally. We average statistics by income category.

Table A4

Agricultural Output: GDP and Satellite Data by Farm Type

	Growth Based on a Vegetation Index				
	(1)	(2)	(3)	(4)	(5)
	All	Smallholders	Large Farms	Smallholders	Large Farms
GDP Agriculture Value-Added Growth (%)	0.317***	-0.056	0.388***	-0.069	0.430***
	(0.084)	(0.150)	(0.083)	(0.158)	(0.087)
Observations	1,356	1,205	1,350	1,205	1,350
R-squared	0.010	0.000	0.016	0.043	0.071
Country and Year Fixed Effects	No	No	No	Yes	Yes

Notes: Data for the agriculture-vegetation index was produced from satellite imagery. The distinction between smallholder and larger corporate farms is based on definitions from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS). We run regressions using panel data from 2000 to 2018 across 87 countries in which the share of agricultural employment is above 25 percent. We also run regressions with fixed effects.

Table A5

Price Practices from 193 Economies

	High income	Upper middle income	Lower middle income	Low income
CPI Weights Reference Year	2013	2011	2010	2011
PPI Weights Reference Year	2011	2010	2010	2009
Inflation Targeting	0.65	0.43	0.39	0.12
CPI Availability	1.00	1.00	1.00	1.00
CPI Expenditure Coverage - National	0.87	0.75	0.70	0.62
CPI Expenditure Coverage - Urban Areas	0.11	0.15	0.20	0.12
CPI Expenditure Coverage - Capital City	0.02	0.09	0.11	0.25
CPI Frequency of Publication - Monthly	0.92	0.92	0.91	1.00
CPI Frequency of Publication - Quarterly	0.08	0.08	0.09	0.00
COICOP CPI Classification System	0.92	0.74	0.57	0.75
PPI Availability	0.79	0.68	0.61	0.41
PPI Timeliness of Publication - Monthly	0.63	0.53	0.39	0.12
PPI Timeliness of Publication - Quarterly	0.16	0.15	0.15	0.16
PPI Latest Industrial Classification	0.56	0.34	0.26	0.09
PPI Insutrial Coverage - MMU	0.29	0.36	0.20	0.16
PPI Insutrial Coverage - More than MMU	0.31	0.15	0.30	0.00

Notes: This table summarizes data compiled by the IMF across 193 economies by Berry et al. (2019).

B. New Data on National Accounts Quality, Capacity, and Integrity

We present a novel database of indicators based on expert audits of national accounts conducted periodically by the World Bank and the IMF: The Reports on the Observance of Standards and Codes (ROSCs) (https://www.worldbank.org/en/programs/rosc). These reports assess criteria of the IMF Data Quality Assessment Framework (DQAF) for 83 countries. Table B1 describes a subset of indicators within this framework, with 9 sub-indicators which focus on quality and integrity of national accounts data. Each indicator was assigned by IMF auditors to one of four categories: observed, largely observed, largely not observed, or not observed. We code these for our analysis as a dummy variable equal to one if the practice is observed or largely observed, and zero otherwise.

This new database enables us to identify additional quality measures beyond those included in Table A3 which focus on GDP compilation practices. This includes measures such as an indicator related to revision policy and practice, which are viewed by the IMF as being central to data quality. This indicator reflects whether revisions and updates of GDP estimates follow a regular and transparent schedule, and whether they are accompanied by explanatory notes and analysis. In addition, these data distinguish between measures of data quality as well as statistical capacity and integrity (i.e., the potential for politically motivated data manipulation). To capture capacity constraints, there is an indicator for whether National Statistical Offices have sufficient staff. The potential for political manipulation is captured with an indicator of whether there is internal governmental access to statistics prior to release and whether internal access is publicly identified. Table 2 in the main text provides summary statistics by income group.

Table B1

Indicators of the IMF Data Quality Assessment Framework (DQAF)

0.1 Legal and institutional environment	0.1.1 The responsibility for collecting, processing, and disseminating the statistics is clearly specified.
	0.1.4 Statistical reporting is ensured through legal mandate and/or measures to encourage response.
0.2 Resources	0.2.1 Staff, facilities, computing resources, and financing are commensurate with statistical programs.
0.4 Other quality management	0.4.2 Processes are in place to monitor quality during the planning and implementation of the statistical program.
1.1 Professionalism	1.1.2 Choice of data sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations
1.2 Transparency	1.2.2 Internal governmental access to statistics prior to their release is publicly identified.
4.3 Revision policy and practice	4.3.1 Revisions and/or updates follow a regular and transparent schedule
	4.3.3 Studies and analyses of revisions are made public (see also 3.5.1).
5.1 Data Accessibility	5.1.4 Statistics are made available to all users at the same time.

Notes: This table summarizes categories and indicators outlined in the IMF DQAF framework and is aligned to the United Nations Fundamental Principles of Official Statistics. IMF staff routinely conduct in depth audits with countries around the world including visits to National Statistics Offices (NSOs) and joint review of data sources and process documentation. We include examples for specific indicators for which we focus our analysis.

C. Data Description

Night-lights data

In our analysis, we use the night-time lights series from the Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS). This data source is cleaned to capture luminosity separate from the effects of cloud coverage, fires, aurora, and ephemeral light (Elvidge et al. 2009). Newer sources of night lights, such as the Visible Infrared Imaging Radiometer Suite (VIIRS), have also emerged; however, this data source is less well suited for growth analysis since it is less regularly cleaned and is accessible for only a few years. We do not aim to adjudicate the best night-time light data source in this paper; rather, we rely on the more often used data series from DMSP-OLS to triangulate national accounts data.

Vegetation Index

The vegetation index we use in our analysis, typically referred to as the Normalized Difference Vegetation Index (NDVI), is estimated by satellite detection of reflectance from plants in specific portions of the visible and infrared spectra. We use a data series for 89 economies where over 25 percent of employment is in agriculture from 2000 to 2018. We include measures for total NDVI per year per country as well as the maximum versus minimum NDVI in a given year and country. Based on definitions from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS), we also disaggregate the NDVI by smallholder farms which are often part of the informal economy versus large-scale commercial agricultural land, which usually is captured in national accounts.

References

Elvidge, Christopher D., Edward H. Erwin, Kimberly E. Baugh, Daniel Ziskin, Benjamin T. Tuttle, Tilottama Ghosh, and Paul C. Sutton. 2009. "Overview of DMSP nighttime lights and future possibilities." In *2009 Joint Urban Remote Sensing Event*, pp. 1-5. IEEE.