In Search of Information: Use of Google Trends' Data to Narrow Information Gaps for Low-income Developing Countries[☆]

Futoshi Narita[a,*], Rujun Yin[a]

[a]*International Monetary Fund, 700 19th Street, NW, Washington, DC 20431, United States*

Abstract

Timely data availability is a long-standing challenge in policy-making and analysis for low-income developing countries. This paper explores the use of Google Trends' data to narrow such information gaps and finds that online search frequencies about a country significantly correlate with macroeconomic variables (e.g., real GDP, inflation, capital flows), conditional on other covariates. The correlation with real GDP is stronger than that of nighttime lights, whereas the opposite is found for emerging market economies. The search frequencies also improve out-of-sample nowcasting performance albeit slightly, demonstrating their potential to facilitate timely assessments of economic conditions in low-income developing countries.

Keywords: Google search volume index, Low-income developing countries, Nighttime lights, Nowcasting, Economic growth, Capital flows
JEL: O11, O47, O57, E37, F17, F37

1. Introduction

Timely data availability in low-income developing countries (LIDCs) is a long-standing challenge to researchers and policy makers. LIDCs have more missing data and longer time lags in data release than more developed economies. For example, as of July 2018, official FDI data for 2017 are available only for less than half of LIDCs, compared to 90 percent for advanced economies.[1] A survey of the IMF staff indicates severer deficiencies in data quality and availability for low-income countries (Independent Evaluation Office, 2016, Figure 2). The lack of reliable and timely information hampers real-time assessment of economic conditions and restricts the ability to set sound policies.

Nontraditional data sources—so-called big data—have proven to be useful in providing operationally

[☆]Acknowledgment: We deeply thank Mamoon Saeed, Kei Moriya, and the Google Trends support team for their excellent technical support and Karina Chavez for her diligent editorial assistance. We also thank Sebastian Acevedo, Emre Alper, Claudia Berg, Alessandro Cantelmo, Rupa Duttagupta, Stefania Fabrizio, Rahul Giri, Wei Guo, Daniel Gurara, Roland Kpodar, Sandra Lizarazo Ruiz, Ali Mansoor, Marco Marini, Giovanni Melina, Machiko Narita, Neree Noumon, Chris Njuguna, Chris Papageorgiou, Saad Quayyum, Mahvash Qureshi, Alessandro Rebucci, Sidra Rehman, Sakina Shibuya, and the participants in the presentation at the IAAE 2019 Annual Conference and the seminars at the IMF for their thoughtful comments and suggestions. We are responsible for any remaining errors. This paper is part of a research project on macroeconomic policy in low-income countries supported by the U.K.'s Department for International Development (DFID, project ID: 60925, IATI Identifier: GB-1-202960). The views expressed in this paper are those of the authors and do not necessarily represent the views of the Internatinoal Monetary Fund (IMF), its Executive Board, IMF management, or the DFID.

[*]Corresponding author

Email addresses: fnarita@imf.org (Futoshi Narita), yinrujunjoy@gmail.com (Rujun Yin)

[1]The calculation is based on the International Financial Statistics database (IMF, 2018b). A fraction of missing values for FDI data since 2000 is 22 percent for LIDCs, compared to 12 percent for all the other non-LIDC economies. See Table B.1 for the country groupings. Note that the situation has been improving, because of country authorities' own efforts and international initiatives to address data gaps, including G-20's Data Gaps Initiative (https://www.imf.org/en/Publications/SPROLLs/g20-data-gaps-initiative).

valuable information in LIDCs.[2] Satellite imagery data, such as nighttime lights, are used to measure economic growth and poverty in countries and sub-regions where data are scarce (Henderson et al., 2012; Jean et al., 2016; Engstrom et al., 2017). In Kenya, researchers analyze mobile phone call records to help combat malaria more effectively (Wesolowski et al., 2012). A sensor technology generates usage statistics to improve performance of water pumps in Kenya and Ethiopia (Thomas et al., 2018, Table B.1).

This paper explores the potential of Google's search volume index (SVI)—a frequency of online search query submissions—to help narrow information gaps in LIDCs. Google's SVI would contain fruitful information about individuals' interests and attentions, considering the growing access to the Internet—especially, through mobile devices in developing countries—and Google's global user share of over 90 percent (StatCounter, 2018). People may search for information online to make economic decisions or to look for some developments in the economy. The SVI may capture these human behaviors in search of information, and that is the information potentially useful for economic analyses (see Appendix A.4, for discussion to formalize this idea). The information search could be more relevant for cross-border activities—travel, trade, foreign investment—that may face larger information barriers than local activities, and thus, it could be particularly useful for analyses on LIDCs, where such external economic activities play a key role (IMF, 2015a). Timely availability of Google's SVI is key in tackling information gaps due to the time lag in releasing official statistics in LIDCs.

To the best of our knowledge, this is the first study to apply Google's SVIs to a macroeconomic analysis on a comprehensive set of developing countries. The existing literature focuses on the use of Google's SVI for more developed countries than LIDCs (Table 1). Following (Choi and Varian, 2012, a working paper version was released in 2009), many researchers started to use Google's SVI to forecast or nowcast socioeconomic indicators.[3] The official statistical authorities and central banks have also adopted Google's SVIs and other big data for policy-making, data compilation, and economic research. But, the efforts are still largely concentrated in advanced or frontier emerging economies (IMF, 2018d, Box 3), likely reflecting the fact that online searches are more prevalent in these countries. We argue, however, that useful information on LIDCs may be able to be extracted from the online searches made in higher-income countries if these searches are *about* LIDCs. Our main analysis covers about 50 LIDCs (less than the total of 59 due to lack of macroeconomic data, while SVIs are available for all countries), and we also extend the analysis to about 90 other emerging and developing economies.

We find that Google's SVI can provide useful information to enhance real-time monitoring of economic conditions in LIDCs. We construct a panel data set of the SVI for each country by setting the country name as a search topic. And to be more granular, we further collect SVIs by category. For example, for Uganda, the SVI under the finance category increases if someone submits a query such as Uganda exchange rate, other things being equal. We choose five categories (finance; business and industrial; law and government; health; travel) and find in-sample significance of some of these SVIs in simple regression models of contemporaneous forecasting (i.e., nowcasting) that predict macroeconomic variables, conditional on lagged covariates. The use of these SVIs also improves out-of-sample performance, albeit slightly, measured by the mean of squared forecasting errors, computed by recursive forecasting regressions.

Using SVIs under various categories altogether seems to help disentangle positive and negative effects from the changes in individuals' attentions to a country. SVIs may signal confounded offsetting effects and it has been an issue in the application of the SVI (e.g., see Vozlyublennaia, 2014; page 18). In normal time, people may pay attention to a country if they are involved in some activities in the country, such as, searching for accommodations. This way, SVIs help identify positive effects on the country's economy. However, people may also pay attention because of natural disasters, conflicts, epidemics, scandals, etc. These events are

[2]In contrast to traditional data that are compiled for specific purposes, big data are collected as a byproduct of other activities (Hammer et al., 2017). The United Nations Economic Commission for Europe (UNECE) provides classification of big data (UNECE, 2013). The Week @ the Beach Index proposed by Laframboise et al. (2014) is an example of the use of nontraditional data sources in economic analysis.

[3]Active areas of research include finance (predicting stock price and volatility, following a seminal paper of an "attention index" by Da et al., 2011); health (including the famous Google Flu Trend by Ginsberg et al., 2009, and its refinement by Lampos et al., 2015); tourism (forecasting tourist arrivals); sociology (measuring issue salience); and political science (voting behaviors). IMF (2015c, Figure 2) uses SVIs to illustrate tourism demand to Samoa.

Table 1: Use of Google's SVI in forecasting/nowcasting economic variables

Author (publication year)	Country under analysis	Variable to predict
Götz and Knetsch (2019)	Germany	GDP
Ferrara and Simoni (2019)	Euro area	GDP
Chamberlin (2010)	United Kingdom	Retail sales
Carrière-Swallow and Labbé (2013)	Chile	Car sales
Barreira et al. (2013)	France, Italy, Portugal, Spain	Car sales
Askitas and Zimmermann (2009)	Germany	Unemployment rate
Fondeur and Karamé (2013)	France	Unemployment rate
Ross (2013)	United Kingdom	Unemployment rate
Reis et al. (2014)	France, Italy	Unemployment rate
Ferreira (2014)	Portugal	Unemployment rate
Chadwick and Sengül (2015)	Turkey	Unemployment rate
Vicente et al. (2015)	Spain	Unemployment rate
Smith (2016)	United Kingdom	Unemployment rate
D'Amuri and Marcucci (2017)	United States	Unemployment rate
Vosen and Schmidt (2011)	United States	Consumption
Wu and Brynjolfsson (2015)	United States	House price
Li et al. (2015b)	China	Consumer price index
Li et al. (2015a)	United States	Oil prices
Bangwayo-Skeete and Skeete (2015)	Caribbean countries	Tourist arrivals
Yang et al. (2015)	China	Tourist arrivals
Li et al. (2017)	China	Tourist arrivals
Artola et al. (2015)	Spain	Tourist arrivals
Siliverstovs and Wochner (2018)	Switzerland	Tourist arrivals
Rivera (2016)	Puerto Rico	Hotel registrations
Da et al. (2011)	United States	Stock prices/returns
Joseph et al. (2011)	United States	Stock prices/returns
Preis et al. (2013)	United States	Stock prices/returns
Vozlyublennaia (2014)	United States	Stock prices/returns
Takeda and Wakao (2014)	Japan	Stock prices/returns
Tantaopas et al. (2016)	Six AEs and four EMEs	Stock prices/returns
Adachi et al. (2017)	Japan	Stock prices/returns
Tang and Zhu (2017)	United States	Stock prices/returns
Welagedara et al. (2017)	United States	Stock prices/returns
Yung and Nafar (2017)	United States	Real estate investment trusts' (REITs) returns
Vlastakis and Markellos (2012)	United States	Stock market volatility
Smith (2012)	Eight AEs	Stock market volatility
Aouadi et al. (2013)	France	Stock market volatility
Hamid and Heiden (2015)	United States	Stock market volatility
Da et al. (2015)	United States	Stock market volatility
Dimpfl and Jank (2016)	United States	Stock market volatility
Moussa et al. (2017)	France	Stock market volatility
Goddard et al. (2015)	Five AEs	Exchange rate volatility
Peltomäki et al. (2018)	25 EMEs	Exchange rate volatility
Afkhami et al. (2017)	United States	Energy price volatility
Campos et al. (2017)	United States	Energy price volatility
Koop and Onorante (2013)	United States	Nine macroeconomic indicators

Source: Authors' survey.

Note: This list may not be exhaustive, and any omissions are purely incidental. See also Buono et al. (2017) for a broader survey on the use of nontraditional data in macroeconomic nowcasting. AEs: advanced economies; EMEs: emerging market economies.

rather associated with negative effects on the economy. Combining SVIs under different categories may help separate these offsetting effects, albeit not perfectly.[4] We generally find that the business-and-industrial and travel categories tend to be associated with positive effects, whereas the finance, law-and-government, and health categories tend to indicate negative effects.

The SVIs show stronger correlation with real GDP than that of nighttime lights for LIDCs, while the opposite is found for emerging market economies (EMEs). The significance of SVIs in the regressions for real GDP shows a stark contrast with the results for nighttime lights extracted from satellite imagery (Henderson et al., 2012), which lost significance once lagged covariates are included in the regressors. This is striking, because nighttime lights are well accepted as a proxy to economic activity in the development literature. For EMEs, however, nighttime lights significantly correlate with real GDP while SVIs are not as significant as in the case of LIDCs. This contrasting finding may indicate some structural differences between LIDCs and EMEs.

In addition to these new empirical findings, this paper also contributes to the literature by providing a foundation for interpreting the SVI. The paper formalizes the underlying conditions where Google's SVI could be associated with people's attention to the entities represented by a query (Appendix A.4). These conditions clarify what can be captured by the SVI and what kind of biases the SVI is subject to, filling the gap in the literature and providing a solid basis for the empirical research using SVIs in general.

The rest of the paper is structured as follows. Section 2 explains how we compile the data from the Google Trends service, while leaving technical details to Appendix A. Section 3 presents the main empirical results, including the comparison with nighttime lights in Section 3.3. Section 4 discusses several extensions, such as the results for EMEs in Section 4.4. Section 5 concludes with policy implications. Appendix B presents supplementary tables.

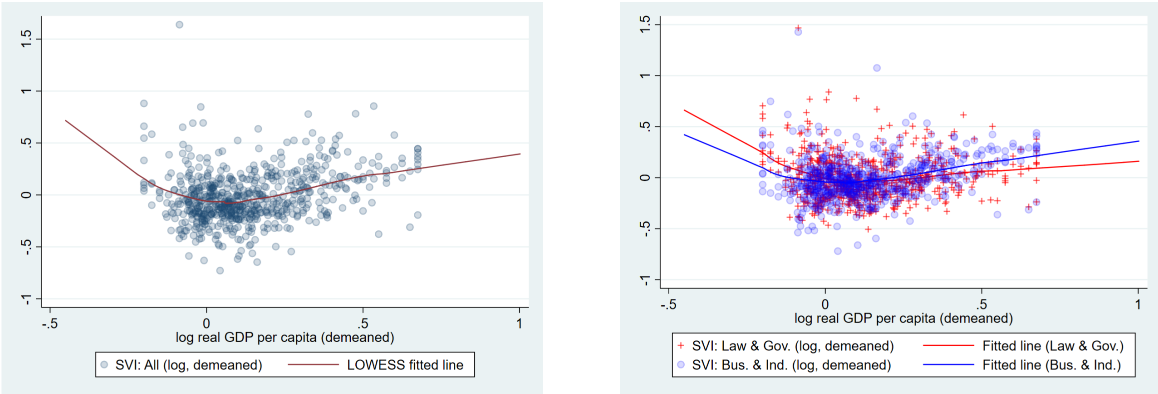
2. Search volume index for a country

The Google Trends service enables us to retrieve an SVI—a normalized measure of the search frequency—of a keyword or a topic. The SVI represents the number of search query submissions to the Google search engine on a keyword or a topic, relative to the total number of query submissions on all kinds of keywords. The SVI is further re-scaled on a range of 0 to 100 so that the resulting time series of an SVI shows 100 at its maximum. A search topic, rather than just a word, can be specified to resolve ambiguity due to homographs—e.g., word "Turkey" can mean a country or a bird (Stephens-Davidowitz and Varian, 2015)—by using Google's Knowledge Graph service. We can specify the locations where the queries were submitted and the categories under which the searches were made. See Appendix A for more details.

We use a country name as a search topic to obtain an SVI that proxies individuals' attention to a LIDC. The SVI based on a country name will increase if more search queries about the country are submitted to the Google servers than any other search queries. We argue that this SVI could reflect the number of people all over the world who get interested in something about the country (see Appendix A.4, for the conditions under which this claim would hold) and that we may be able to extract useful information about the country from the SVI. We use Google's Knowledge Graph service to resolve ambiguity of country names, including language issues (e.g., "Côte d'Ivoire" and "Ivory Coast") and adjust SVIs to make them comparable across countries (Appendix A.3). The SVIs constructed as such exhibit some positive correlation with the country income levels.

To separate positive and negative sentiments, we retrieve SVIs by category. A common issue with the SVI is the difficulty in labeling search terms with positive or negative sentiment and identifying how they are linked to economic indicators. Among the 25 major categories, we choose five categories—finance; business and industrial; law and government; health; and travel—to capture searches related to economic activities (finance; business and industrial; travel) and at the same time to control for searches related to negative

[4]The use of many SVIs under different categories is a standard approach in the literature. The most common approach is reducing the number of SVIs by extracting the principle components prior to regression analyses (Scott and Varian, 2015; Acevedo, 2016). In our case, we only use five SVIs and thus do not have to reduce the number of variables. In addition, keeping original SVIs helps us interpret the estimation results.



(a) U-shaped relationship between the SVI and real GDP per capita

(b) Different degrees of the responsiveness of SVIs to postive or negative sentiment.

Figure 1: SVI and real GDP per capita in LIDCs: within-country variations

Sources: Google Trends, World Development Indicators (World Bank, 2018), and the authors' calculations.

Note: The sample includes 53 LIDCs from 2004 to 2016, used in the main regression in Section 3.2. The SVI data are taken from Google Trends' private API. To preserve anonymity, demeaned real GDP per capita are winsorized at one percent and slightly modified to prevent to be isolated in the graph, while the fitted lines are drawn based on original data. All variables are transformed in natural logarithm and demeaned by country. Fitted lines are drawn using the locally weighted scatterplot smoothing (LOWESS; Cleveland, 1979). API: Application Programming Interface; LIDCs: low-income developing countries.

incidents that may adversely affect the economy (law and government; health). Note that SVIs under more granular subcategories (as shown in B.2) tend to return zeros due to lower search frequencies than Google's reporting threshold. It is an empirical question how successful this strategy is. In line with the concern about mixing positive and negative sentiments, within-variations of the overall SVI shows a U-shaped relationship with real GDP per capita (Figure 1a). Similar plots for the SVI under the business-and-industrial category and the SVI under the law-and-government category indicate that the former is more responsive to positive sentiment while the latter is more responsive to negative sentiment (Figure 1b).

Some cases illustrate underlying relationships between SVIs and economic activities. For example, the SVI for Myanmar under the travel category seems to capture the increasing trend of tourist arrivals to Myanmar since 2011 (Figure 2). From the beginning of 2011, Myanmar underwent a series of political reforms (IMF, 2015b). The following sections investigate whether this conjecture could be generalized, based on regression analyses.

3. Can Google's SVIs improve nowcasting performance for LIDCs?

3.1. *Nowcasting model*

To examine potential of Google's SVIs, we consider a simple nowcasting model using SVIs. We construct a panel data set of SVIs (the yearly averages of monthly data) from 2004 to 2017 for 59 LIDCs, combined with macroeconomic data taken from several databases (see Table B.3 for variable definitions and data sources; Table B.4 for summary statistics; and Table B.5 for pairwise correlation coefficients for selected variables). We postulate a simple linear regression as follows:

$$Y_{it} = \rho Y_{i,t-1} + \beta SVI_{it} + \gamma X_{i,t-1} + \alpha_i + D_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} denotes a variable to predict (real GDP growth, real exports, travel arrivals, inflation, exchange rates, private capital inflows, FDI inflows); SVI_{it} denotes a vector of SVIs under the selected five categories; X_{it} denotes a vector of other control variables; α_i and D_t are country fixed effects and time dummies, respectively; and ε_{it} denotes the residuals. See Table B.3 for how each variable is constructed and transformed (e.g., in natural logarithm or in percent change).

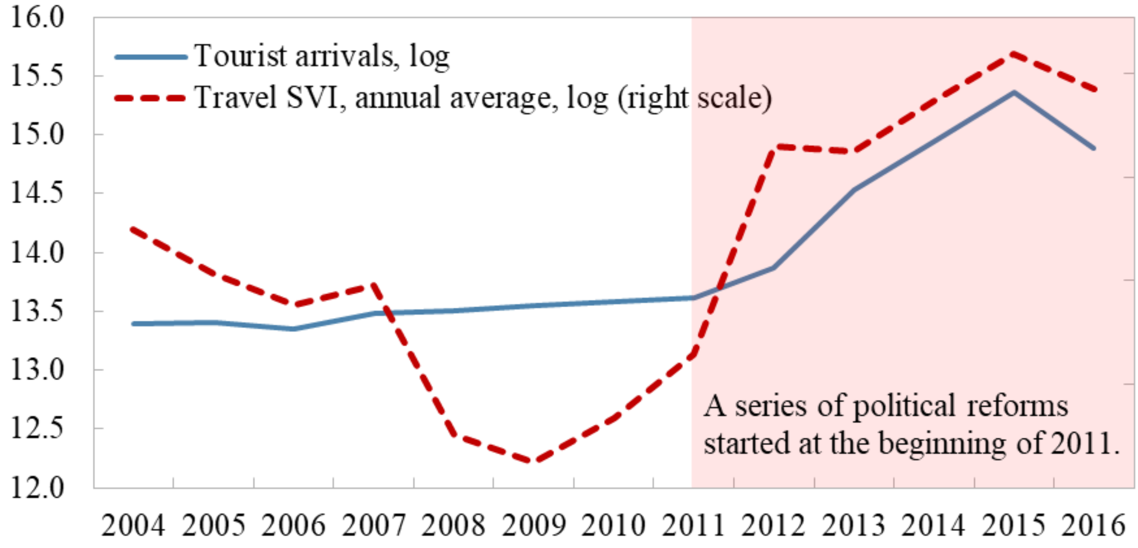


Figure 2: SVI under the travel category and tourist arrivals in Myanmar
Sources: Google Trends, World Development Indicators (World Bank, 2018), and the authors' calculations.
Note: The SVI data are taken from the Google Trends website (`https://trends.google.com/trends/explore?cat=67&date=all&q=%2Fm%2F04xn_`). SVI: search volume index.

This specification is motivated by real-time assessment of the economy when only lagged data are available. We put control variables X_{it} with a one-year lag, whereas the SVIs are contemporaneous, because our purpose is to explore the benefits from timely observation of SVIs in real-time monitoring of the economy where timely availability of macroeconomic statistics is an issue.[5] For example, we consider a situation to assess real GDP growth for the year 2016 as of January 2017 when actual real GDP and other related macroeconomic data for 2016 were not yet available, although SVIs for 2016 were available. Control variables X_{it} are chosen based on the empirical literature on variables to nowcast (e.g., for economic growth regression, Barro, 2015; for the determinants of capital flows, Araujo et al., 2017; Hashimoto and Wacker, 2016; Choi and Hashimoto, 2018), although many of the control variables that are used in the literature are not included due to lack of observations for many LIDCs. For example, including the gross enrollment ratio to secondary education reduces the sample size by one-third, while the results do not change significantly.

The purpose of the exercise is to find useful correlation between SVIs and economic variables, instead of establishing causality. We aim to predict Y_{it} by modeling the expected value of Y_{it} conditional on all the information available in a reduced form, instead of estimating structural causation between variables of interest. See Kleinberg et al., 2015 for a useful distinction between prediction and causation.

High correlation across SVIs by category—ranging from 0.77 to 0.92 (Table A.5)—would not be a matter of concern in predicting Y_{it}.[6] This is because the expected value of Y_{it} conditional on highly collinear variables should be similar to the one conditional on only one of the collinear variables, because highly collinear variables should carry very similar amounts of information about Y_{it}.[7] Since the number of SVIs that we use is five (or ten in extensions), there is little concern on a spurious perfect fit due to the large number of regressors. However, such high correlation would pose a challenge in separating the category SVIs into those that capture positive sentiments and those that capture negative sentiments.

Our model specification suffers from endogeneity issues. The issues stem from two-way causality between Y_{it} and SVI_{it} and from the inclusion of country fixed effects together with the lagged dependent variable

[5] Our purpose is not *backcasting*, which aims to assess real-time measurement errors based on ex post information. Although we use ex post data of SVIs, we consider this exercise as *nowcasting*, because the focus is still real-time economic monitoring.
[6] Within correlation (i.e., correlation among SVIs demeaned by country) is somewhat lower (0.40-0.80, in most cases).
[7] More generally, see Goldberger (1991, Chapter 23.3), echoed by (Hansen, 2019), for an argument that the issue of multicollinearity may be overemphasized.

(the so-called Nickell bias problem; e.g., see Barro, 2015). In general, endogeneity issues do affect the performance of prediction. For example, in the standard demand-supply estimation, an increase in quantity by itself is only a mixed signal on the price. The price would increase or decrease, depending on whether the quantity increase is driven by stronger demand or supply. The problem is, however, arguably much less severe in the case where two-way causality shares the same direction, i.e., an increase in one variable always implies an increase (or always implies a decrease) in the other variable, keeping other things constant. We argue that the relationships between SVIs and economic variables of our interest fall in this case. After all, how severe the endogeneity problem is can be assessed by the performance of prediction, although the estimated coefficients would be anyway biased.

3.2. In-sample regression results

We find some of the SVIs show significance in the simple nowcasting model, contributing to a better fit of the model. We confirm that these findings are robust to the issue of sampling, conducted in constructing SVIs (see Appendix A for details), by repeating the same exercise for five separate vintages of the SVIs constructed during April-June 2018. For ease of exposition, we refer to the SVI under a category in a concise way; for example, the SVI under the business-and-industrial category is referred to the business-industrial SVI, and so on. The findings are broadly similar when we do not include lagged covariates. While interpreting the estimated coefficients needs some caution because of the estimation issues mentioned in the previous section (i.e., endogeneity issues and the issue of multicollinearity among the SVIs), specific findings are as follows:

- Economic activities (Table 2). The business-industrial SVI exhibits a significant positive correlation with real GDP, indicating that a 10 percent increase in business-related attention would be associated with a 0.7 percent increase in real GDP. The law -government SVI and the health SVI, on the other hand, show significant negative correlations, implying that these SVIs may capture slowdowns in economic activities due to public concerns on legal, political, or health issues. These SVIs show a broadly similar pattern of correlation with real exports and tourist arrivals—with larger magnitudes—, in line with a conjecture that people's attention from outside of the country is the source of the observed correlations. The travel SVI is positively correlated with tourist arrivals. We also try tourism receipts, but the correlation is not as robust as for tourist arrivals, possibly because the SVI is more associated with the number of people interested in visiting the country, rather than how much they spend in the country.

- Prices (Table 3). There is strong positive correlation between inflation and the finance SVI—a 10 percent increase in finance-related attention would be associated with an increase in inflation by 0.3 percentage points. The results for the nominal exchange rate imply that the finance SVI may reflect currency depreciation pressures and that its pass-through to inflation may explain the results for inflation. Correlation between the finance SVI and the real effective exchange rate (REER) is not significant, possibly due to relatively high pass-through in LIDCs. The law-government SVI seems to be correlated with REER appreciation, which we admit is not so intuitive because the law-government SVI is negatively associated with economic activities (as is shown in Table 2). The travel SVI is significantly associated with lower prices, which would be due to people's travel interests to a destination with cheaper goods and services.

- Capital flows (Table 4). We find positive associations between gross capital inflows and the business-industrial SVI. Motivated by Araujo et al. (2017), we separately examine FDI and non-FDI flows and find somewhat stronger correlation for non-FDI flows. The finance SVI show no significant association, possibly because the SVI may be more associated with individuals' behaviors (e.g., checking the exchange rate) and personal investment to these countries is not yet significant. The behaviors of institutional investors may be better captured by the business-industrial SVI. The travel SVI is negatively correlated with capital flows, which may reflect lower financing needs due to higher travel service receipts.

The findings are broadly robust to model uncertainty (Table 5). We employ the Bayesian model averaging (BMA) methodology to examine robustness of our findings to specification uncertainty (Leamer, 1978). The estimation is implemented using Stata command **bma** (De Luca and Magnus, 2011). The results show that our findings are mostly robust to specification uncertainty, although the correlations with inflation and capital flows are not so strong as they appear in Tables 3 and 4.

3.3. Comparison with nighttime lights

Nighttime lights (NLs) extracted from processed satellite imagery can also serve as a nontraditional source of information for real-time economic monitoring, like SVIs. Since the seminal application by Henderson et al. (2012), NLs have gained popularity as a proxy to the degree of economic activity (for a recent survey on the economic applications of satellite data, see Donaldson and Storeygard, 2016). While Henderson et al. (2012) compile annual data based on the Defense Meteorological Satellite Program Operational Linescan System (DMSP OLS) data, a newer data set based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) is available monthly since April 2012 (until October 2018 as of November 18, 2018), although its annual data set—with additional data cleaning—is available only for 2015 and 2016.[8] We use the annual data compiled by the R package **Rnightlights**, developed by Njuguna (2018), while cross-checking them with the data compiled by Henderson et al. (2012). The correlation between the two NL data are almost one (Table B.5).

We benchmark SVIs with NLs and find that SVIs may contain stronger signals on economic activity than NLs in LIDCs, while we find the opposite for EMEs. The significance of SVIs broadly remains while NLs are not statistically significant for LIDCs (Table 6, columns 1-4). For EMEs, however, the opposite is found—NLs are significant in the case of the OLS while SVIs are not (Table 6, columns 5-6). The weaker performance of NLs for lower income countries is in line with the finding in the literature that the association between NLs and output tends to be weaker for lower output areas (Chen and Nordhaus, 2011, Fig. 2, p. 8590). Further investigation indicates that the significance of NLs is lost for LIDCs when regressors include the lag of covariates (Table B.6), whereas it is not lost for EMEs (Table B.7). The contrasting results imply that there are some interesting structural differences between LIDCs and EMEs. For example, SVIs may better capture external factors, which may be relatively more important in LIDCs, whereas NLs may better reflect the level of domestic economic activity, which may play a larger role in EMEs than in LIDCs. The comparison between LIDCs and EMEs is also discussed in Section 4.4.

The weak linkages between NLs and real GDP in LIDCs may need to be revisited to reflect methodological refinements proposed in the literature. For example, Zhang et al. (2016) point out an issue of inconsistencies in the temporal signal and propose a refined method to generate a globally consistent time series of NLs from the DMSP OLS data. Their method leads to higher correlations between NLs and GDP than those based on the standard method (Elvidge et al., 2009; Elvidge et al., 2014). Also, Hu and Yao (2019) uses a nonparametric estimation framework to find important nonlinearity in the relationship between NLs and real GDP. Debbich (2019) combines NLs and the data on radiant heat from gas flaring to strengthen the estimated relationship between NLs and real GDP for oil-producing countries such as Yemen.

3.4. Out-of-sample nowcasting

We also examine out-of-sample performance of nowcasting. We conduct recursive nowcasting using 2012 as the starting year and calculate the mean squared error (MSE) of prediction for 2013-2016.[9] Namely, we predict the value of the variable of interest for 2013 by feeding observations available in 2013 (i.e., SVIs for 2013 and other variables for 2012) using the model estimated by the observations up to 2012. We then repeat this to predict values for 2014, 2015, and 2016, incrementally using more data to estimate the model.

[8] See `https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html`. Both original NL data sources are compiled by the initiatives under the National Oceanic and Atmospheric Administration (see the note under Table 6).

[9] As our nowcasting models include country fixed effects and time dummies, we follow Calhoun (2014) to set the prediction period to be close to the square root of the entire sample period ($4 \simeq \sqrt{13}$). The results may depend on the choice of the starting year in general (Rossi and Inoue, 2012).

Table 2: Economic activities and the search volume index (SVI) in LIDCs

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables	Real GDP		Real exports		Tourist arrivals	
SVI: Finance		0.00		-0.00		0.01
		(0.01)		(0.04)		(0.08)
SVI: Business and industrial		0.07***		0.16*		0.25
		(0.02)		(0.09)		(0.19)
SVI: Law and government		-0.07***		-0.20***		-0.36***
		(0.02)		(0.07)		(0.11)
SVI: Health		-0.03**		-0.02		-0.23**
		(0.02)		(0.03)		(0.11)
SVI: Travel		0.00		0.02		0.23**
		(0.01)		(0.05)		(0.09)
Lagged dependent variable	0.85***	0.84***	0.83***	0.83***	0.68***	0.65***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
Population (lag)	-0.03	-0.01	0.15	0.29	-1.12*	-0.65
	(0.13)	(0.12)	(0.41)	(0.44)	(0.56)	(0.47)
Internet users (lag)	-0.00	-0.00	0.02	0.02	0.02	0.04
	(0.01)	(0.00)	(0.02)	(0.02)	(0.03)	(0.03)
Real GDP (lag)			-0.28*	-0.31*	-0.37	-0.48**
			(0.15)	(0.16)	(0.25)	(0.23)
Trade openness (lag)	0.02	0.01	-0.02	-0.05	-0.24**	-0.29***
	(0.01)	(0.01)	(0.09)	(0.08)	(0.11)	(0.10)
Fiscal spending (lag)	0.04***	0.03***	0.06	0.04	0.24***	0.22***
	(0.01)	(0.01)	(0.05)	(0.05)	(0.07)	(0.06)
REER, log level (lag)	-0.03	-0.03	0.10	0.09	-0.36***	-0.36***
	(0.03)	(0.03)	(0.09)	(0.09)	(0.13)	(0.11)
Inflation (lag)	-0.00*	-0.00*	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Trading partners' growth (lag)	0.00	-0.00	-0.00	-0.00	0.03**	0.03**
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Export price growth (lag)	0.00	0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Capital account openness (lag)	0.01	-0.01	0.14	0.09	0.30***	0.18
	(0.02)	(0.02)	(0.09)	(0.09)	(0.10)	(0.12)
Age dependency ratio (lag)	0.00	-0.00	0.00	0.00	0.01	0.01
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
Observations	644	644	633	633	575	575
Number of countries	53	53	53	53	52	52
Adjusted R-squared	0.961	0.964	0.797	0.802	0.743	0.763
Country fixed effects	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SVI: search volume index.

Table 3: Price developments and the search volume index (SVI) in LIDCs

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables	Inflation (percent change)		Nominal exchange rate (local currencies to one U.S. dollar, percent change)		REER (percent change)	
SVI: Finance		3.36***		7.83***		-2.49
		(1.04)		(2.32)		(1.68)
SVI: Business and industrial		-2.93*		-0.57		-0.90
		(1.47)		(3.48)		(2.73)
SVI: Law and government		0.48		-5.72**		4.68**
		(1.20)		(2.22)		(2.21)
SVI: Health		1.53		0.05		0.19
		(0.99)		(1.58)		(1.38)
SVI: Travel		-2.90**		-2.99		0.69
		(1.25)		(2.52)		(1.89)
Lagged dependent variable	0.34***	0.32***	0.13***	0.11***	0.04	0.04
	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)
Population (lag)	0.25	0.09	-6.19	1.65	14.24	10.67
	(14.79)	(15.96)	(16.98)	(18.78)	(9.45)	(11.28)
Internet users (lag)	0.96**	0.82*	0.82	0.52	-0.36	-0.39
	(0.42)	(0.43)	(0.81)	(0.79)	(0.69)	(0.67)
Real GDP (lag)	0.75	1.49	2.59	3.40	-1.43	-1.43
	(3.15)	(3.16)	(4.77)	(5.00)	(5.81)	(5.45)
Trade openness (lag)	-0.10	0.45	-6.89***	-6.37**	7.23***	7.55***
	(1.08)	(1.08)	(2.27)	(2.45)	(2.03)	(2.22)
Fiscal spending (lag)	-1.58	-1.83	-3.62*	-5.04**	2.22	2.90
	(1.12)	(1.14)	(1.90)	(1.95)	(2.19)	(2.16)
REER, percent change (lag)	-0.20***	-0.20***				
	(0.03)	(0.03)				
Inflation (lag)			-0.09**	-0.10**	0.24**	0.25**
			(0.04)	(0.04)	(0.11)	(0.12)
Trading partners' growth (lag)	-0.07	-0.06	-0.36	-0.35	-0.51	-0.52
	(0.21)	(0.20)	(0.32)	(0.31)	(0.33)	(0.34)
Import price growth (lag)	0.06	0.08	-0.11	-0.12	0.17	0.20*
	(0.06)	(0.06)	(0.09)	(0.10)	(0.11)	(0.11)
Capital account openness (lag)	-2.98	-1.84	-4.69	-4.57	0.45	1.06
	(3.10)	(2.70)	(4.15)	(3.52)	(3.15)	(3.06)
Age dependency ratio (lag)	-0.02	-0.02	0.31*	0.30*	-0.29***	-0.27**
	(0.13)	(0.13)	(0.17)	(0.17)	(0.10)	(0.10)
Observations	642	642	671	671	641	641
Number of countries	54	54	55	55	54	54
Adjusted R-squared	0.306	0.319	0.304	0.326	0.153	0.158
Country fixed effects	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SVI: search volume index.

Table 4: Capital flows and the search volume index (SVI) in LIDCs

Dependent variables	(1) Total capital inflows	(2)	(3) Private capital inflows	(4)	(5) FDI inflows	(6)	(7) Non-FDI inflows	(8)
SVI: Finance		-0.23		-0.17		-0.12		-0.32
		(0.23)		(0.20)		(0.27)		(0.29)
SVI: Business and industrial		1.19***		0.94***		1.24**		1.49***
		(0.30)		(0.35)		(0.53)		(0.50)
SVI: Law and government		-0.22		0.06		-0.32		-0.05
		(0.32)		(0.26)		(0.32)		(0.29)
SVI: Health		-0.47**		-0.36		-0.39*		-0.19
		(0.21)		(0.25)		(0.22)		(0.21)
SVI: Travel		-0.28		-0.58***		-0.21		-0.64**
		(0.18)		(0.19)		(0.24)		(0.28)
Lagged dependent variable	0.16**	0.14*	0.08	0.06	0.23***	0.21***	0.07	0.02
	(0.08)	(0.08)	(0.10)	(0.09)	(0.06)	(0.06)	(0.08)	(0.08)
Population (lag)	0.01	-1.00	-0.62	-2.22*	0.22	-0.01	0.78	-0.45
	(1.34)	(1.39)	(1.30)	(1.19)	(1.23)	(1.30)	(1.52)	(1.62)
Internet users (lag)	0.15*	0.10	0.19*	0.15	0.13	0.09	0.15	0.06
	(0.08)	(0.09)	(0.10)	(0.10)	(0.13)	(0.13)	(0.14)	(0.13)
Real GDP (lag)	0.47	0.56	1.04*	1.14*	0.96	0.94	-0.06	0.18
	(0.63)	(0.62)	(0.61)	(0.59)	(0.90)	(0.86)	(0.93)	(0.97)
Trade openness (lag)	0.55*	0.51*	0.68**	0.67**	0.55*	0.53*	1.27***	1.33***
	(0.29)	(0.26)	(0.29)	(0.26)	(0.29)	(0.30)	(0.40)	(0.36)
Fiscal spending (lag)	0.26	0.26	0.30	0.33	0.11	0.07	-0.25	-0.31
	(0.24)	(0.26)	(0.23)	(0.22)	(0.26)	(0.27)	(0.35)	(0.37)
REER, log level (lag)	-0.15	-0.14	-0.13	-0.12	-0.05	0.06	0.04	0.16
	(0.40)	(0.36)	(0.41)	(0.38)	(0.39)	(0.36)	(0.69)	(0.68)
Inflation (lag)	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Trading partners' growth (lag)	0.05	0.05	0.05	0.05	0.04	0.04	-0.00	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Export price growth (lag)	0.01	0.01	-0.01	-0.00	0.01	0.01	-0.03**	-0.02
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Capital account openness (lag)	0.11	0.02	0.45	0.55	-0.39	-0.54	1.73***	1.93***
	(0.52)	(0.59)	(0.42)	(0.44)	(0.45)	(0.44)	(0.48)	(0.45)
Age dependency ratio (lag)	0.00	0.01	0.00	0.00	-0.02	-0.02	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Observations	461	461	454	454	535	535	377	377
Number of countries	49	49	49	49	49	49	48	48
Adjusted R-squared	0.424	0.437	0.419	0.433	0.339	0.348	0.390	0.414
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SVI: search volume index.

Table 5: Bayesian model averaging results for LIDCs

Dependent variables	(1) Real GDP	(2) Real exports	(3) Tourist arrivals	(4) Inflation	(5) Nominal exchange rate	(6) Private capital inflows	(7) FDI inflows
SVI: Finance	0.00	0.00	0.01	**1.72**	**6.20**	0.00	0.01
	[0.04]	[0.06]	[0.09]	[0.67]	[0.96]	[0.06]	[0.07]
SVI: Business and industrial	**0.07**	**0.11**	**0.19**	-0.12	-0.39	**0.40**	**0.36**
	[1.00]	[0.76]	[0.65]	[0.08]	[0.10]	[0.56]	[0.52]
SVI: Law and government	**-0.07**	**-0.17**	**-0.34**	0.04	**-3.89**	-0.00	-0.05
	[1.00]	[0.97]	[1.00]	[0.05]	[0.68]	[0.06]	[0.12]
SVI: Health	**-0.03**	0.00	**-0.19**	0.05	-0.14	-0.16	-0.08
	[0.94]	[0.05]	[0.81]	[0.06]	[0.07]	[0.33]	[0.20]
SVI: Travel	-0.00	0.00	**0.24**	**-1.49**	-1.63	-0.14	-0.01
	[0.04]	[0.06]	[0.97]	[0.53]	[0.38]	[0.32]	[0.06]
Lagged dependent variable	**0.83**	**0.81**	**0.65**	**0.34**	**0.10**	0.03	**0.26**
	[1.00]	[1.00]	[1.00]	[1.00]	[0.83]	[0.27]	[1.00]
Population (lag)	0.00	**0.01**	-0.05	-0.03	0.31	-0.04	0.01
	[0.04]	[0.06]	[0.10]	[0.04]	[0.05]	[0.06]	[0.04]
Internet users (lag)	-0.00	0.00	0.00	0.24	0.03	**0.11**	0.05
	[0.04]	[0.08]	[0.08]	[0.30]	[0.05]	[0.55]	[0.30]
Real GDP (lag)		-0.07	**-0.21**	-0.01	0.07	**1.38**	0.56
		[0.38]	[0.53]	[0.04]	[0.05]	[0.80]	[0.41]
Trade openness (lag)	0.00	-0.00	**-0.23**	-0.00	**-7.12**	**0.76**	0.21
	[0.10]	[0.05]	[0.90]	[0.04]	[0.98]	[0.95]	[0.37]
Fiscal spending (lag)	**0.03**	0.00	**0.16**	-0.17	**-2.66**	0.13	0.03
	[0.98]	[0.05]	[0.82]	[0.13]	[0.70]	[0.33]	[0.11]
REER, log level (lag)	-0.00	0.02	**-0.30**			-0.00	0.00
	[0.11]	[0.15]	[0.88]			[0.05]	[0.04]
REER, percent change (lag)				**-0.19**			
				[1.00]			
Inflation (lag)	-0.00	-0.00	-0.00		**-0.05**	-0.00	-0.00
	[0.09]	[0.06]	[0.04]		[0.53]	[0.08]	[0.05]
Trading partners' growth (lag)	0.00	-0.00	0.01	-0.00	-0.02	0.01	0.00
	[0.05]	[0.06]	[0.24]	[0.04]	[0.07]	[0.11]	[0.09]
Export price growth (lag)	-0.00	0.00	-0.00			-0.00	0.00
	[0.36]	[0.07]	[0.04]			[0.05]	[0.07]
Import price growth (lag)				0.01	-0.02		
				[0.08]	[0.11]		
Capital account openness (lag)	-0.00	0.01	0.01	-0.18	-0.44	0.04	-0.02
	[0.04]	[0.09]	[0.06]	[0.07]	[0.09]	[0.09]	[0.05]
Age dependency ratio (lag)	-0.00	0.00	0.00	- 0.00	**0.25**	-0.00	-0.00
	[0.04]	[0.05]	[0.08]	[0.04]	[0.70]	[0.05]	[0.07]
Observations	644	633	575	642	671	454	535
Number of countries	54	54	53	54	55	49	49
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Posterior inclusion probability (PIP) are reported in brackets. The coefficients are bolded if PIP exceeds 0.5, corresponding to what is known as the median probability model (Barbieri and Berger, 2004). The estimation is implemented using Stata command **bma** (De Luca and Magnus, 2011). See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SVI: search volume index.

Table 6: Search volume index (SVI) and nighttime lights (NLs)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables	Real GDP					
	LIDCs				EMEs	
	OLS	BMA	OLS	BMA	OLS	BMA
SVI: Finance	0.01	0.01	0.01	0.00	0.00	0.00
	(0.02)	[0.25]	(0.01)	[0.06]	(0.01)	[0.04]
SVI: Business and industrial	0.02	0.00	0.06***	**0.05**	-0.01	-0.00
	(0.03)	[0.14]	(0.02)	[1.00]	(0.01)	[0.10]
SVI: Law and government	-0.06***	**-0.07**	-0.08***	**-0.08**	-0.00	-0.00
	(0.02)	[0.99]	(0.02)	[1.00]	(0.01)	[0.06]
SVI: Health	-0.01	0.00	-0.01	-0.00	-0.02	-0.01
	(0.01)	[0.06]	(0.02)	[0.07]	(0.01)	[0.19]
SVI: Travel	0.02	0.00	0.00	-0.00	0.01	0.00
	(0.02)	[0.19]	(0.02)	[0.05]	(0.01)	[0.05]
NLs from HSW (2012)	0.01	0.00				
	(0.02)	[0.06]				
NLs from HSW (2012) (lag)	-0.01	-0.00				
	(0.01)	[0.06]				
NLs from *Rnightlights*			0.01	0.00	0.02**	0.00
			(0.01)	[0.08]	(0.01)	[0.11]
NLs from *Rnightlights* (lag)			-0.01	-0.00	-0.02**	-0.00
			(0.01)	[0.04]	(0.01)	[0.06]
Control variables included	YES	YES	YES	YES	YES	YES
Observations	241	241	545	545	711	711
Sample period	2004-2008	2004-2008	2004-2013, 2015-2016	2004-2013, 2015-2016	2004-2013, 2015-2016	2004-2013, 2015-2016
Number of countries	53	53	53	53	70	70
Adjusted R-squared	0.937	-	0.969	-	0.961	-
Country fixed effects	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES
Excluding periods of jumps	NO	NO	NO	NO	NO	NO

Sources: Chinn and Ito (2006); Earth Observation Group; Financial Flows Analytics (IMF, 2018a); GADM (2018); Google Trends; Henderson et al. (2012); International Financial Statistics (IMF, 2018b); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' estimation.

Note. For ordinary least squares (OLS), cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. For Bayesian model averaging (BMA), posterior inclusion probability (PIP) are reported in brackets. The coefficients are bolded if PIP exceeds 0.5, corresponding to what is known as the median probability model (Barbieri and Berger, 2004). The estimation is implemented using Stata command **bma** (De Luca and Magnus, 2011). The NLs from HSW (2012) line shows the coefficients on NL data (variable *lndn*) compiled by Henderson et al. (2012), available for 1992-2008. The NLs from Rnightlights line shows the coefficients on NL data compiled by R package **Rnightlights** developed by Njuguna (2018), available for 1992-2013 based on DMSP OLS data (also used by Henderson et al., 2012) and for 2015-2016 based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) data. The DMSP OLS data are based on the processed images provided by National Geophysical Data Center, while images are collected by U.S. Air Force Weather Agency. The VIIRS DNB data are produced by the Earth Observation Group, NOAA/NCEI. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. Among EMEs, the NL data exclude countries identified as outliers by Henderson et al. (2012, footnote 16, p. 1011; Bahrain, Equatorial Guinea, Serbia, Montenegro). For the data compiled by **Rnightlights**, several large economies are also excluded due to their heavy computational burden (Brazil, Chile, China, Indonesia, India, Mexico, Peru, Russia). DMSP OLS: Defense Meteorological Satellite Program Operational Linescan System; EMEs: emerging market economies; LIDCs: low-income developing countries; NCEI: National Centers for Environmental Information; NOAA: National Oceanic and Atmospheric Administration; REER: real effective exchange rate; SVI: search volume index.

We compare the best predicting models selected from the pool of variables with and without SVIs. As including irrelevant variables to a model may increase the MSE, we conduct an exhaustive search from the pool of SVIs and control variables to identify the set of variables with which the linear regression model minimizes the MSE, combined with country fixed effects and time dummies. We then do this again only for control variables, without SVIs, and compare the MSEs between the two best predicting models.[10]

This way, we find that adding SVIs to the pool of variables improves performance in nowcasting economic indicators. We find that for all economic indicators to predict, the MSE of the best model is lower when including SVIs in the pool of selection, in the case of LIDCs (Table 7, Panel A). The differences in MSEs between the best models with and without SVIs are not very large in general nor statistically significant. Note, however, that most of our comparisons are between nested models and the standard statistical inference based on the Diebold-Mariano test (Diebold and Mariano, 1995) across nested models may not be valid, especially in the presence of autocorrelation or cross-panel dependency (e.g., see Diebold, 2015, for the review of the literature). The SVIs included in the best model are generally in line with the in-sample analysis, but not always the same. For example, for real GDP, while the law-government SVI is always selected in the top 10 models in terms of the MSE, as is significant in the in-sample results, the business-industrial SVI is not selected, but instead, the finance SVI is selected (Table B.8). Further investigation would be interesting to reconcile in-sample and out-of-sample results, as is discussed in the literature (e.g., Inoue and Kilian, 2005; Diebold, 2015, and associated comment papers).

4. Extensions

4.1. Jumps in SVIs

We observe jumps (or positive outliers) in SVIs occasionally. These acute increases in the SVIs are associated with critical events, including natural disasters, major policy changes, and key developments in the business environment. We identify 178 jumps in the SVI for the "all" category (i.e., with no category specified) out of 804 observations in our sample for LIDCs, using a methodology in the finance literature (Lee and Mykland, 2008). The difference between the squared percent change and the consecutive absolute percent change (called bi-power variations) indicates a huge change in the SVI within a period (see Appendix A.5 for details). The reason for not using each SVI by category for the jump detection is to focus on very acute increases in individuals' attention that are significant enough to stand out in the SVI with no category specified, even though their causes would be category-specific.

Excluding the periods when a jump occurred seems to sharpen estimation results. As each jump would have a very different implication from one another, we exclude those periods with jumps from the sample and re-estimate our models. The results show more statistical significance in many cases, while there is no significant change for inflation and the significance rather weakens for real exports and FDI inflows (Table B.10). This implies that jumps in SVIs could indicate the periods when their relationships with economic variables become unstable or strongly nonlinear, and that excluding such periods either strengthens the true linear relationships or weakens the spurious significance.

4.2. Lagged effects

There could be time lags for people's attentions to materialize as actual economic actions. Search of background information would happen before travel or investment take place. In this regard, SVIs could rather serve as a leading indicator.

In our specifications, lagged SVIs do not show significant correlation as clearly as contemporaneous SVIs do (Table B.11). This is probably because our models are at the annual frequency and the one-year lag could be too long. An exception is the case of private capital flows where lagged SVIs work better. For real GDP, lagged SVIs seem to complement contemporaneous SVIs. More meaningful leading signals could possibly be found in the SVIs at a higher frequency such as monthly, although limited availability of other indicators at a higher frequency would pose a challenge for such an analysis.

[10] We also compare the averages of the lowest 10 MSEs, instead of only the lowest MSE, and find very similar results.

Table 7: Out-of-sample performance of nowcasting

	Real GDP	Real exports	Tourist arrivals	Inflation	Nominal exchange rate	Private capital inflows	FDI inflows
	Panel A. MSE of the best model with fixed and time effects — LIDCs						
Controls only	0.37	1.60	7.58	0.12	0.83	60.84	122.40
Controls + SVIs	0.36	1.59	6.89	0.11	0.73	55.96	117.68
Difference (in percent)	**-2.6**	**-1.0**	**-10.0**	**-7.4**	**-14.4*****	**-8.7*****	**-4.0**
	Panel B. MSE of the best model with fixed and time effects — EMEs						
Controls only	0.09	0.77	1.92	0.45	1.00	75.45	47.90
Controls + SVIs	0.08	0.76	1.92	0.44	0.95	75.22	47.90
Difference (in percent)	**-3.1**	**-1.4**	**0.0**	**-2.0**	**-5.5*****	**-0.3**	**0.0**

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. We conduct recursive nowcasting using a panel data set from 2004 to 2016. We set 2012 as the starting year and calculate the mean squared error (MSE) of prediction for 2013-2016. We predict the value of the variable of interest for 2013, by feeding observations available in 2013 (i.e., SVIs for 2013 and other controls for 2012) using the model estimated by the observations up to 2012. We then repeat this to predict values for 2014, 2015, and 2016, incrementally using more data to estimate the model. We include country fixed effects and time dummies, from which we back out the averaged constant term so that country fixed effects and time effects are redefined as deviations from the constant term, and thus, ex ante time effects for prediction years can be assumed to be zero. Panel A shows the results for LIDCs and Panel B shows the results for EMEs. The Control variables + SVIs lines show the minimum MSEs identified by an exhaustive search from the pool of all variables to be included in the model. The Controls only lines show the minimum MSEs identified by an exhaustive search from the pool of control variables, excluding the SVIs. See Tables B.8 and B.9 for the best model specifications chosen in this procedure. To overcome a computational challenge stemming from the exhaustive search across variables to include, we follow the algorithm proposed by Somaini and Wolak (2016) to speed up the calculation to estimate regressions with two-way fixed effects. The Difference (in percent) lines show the differences of the above two lines in percent of the second line. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively, based on a Diebold-Mariano test (Diebold and Mariano, 1995) using cluster-robust standard errors, although it should be noted that most of these model comparisons are between nested models and conducting statistical inference across nested models is not trivial, especially when forecasting errors could exhibit autocorrelation or cross-panel dependency (e.g., see Diebold, 2015, for a review of the literature). The nominal exchange rate is the local currency per U.S. dollar, transformed to annual percent changes, period average. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. For inflation and nominal exchange rate, we divide them by 100 to be comparable to other logged variables for this table. EMEs: emerging market economies; LIDCs: low-income developing countries; SVI: search volume index.

4.3. Searches made domestically

We further examine SVIs on the searches made domestically. We construct an additional data set of SVIs by changing the location from "worldwide" to each country of interest—e.g., searches about Bangladesh made in Bangladesh. We refer to these SVIs as domestic SVIs. The domestic SVIs would capture individuals' attention to a country in that country. The domestic SVIs are more likely to be subject to the issue of low responses and the reporting cut-off, but they would potentially capture certain activities (especially those that happened locally) better than the worldwide SVIs.

Including domestic SVIs do not generally change the regression results, implying that the major source of information from worldwide SVIs is attention from foreign locations. Results do not change for most of the cases, except capital flows, which now show weaker correlation (Appendix Table 12). The domestic business-industrial SVI is negatively associated with inflation, which may reflect the importance of inflation for local businesses.

4.4. Does it work for EMEs too?

We also investigate whether Google's SVIs would be useful for macroeconomic analyses in EMEs. Our work can naturally extend to EMEs, many of which share common characteristics with LIDCs (see Table B.1 for the list of the EMEs and Table B.13 for summary statistics for EMEs).

For EMEs, the correlations between SVIs and macroeconomic variables are not as robust as those for LIDCs (Table B.14). As discussed in Section 3.3, the weaker correlations might imply relatively weaker influences of the external factors to EMEs than LIDCs—due to larger domestic markets in EMEs—because SVIs may better capture external factors related to online searches from abroad. Another reason could be that investors' behaviors to gain information about EMEs through the Internet may not be significant signals among other key factors in more matured and complicated financial markets in EMEs than those in LIDCs. Similarly, adding SVIs does not improve the nowcasting accuracy for EMEs as much as it does for LIDCs (Table 7, Panel B; Table B.9).

4.5. Machine learning algorithms

We apply some of machine learning algorithms, which are increasingly becoming popular in economic analyses (Varian, 2014; Mullainathan and Spiess, 2017). As these algorithms are mainly used to produce out-of-sample projections, we compare out-of-sample performances with our best linear models in Section 3.4. We examine LASSO-type regressions (implemented using **lassopack** by Ahrens et al., 2018) and Random Forest (implemented using the scikit-learn package in Python by Pedregosa et al., 2011).

The results are broadly similar to our findings in Section 3.4. LASSO-type regressions select a similar set of SVIs in the final projection models. For these methods, out-of-sample MSEs are larger than our best linear models, because their final projection models are still linear and our best linear models achieve the minimum MSEs by construction. As for Random Forest (implemented using the module **ensemble.ExtraTreesRegressor** of the scikit-learn package), we find slight improvements in MSEs compared to our best linear models and the MSEs are slightly smaller for the models that include SVIs than those do not. A more comprehensive analysis would potentially find further refinements of our results so far.

5. Conclusion

This paper presents an effort to use advanced technology to address the recurrent issue of lack of information in policy-making and analysis for developing economies. While progress has been made in timely provision of official data, nontraditional data obtained through recent technology have enormous potential to fill information gaps in developing economies. We investigate how much information we could obtain from Internet search frequencies to strengthen the capacity to monitor and assess current economic developments.

Our findings provide important steps forward in utilizing Google Trends' data in economic analyses. The in-sample and out-of-sample performances of a simple nowcasting model demonstrates the usefulness of the information contained in Google's SVI. The contrasting results between LIDCs and EMEs regarding the comparison of SVIs and another new source of information—nighttime lights—not only demonstrate the stronger case for the use of SVIs for LIDCs but also suggest the need to further investigate any structural differences between these country groups.[11] The estimated regression models indicate whether positive or negative effects are to be expected for each SVI and provide quantitative implications from the changes in SVIs. The results also indicate that jumps or outliers in SVIs may need to be separately treated because the estimated linear relationships are likely to break on these occasions. Monitoring SVIs can complement the use of judgment required in making forecasts, particularly for low-income countries where statistical models are generally less reliable than advanced economies due to data availability (Independent Evaluation Office, 2014, paragraph 34, p. 13).

There are several lessons learned about the use of Google Trends' data in economic analyses. First, monitoring SVIs under several categories is recommended to separate positive and negative signals. Second, interpreting jumps in SVIs warrants caution as they may likely indicate a departure from the normal relationships. What causes a jump can be identified by typing the country name and the period when a jump occurred into an online search engine. Lastly, a more granular analysis using specific search terms would be attractive but indeed challenging. This is not only because such an analysis would highly depend on the choice of terms (e.g., see a discussion by Smith, 2016, cited by Harchaoui and Janssen, 2018), but also because using more than one search term often leads to very low frequencies and sometimes falls below a threshold to be cut off, resulting in a zero response. For this reason, we use Google's Knowledge Graph service to identify a topic rather than a term and keep our topic as broad as a country, while achieving granularity by using SVIs under various categories. These practical solutions, however, rely on nontransparent methodologies and could undermine the credibility of the analyses.

There is still more to be explored to fully realize the potential benefits of using Google's SVIs. Our results at the annual frequency makes the case for more practical analyses on the use of Google's SVIs in constructing high frequency indicators of economic activities, as SVIs are available monthly (or even weekly for past 5 years, via the web service). In practice, nowcasting models may need to be tailored to the country of applications for more accuracy. Taking care of jumps in SVIs would be more important in such analyses, as these jumps can be noise or may serve as forewarning for a surge or a decline in the economy. Lastly, more flexible methodologies to analyze data, such as machine learning algorithms discussed by Varian (2014) and Mullainathan and Spiess (2017), could help extract more useful information from the SVIs. We have examined some machine learning algorithms in Section 4.5 and found broadly similar results so far, but further exploration in this direction is interesting.

The use of SVIs to cross-check the validity of official statistics would be interesting, but we need to be cautious. As is the case for nighttime lights (Henderson et al., 2012), the SVIs may possibly be used to cross-check the validity of official statistics, particularly in the context of a large share of the informal economy in LIDCs and other developing economies. If official statistics (e.g., real GDP) appeared much lower than the levels implied by observed SVIs, then it might indicate that a sizable portion of economic activities might not be captured by official statistics. This is the same logic behind the sociological literature on measuring issue salience (e.g., Stephens-Davidowitz, 2017). We need to be cautious, however, because a deviation between SVIs and official statistics would not necessarily be a proof of inaccuracy in the official statistics. Other factors include noises in SVIs themselves, unannounced changes in measurement of SVIs, and structural breaks in the relationships between SVIs and economic activities. Reis et al. (2014, section 6) list the challenges in using Google's SVIs in compiling official statistics, including transparency, auditability, consistency in measurement over time, and continuity of the Google Trends service in the future.

Further research is also needed for a more systematic use of Google's SVIs in policy decision making. Although the frequency of online search per se should be as objective as transaction data—unlike qualitative indicators based on subjective judgments—, it is still influenced by uncertainties stemming from the natural language processing algorithm used to compile category-specific SVIs (whose details are not disclosed to the public) and from Google's Knowledge Graph service that may not perfectly distinguish topics very close to each other (e.g., the Republic of the Congo versus the Democratic Republic of the Congo). The objectivity of Google's SVIs can be examined by comparing them with survey data (Vosen and Schmidt,

[11] See also the last paragraph of Section 3.3 for qualifications on methodological issues in estimating the linkage between nighttime lights and real GDP.

2011). In addition, while we provide certain conditions in Appendix A.4 where the SVI could represent people's attention without bias, the SVI may send a biased signal if these conditions do not hold. Lastly, as is known as Campbell's law (Campbell, 1979), a predominant use of Google's SVIs in policy decision making could provide undesirable incentives to manipulate frequencies of particular search terms—manually or automatically using Internet bots, distorting the useful relationships between SVIs and macroeconomic data. Addressing these concerns and caveats is left for future research.

References

Acevedo, S., 2016. Nowcasting GDP with Google Trends. Proof of Concept for the IMF's Big Data Challenge, Unpublished.

Adachi, Y., Masuda, M., Takeda, F., 2017. Google search intensity and its relationship to the returns and liquidity of Japanese startup stocks. Pacific-Basin Finance Journal 46, 243–257.

Afkhami, M., Cormack, L., Ghoddusi, H., 2017. Google search keywords that best predict energy price volatility. Energy Economics 67, 17–27.

Ahrens, A., Hansen, C.B., Schaffer, M.E., 2018. LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation. Statistical Software Components, Boston College Department of Economics. URL: `https://ideas.repec.org/c/boc/bocode/s458458.html`.

Aouadi, A., Arouri, M., Teulon, F., 2013. Investor attention and stock market activity: Evidence from France. Economic Modelling 35, 674–681.

Araujo, J.D., David, A.C., van Hombeeck, C., Papageorgiou, C., 2017. Joining the club? Procyclicality of private capital inflows in lower income developing economies. Journal of International Money and Finance 70, 157–182. doi:`10.1016/j.jimonfin.2016.08.006`.

Artola, C., Pinto, F., de Pedraza García, P., 2015. Can internet searches forecast tourism inflows? International Journal of Manpower 36, 103–116. doi:`10.1108/IJM-12-2014-0259`.

Askitas, N., Zimmermann, K.F., 2009. Google econometrics and unemployment forecasting. Applied Economics Quarterly 55, 107–120.

Bangwayo-Skeete, P.F., Skeete, R.W., 2015. Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. Tourism Management 46, 454–464.

Barbieri, M.M., Berger, J.O., 2004. Optimal predictive model selection. The annals of statistics 32, 870–897. doi:`10.1214/009053604000000238`.

Barreira, N., Godinho, P., Melo, P., 2013. Nowcasting unemployment rate and new car sales in south-western Europe with Google Trends. NETNOMICS: Economic Research and Electronic Networking 14, 129–165. doi:`10.1007/s11066-013-9082-8`.

Barro, R.J., 2015. Convergence and modernisation. The Economic Journal 125, 911–942. doi:`10.1111/ecoj.12247`.

Buono, D., Mazzi, G.L., Kapetanios, G., Marcellino, M., Papailias, F., 2017. Big data types for macroeconomic nowcasting. Eurostat Review on National Accounts and Macroeconomic Indicators 1/2017, 93–145.

Calhoun, G., 2014. Out-of-sample comparisons of overfit models. Economics Working Papers, No. 11002. Iowa State University.

Campbell, D.T., 1979. Assessing the impact of planned social change. Evaluation and program planning 2, 67–90.

Campos, I., Cortazar, G., Reyes, T., 2017. Modeling and predicting oil VIX: Internet search volume versus traditional variables. Energy Economics 66, 194–204.

Carrière-Swallow, Y., Labbé, F., 2013. Nowcasting with google trends in an emerging market. Journal of Forecasting 32, 289–298.

Chadwick, M.G., Sengül, G., 2015. Nowcasting the Unemployment Rate in Turkey: Let's Ask Google. Central Bank Review 15, 15.

Chamberlin, G., 2010. Googling the present. Economic & Labour Market Review 4, 59–95.

Chen, X., Nordhaus, W.D., 2011. Using luminosity data as a proxy for economic statistics. Proceedings of the National Academy of Sciences 108, 8589–8594.

Chinn, M.D., Ito, H., 2006. What matters for financial development? Capital controls, institutions, and interactions. Journal of development economics 81, 163–192. URL: `https://web.pdx.edu/~ito/Chinn-Ito_website.htm`.

Choi, H., Varian, H., 2012. Predicting the present with Google Trends. Economic Record 88, 2–9.

Choi, S., Hashimoto, Y., 2018. Does transparency pay? Evidence from IMF data transparency policy reforms and emerging market sovereign bond spreads. Journal of International Money and Finance 88, 171–190.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74, 829–836.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. The Journal of Finance 66, 1461–1499.

Da, Z., Engelberg, J., Gao, P., 2015. The sum of all FEARS investor sentiment and asset prices. The Review of Financial Studies 28, 1–32. doi:`10.1093/rfs/hhu072`.

D'Amuri, F., Marcucci, J., 2017. The predictive power of Google searches in forecasting US unemployment. International Journal of Forecasting 33, 801–816.

De Luca, G., Magnus, J.R., 2011. Bayesian model averaging and weighted average least squares: equivariance, stability, and numerical issues. Stata Journal 11, 18–544.

Debbich, M., 2019. Assessing oil and non-oil gdp growth from space: An application to yemen 2012-17. IMF Working Paper 19/221.

Diebold, F.X., 2015. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. Journal of Business & Economic Statistics 33, 1–1. doi:10.1080/07350015.2014.983236.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & economic statistics 13, 253–263. doi:10.1080/07350015.1995.10524599.

Dimpfl, T., Jank, S., 2016. Can internet search queries help to predict stock market volatility? European Financial Management 22, 171–192. doi:10.1111/eufm.12058.

Donaldson, D., Storeygard, A., 2016. The view from above: Applications of satellite data in economics. Journal of Economic Perspectives 30, 171–98.

Elvidge, C., Ziskin, D., Baugh, K., Tuttle, B., Ghosh, T., Pack, D., Erwin, E., Zhizhin, M., 2009. A fifteen year record of global natural gas flaring derived from satellite data. Energies 2, 595–622.

Elvidge, C.D., Hsu, F.C., Baugh, K.E., Ghosh, T., 2014. National trends in satellite-observed lighting. Global urban monitoring and assessment through earth observation 23, 97–118.

Engstrom, R., Hersh, J., Newhouse, D., 2017. Poverty from space: using high-resolution satellite imagery for estimating economic well-being. Policy Research Working Paper.

Ferrara, L., Simoni, A., 2019. When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage. Banque de France Working Paper, No.717.

Ferreira, P., 2014. Improving prediction of unemployment statistics with Google Trends: part 2. Eurostat Working Paper.

Fondeur, Y., Karamé, F., 2013. Can Google data help predict French youth unemployment? Economic Modelling 30, 117–125. doi:10.1016/j.econmod.2012.07.017.

GADM, 2018. Database of global administrative areas. https://gadm.org/, accessed in September 2018.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L., 2009. Detecting influenza epidemics using search engine query data. Nature 457, 1012–1014.

Goddard, J., Kita, A., Wang, Q., 2015. Investor attention and FX market volatility. Journal of International Financial Markets, Institutions and Money 38, 79–96.

Goldberger, A.S., 1991. A course in econometrics harvard university press. Cambridge, MA .

Götz, T.B., Knetsch, T.A., 2019. Google data in bridge equation models for German GDP. International Journal of Forecasting 35, 45–66.

Hamid, A., Heiden, M., 2015. Forecasting volatility with empirical similarity and Google Trends. Journal of Economic Behavior & Organization 117, 62–81.

Hammer, C., Kostroch, D.C., Quirós, G., STA Internal Group, 2017. Big data: Potential, challenges and statistical implications. IMF Staff Discussion Note.

Hansen, B.E., 2019. Econometrics. https://www.ssc.wisc.edu/~bhansen/econometrics/, accessed in October 2019.

Harchaoui, T.M., Janssen, R.V., 2018. How can big data enhance the timeliness of official statistics?: The case of the us consumer price index. International Journal of Forecasting 34, 225–234.

Hashimoto, Y., Wacker, K., 2016. The role of information for international capital flows: new evidence from the SDDS. Review of World Economics 152, 529–557. doi:10.1007/s10290-016-0250-4.

Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. American Economic Review 102, 994–1028. doi:10.1257/aer.102.2.994.

Hu, Y., Yao, J., 2019. Illuminating economic growth. IMF Working Paper 19/77.

Independent Evaluation Office, 2014. IMF forecasts: process, quality, and country perspectives. URL: http://www.ieo-imf.org/ieo/pages/EvaluationImages261.aspx. Evaluation Report.

Independent Evaluation Office, 2016. Behind the scenes with data at the IMF : an IEO evaluation. URL: http://www.ieo-imf.org/ieo/pages/EvaluationImages261.aspx. evaluation Report.

Inoue, A., Kilian, L., 2005. In-sample or out-of-sample tests of predictability: Which one should we use? Econometric Reviews 23, 371–402. doi:10.1081/ETC-200040785.

International Monetary Fund, 2015a. Macroeconomic developments in low-income developing countries. Policy Paper.

International Monetary Fund, 2015b. Myanmar: 2015 Article IV Consultation Staff Report. Country Report No. 15/267.

International Monetary Fund, 2015c. Samoa: 2015 Article IV Consultation Staff Report. Country Report No. 15/191.

International Monetary Fund, 2018a. Financial Flows Analytics. Confidential and internally accessed in November 2018.

International Monetary Fund, 2018b. International Financial Statistics. http://data.imf.org/?sk=4C514D48-B6BA-49ED-8AB9-52B0C1A0179B, accessed in November 2018.

International Monetary Fund, 2018c. Macroeconomic developments in low-income developing countries—2018. Policy Paper, March.

International Monetary Fund, 2018d. Overarching strategy on data and statistics at the fund in the digital age. Policy Paper, March.

International Monetary Fund, 2018e. Statistical Appendix of the *October 2018 World Economic Outlook*. Policy Paper.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353, 790–794.

Joseph, K., Wintoki, M.B., Zhang, Z., 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting 27, 1116–1127.

Kleinberg, J., Ludwig, J., Mullainathan, S., Obermeyer, Z., 2015. Prediction policy problems. American Economic Review 105, 491–95. doi:10.1257/aer.p20151023.

Koop, G., Onorante, L., 2013. Macroeconomic nowcasting using Google probabilities. Retrieved at: https://www.ecb.europa.eu/events/pdf/conferences/140407/OnoranteKoop_MacroeconomicNowcastingUsingGoogleProbabilities.pdf.

Laframboise, M.N., Mwase, N., Park, M.J., Zhou, Y., 2014. Revisiting tourism flows to the Caribbean: What is driving arrivals?

IMF Working Paper 14/229.

Lampos, V., Miller, A.C., Crossan, S., Stefansen, C., 2015. Advances in nowcasting influenza-like illness rates using search query logs. Scientific reports 5, 12760.

Leamer, E.E., 1978. Specification searches: Ad hoc inference with nonexperimental data. volume 53. John Wiley & Sons Incorporated.

Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. The Review of Financial Studies 21, 2535–2563. doi:10.1093/rfs/hhm056.

Li, X., Ma, J., Wang, S., Zhang, X., 2015a. How does Google search affect trader positions and crude oil prices? Economic Modelling 49, 162–171.

Li, X., Pan, B., Law, R., Huang, X., 2017. Forecasting tourism demand with composite search index. Tourism management 59, 57–66.

Li, X., Shang, W., Wang, S., Ma, J., 2015b. A MIDAS modelling framework for Chinese inflation index forecast incorporating Google search data. Electronic Commerce Research and Applications 14, 112–125. doi:10.1016/j.elerap.2015.01.001.

Moussa, F., Delhoumi, E., Ouda, O.B., 2017. Stock return and volatility reactions to information demand and supply. Research in International Business and Finance 39, 54–67.

Mullainathan, S., Spiess, J., 2017. Machine learning: an applied econometric approach. Journal of Economic Perspectives 31, 87–106.

Njuguna, C., 2018. Rnightlights: Satellite nightlight data extraction. Comprehensive R Archive Network, https://CRAN.R-project.org/package=Rnightlights (for a developer version, https://github.com/chrisvwn/Rnightlights, accessed in September 2018).

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Peltomäki, J., Graham, M., Hasselgren, A., 2018. Investor attention to market categories and market volatility: The case of emerging markets. Research in International Business and Finance 44, 532–546. doi:10.1016/j.ribaf.2017.07.124.

Preis, T., Moat, H.S., Stanley, H.E., 2013. Quantifying trading behavior in financial markets using Google Trends. Scientific reports 3, 1684.

Reis, F., Ferreira, P., Perduca, V., 2014. The use of web activity evidence to increase the timeliness of official statistics indicators. Eurostat Working Paper.

Rivera, R., 2016. A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. Tourism Management 57, 12–20.

Ross, A., 2013. Nowcasting with Google Trends: a keyword selection method. Fraser of Allander Economic Commentary 37, 54–64.

Rossi, B., Inoue, A., 2012. Out-of-sample forecast tests robust to the choice of window size. Journal of Business & Economic Statistics 30, 432–453.

Scott, S.L., Varian, H.R., 2015. Bayesian variable selection for nowcasting economic time series, in: Goldfarb, A., Greenstein, S.M., Tucker, C.E. (Eds.), Economic Analysis of the Digital Economy. University of Chicago Press. chapter 4, pp. 119–135.

Siliverstovs, B., Wochner, D.S., 2018. Google trends and reality: Do the proportions match?: Appraising the informational value of online search behavior: Evidence from Swiss tourism regions. Journal of Economic Behavior & Organization 145, 1–23.

Smith, G.P., 2012. Google internet search activity and volatility prediction in the market for foreign currency. Finance Research Letters 9, 103–110.

Smith, P., 2016. Google's MIDAS touch: Predicting UK unemployment with Internet search data. Journal of Forecasting 35, 263–284.

Somaini, P., Wolak, F.A., 2016. An algorithm to estimate the two-way fixed effects model. Journal of Econometric Methods 5, 143–152. doi:10.1515/jem-2014-0008.

StatCounter, 2018. Global Stats. http://gs.statcounter.com/search-engine-market-share, accessed in July 2018.

Stephens-Davidowitz, S., 2017. Everybody lies: big data, new data, and what the internet can tell Us about who we really are. HarperCollins New York.

Stephens-Davidowitz, S., Varian, H., 2015. A hands-on guide to Google data. Tech. Rep., http://people.ischool.berkeley.edu/~hal/Papers/2015/primer.pdf, Google.

Takeda, F., Wakao, T., 2014. Google search intensity and its relationship with returns and trading volume of Japanese stocks. Pacific-Basin Finance Journal 27, 1–18.

Tang, W., Zhu, L., 2017. How security prices respond to a surge in investor attention: Evidence from Google Search of ADRs. Global Finance Journal 33, 38–50.

Tantaopas, P., Padungsaksawasdi, C., Treepongkaruna, S., 2016. Attention effect via internet search intensity in Asia-Pacific stock markets. Pacific-Basin Finance Journal 38, 107–124.

Thomas, E., Andrés, L.A., Borja-Vega, C., Sturzenegger, G., 2018. Innovations in WASH Impact Measures: Water and Sanitation Measurement Technologies and Practices to Inform the Sustainable Development Goals. The World Bank. doi:10.1596/978-1-4648-1197-5.

United Nations Economic Commission for Europe, 2013. Classification of types of big data. URL: https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data.

Varian, H.R., 2014. Big data: New tricks for econometrics. Journal of Economic Perspectives 28, 3–28.

Vicente, M.R., López-Menéndez, A.J., Pérez, R., 2015. Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing? Technological Forecasting and Social Change 92, 132–139.

Vlastakis, N., Markellos, R.N., 2012. Information demand and stock market volatility. Journal of Banking & Finance 36, 1808–1821.

Vosen, S., Schmidt, T., 2011. Forecasting private consumption: survey-based indicators vs. Google trends. Journal of Forecasting 30, 565–578.

Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. Journal of Banking & Finance 41, 17–35. doi:10.1016/j.jbankfin.2013.12.010.

Welagedara, V., Deb, S.S., Singh, H., 2017. Investor attention, analyst recommendation revisions, and stock prices. Pacific-Basin Finance Journal 45, 211–223.

Wesolowski, A., Eagle, N., Tatem, A.J., Smith, D.L., Noor, A.M., Snow, R.W., Buckee, C.O., 2012. Quantifying the impact of human mobility on malaria. Science 338, 267–270.

World Bank, 2018. World Development Indicators. https://datacatalog.worldbank.org/dataset/world-development-indicators, accessed in November 2018.

Wu, L., Brynjolfsson, E., 2015. The future of prediction: How Google searches foreshadow housing prices and sales, in: Goldfarb, A., Greenstein, S.M., Tucker, C.E. (Eds.), Economic analysis of the digital economy. University of Chicago Press. chapter 3, pp. 89–118.

Yang, X., Pan, B., Evans, J.A., Lv, B., 2015. Forecasting Chinese tourist volume with search engine data. Tourism Management 46, 386–397.

Yung, K., Nafar, N., 2017. Investor attention and the expected returns of REITs. International Review of Economics & Finance 48, 423–439.

Zhang, Q., Pandey, B., Seto, K.C., 2016. A robust method to generate a consistent time series from dmsp/ols nighttime light data. IEEE Transactions on Geoscience and Remote Sensing 54, 5821–5831.

Appendix A. Techinical details

Appendix A.1. Introduction to Google's Search Volume Index (SVI)

The Google Trends service compiles an index, SVI, which measures how many times a keyword (or key words under a topic) has been submitted to the Google search engine. A search topic, rather than just a keyword, can also be specified to deal with ambiguity of a search word due to homographs. Figure A.1 shows an example of the SVI of search queries on country "Kenya" as a topic, from all over the world (specified as "Worldwide"), classified as the finance category (specified as "Finance"). Data points A and B are first calculated as the ratios of searches related to topic "Kenya," divided by the total searches for all queries from the same location ("Worldwide"), under the same category ("Finance"), for each period (October 2011 and July 2015, respectively). In this case, point B is the maximum of such ratios over time, and therefore, the SVI for July 2015 shows 100 and the SVI for October 2011 shows 58, which is computed as the ratio of A to B multiplied by 100.

Figure A.1: Google Trends search for "Kenya" as a search topic
Source: Google Trends' website (`https://trends.google.com/trends/`).

The SVI is constructed from sub-samples of total search data, randomly selected periodically to take a balance between usefulness and anonymity. Although all the queries submitted are stored, the Google Trends service conducts a random sampling and uses only a fraction of the entire search data to construct an SVI. Too small observations are also concealed. Re-sampling is done periodically (e.g., daily), which complicates the replication of the data downloaded previously. It is then recommended that researchers repeat downloading the same data to take the average and focus on inferred population moments, while it is also reported that the sampling generally gives reasonably precise estimates, and more than a single sample may not be needed in practice (Stephens-Davidowitz and Varian, 2015).

We retrieve monthly SVIs via Google Trends' Application Programming Interface (API), which has two major differences from the website. The SVI from the API is compiled from a 10-percent sample of total Google searches, compared with a 1 percent sampling rate for the website. On the other hand, the API provides monthly data only, while the website (`https://trends.google.com/trends/`) provides daily (if you query less than 90 days), weekly (if less than 5 years), and monthly data. The access to the API is provided through a proprietary arrangement. Data downloaded through the API need to be kept confidential and to be disposed within 12 months after the publication of the project outcome. We use program codes written in Python to retrieve data through the API.

Appendix A.2. Two-layer normalization of the SVI

Two layers of normalization are conducted in constructing an SVI. The Google servers store the information about "search volume," which is the total number of searches on query q submitted to the Google search service from location l in time t, denoted by $SV_{t,l}(q)$. However, $SV_{t,l}(q)$ is not available from the Google Trends service. Instead, we observe an SVI, which is defined in two normalization steps as follows. First, search volume on a query is normalized by the total search volume on all queries. That is, the search volume ratio (SVR), which is the ratio of search volume on query q to search volume of all the queries that were submitted in time t at location l—denoted by $SVR_{t,l}(q)$—, is constructed as follows:

$$SVR_{t,l}(q) \stackrel{\text{def}}{=} \frac{SV_{t,l}(q)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}. \tag{A.1}$$

Second, the SVR is further normalized such that the highest value under a particular data request takes 100, which defines the SVI—denoted by $SVI_{t,l}(q)$—as follows:

$$SVI_{t,l}(q) \stackrel{\text{def}}{=} \frac{SVR_{t,l}(q)}{\max_{t \in T_0} SVR_{t,l}(q)} \times 100, \tag{A.2}$$

where T_0 is the set of the time periods under the data request. Letting t^* denote the time that attains the maximum under the data request (and hence it will change under a different data request), we have:

$$SVI_{t,l}(q) = \left[\frac{SV_{t,l}(q)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right]\left[\frac{SV_{t^*,l}(q)}{\sum_{\tilde{q}} SV_{t^*,l}(\tilde{q})}\right]^{-1} \times 100 = \left[\frac{SV_{t,l}(q)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times \text{constant}. \tag{A.3}$$

Therefore, $SVI_{t,l}(q)$ is an index proportional to the frequency of searches on query q relative to the frequency of searches on all the queries submitted at location l at time t.

The two layers of normalization applied to the SVI are intended to provide an accessible and meaningful metric. The first step of the normalization controls for trivial changes in search volumes, including due to a general trend increase in search volumes observed for virtually all queries and a tendency to observe higher search volumes for queries originated from more populated locations (Stephens-Davidowitz and Varian, 2015). The second step of the normalization scales the SVI to take a value between 0 to 100 for any selection of query, time, and location, which makes the SVI accessible to wide users.

However, the two-layer normalization complicates the analysis of the SVI. For example, an increase in an SVI for query q from time t_1 to time t_2 ($> t_1$), while keeping the location the same, does not necessarily mean that query q was searched more often in time t_2. Taking two SVIs yields

$$\frac{SVI_{t_2,l}(q)}{SVI_{t_1,l}(q)} = \left[\frac{SV_{t_2,l}(q)}{SV_{t_1,l}(q)}\right]\left[\frac{\sum_{\tilde{q}} SV_{t_2,l}(\tilde{q})}{\sum_{\tilde{q}} SV_{t_1,l}(\tilde{q})}\right]^{-1}, \tag{A.4}$$

which fluctuates not only because of the change in the search volume for query q from time t_1 to time t_2, but also because of the change in the total search volume for all the queries submitted from time t_1 to time t_2. In general, there is an increasing trend in the total number of searches over time, and thus, this ratio would increase only if the search volume for query q increased at a faster pace than the increasing trend in the total number of searches.

In addition, the scaling adjustment made per data request prevents researchers from directly comparing different SVIs in levels. The units of SVIs differ across data requests to the Google Trends service. This would not be a problem if researchers could download all the SVIs of interest at once in one data request. But this is not the case in practice, not only because researchers may have second thoughts on which SVIs are needed for their analyses, but also because there are limits on the size of data requests (i.e., "quota limits"), which prevent such a massive data request at once.

Appendix A.3. Making SVIs comparable

SVIs are not comparable as they are, due to the normalization, but there is a way to make them comparable across queries—i.e., across countries in our case. Although we cannot infer search volumes in levels—i.e., $SV_{t,l}(q)$ itself—due to the first layer of the normalization, we can control for the scaling per data request made in the second layer of the normalization. After downloading two SVIs to be compared for a given period, we submit one data request for the averages of the two SVIs over the period of interest and use these values to adjust one of the two SVIs to be in the same unit of the other.

The specific procedure is as follows. Consider two SVIs, denoted by $SVI^1_{t,l}(q_1)$ and $SVI^2_{t,l}(q_2)$, to be compared for the same T periods, where superscripts 1 and 2 indicate that they are downloaded in two separate data requests. The scaling per data request results in two constants, C_1 and C_2, associated with these SVIs as follows:

$$SVI^1_{t,l}(q_1) = \left[\frac{SV_{t,l}(q_1)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times C_1, \quad SVI^2_{t,l}(q_2) = \left[\frac{SV_{t,l}(q_2)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times C_2. \tag{A.5}$$

Downloading the averages of these SVIs over time in one data request, indicated by superscript 3 and associated with a scaling constant C_3, provides the two values as follows:

$$\frac{1}{T}\sum_t SVI^3_{t,l}(q_1) = \frac{1}{T}\sum_t \left[\frac{SV_{t,l}(q_1)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times C_3, \tag{A.6}$$

$$\frac{1}{T}\sum_t SVI^3_{t,l}(q_2) = \frac{1}{T}\sum_t \left[\frac{SV_{t,l}(q_2)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times C_3. \tag{A.7}$$

Combining these, we adjust $SVI^2_{t,l}(q_2)$ as follows:

$$SVI^1_{t,l}(q_2) = SVI^2_{t,l}(q_2) \times \left[\frac{\frac{1}{T}\sum_t SVI^3_{t,l}(q_2)}{\frac{1}{T}\sum_t SVI^2_{t,l}(q_2)}\right] \times \left[\frac{\frac{1}{T}\sum_t SVI^1_{t,l}(q_1)}{\frac{1}{T}\sum_t SVI^3_{t,l}(q_1)}\right] = \left[\frac{SV_{t,l}(q_2)}{\sum_{\tilde{q}} SV_{t,l}(\tilde{q})}\right] \times C_1, \tag{A.8}$$

where $SVI^1_{t,l}(q_2)$ denotes the adjusted SVI for query q_2, which has the common scaling constant C_1 with $SVI^1_{t,l}(q_1)$. This way, $SVI^1_{t,l}(q_1)$ and $SVI^1_{t,l}(q_2)$ become comparable with each other.

We apply this adjustment bilaterally for all two pairs of SVIs of interest and make all SVIs associated with one common constant. The common constant is denoted by C_0 henceforth. The value 100 under these comparable SVIs now indicates the highest among all the SVIs over time across queries (i.e., country names) in our data set.

We cannot apply this adjustment for SVIs across categories, unfortunately. The Google Trends service does not provide the averages of SVIs across categories, which we need in the adjustment procedure. Therefore, we cannot make SVIs under different categories comparable. For example, in our data set, for Uganda, the SVI under the travel category is higher than that of the finance category, but it does not necessarily mean that more queries are submitted under the travel category than the finance category.

Appendix A.4. Conditions for proper measurement of people's attention

We establish a simple set of conditions, under which the SVI does capture the degree of people's attention to the subject of the search. Following the idea of Da et al. (2011), we assume that the search volume on query q at time t in location l is associated with some degree of people's attention to the entity represented by query q, denoted by $A_{t,l}(q)$. We need to be careful in establishing the relationship between $A_{t,l}(q)$ and $SVI_{t,l}(q)$, the latter of which requires access to the Internet and the use of the Google search service.

We first simply assume that $A_{t,l}(q)$ is the fraction of people who are interested in the entity represented by query q:

$$A_{t,l}(q) \overset{\text{def}}{=} \frac{N_{t,l}(q)}{\text{Population}_{t,l}}, \tag{A.9}$$

where $\text{Population}_{t,l}$ denotes the total population and $N_{t,l}(q)$ the number of people who are interested in the entity represented by query q, at location l in time t, regardless of the access to the Internet and the use of the Google search. This way, we put aside the issue of the intensive margin of people's attention, such

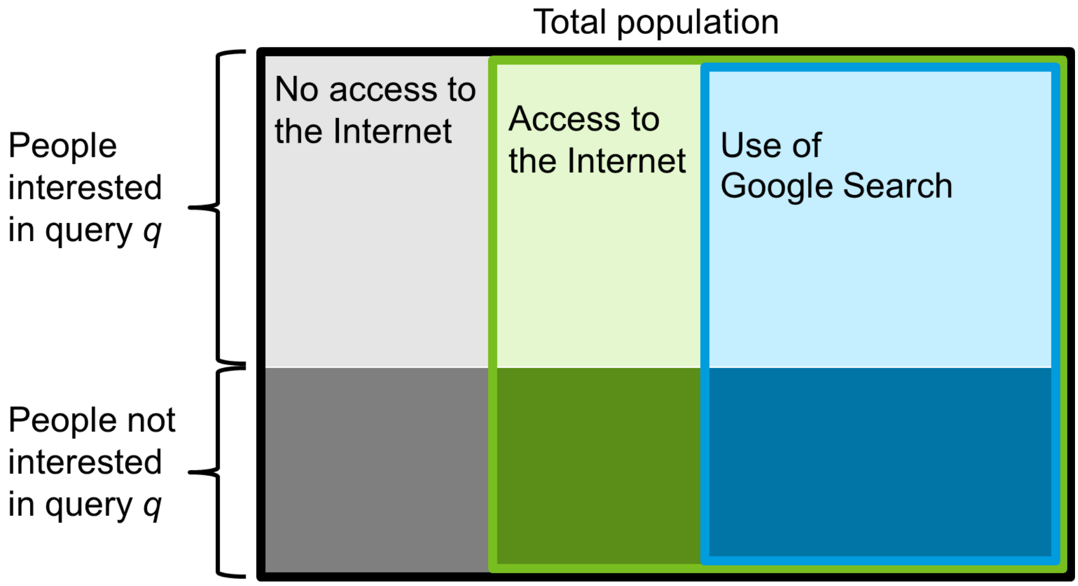


Figure A.2: Ideal case: use of Google Search as a random sampling to infer people's attention
Source: Authors.
Note: This is a conceptual diagram. The entire rectangle represents the whole population. Those who are interested in query q may not always have access to the Internet or use the Google search. In this diagram, there is no selection bias in use of the Google search with respect to whether people are interested in query q. As a result, the proportion of the people who are interested in query q (i.e., the upper light-colored part) and the people who are not (i.e., the lower dark-colored part) is the same even within those who use the Google search (the light-blue and the dark-blue parts, respectively). This is the ideal situation that is ensured by Assumptions 1-3 (especially, Assumption 2), although the assumptions can be violated in practice.

as the case where some people may be more attentive than others. Still, we need to take it into account that only part of the people interested in query q have access to the Internet and use the Google search to submit query q (Figure A.2).

We make three assumptions to establish a meaningful relationship between the SVI and peoples attention. The first assumption is about the number of searches on Google per person on average, conditional on making at least one search, which is denoted by $\bar{Q}_{t,l}(q)$. We have:

$$SV_{t,l}(q) = \bar{Q}_{t,l}(q)G_{t,l}(q), \quad SVI_{t,l}(q) = \left[\frac{\bar{Q}_{t,l}(q)G_{t,l}(q)}{\sum_{\tilde{q}}\bar{Q}_{t,l}(\tilde{q})G_{t,l}(\tilde{q})}\right] \times C_0. \tag{A.10}$$

where $G_{t,l}(q)$ denotes the number of people who search query q (at least once) on Google and C_0 is the common constant discussed in Appendix A.3. The second assumption simplifies the relationship between $G_{t,l}(q)$ and the number of people interested in query q, $N_{t,l}(q)$. The third assumption deals with the difficulty stemming from multiple counting in the sum over all submitted queries.

Assumption 1. Focus on the extensive margin
The average number of Google searches regarding query q per person, conditional on making at least one search, is constant across queries: $\bar{Q}_{t,l}(q) = \bar{Q}_{t,l}$ for any t, l, and q.

We make Assumption 1, for convenience, to focus only on the extensive margin of the search volume. Under this assumption, we have

$$SVI_{t,l}(q) = C_0\frac{SV_{t,l}(q)}{\sum_{\tilde{q}}SV_{t,l}(\tilde{q})} = C_0\frac{\bar{Q}_{t,l}G_{t,l}(q)}{\bar{Q}_{t,l}\sum_{\tilde{q}}G_{t,l}(\tilde{q})} = C_0\frac{G_{t,l}(q)}{\sum_{\tilde{q}}G_{t,l}(\tilde{q})}. \tag{A.11}$$

That is, the SVI is proportional to the fraction of people who submitted query q over the total number of people who use the Google search at location l in time t. Assumption 1 claims that the pattern of such multiple search query submissions does not change significantly or systemically across queries. It should be practically reasonable to consider that levels of SVIs for different queries are not entirely dominated by different degrees of multiple searches across queries (i.e., the intensive margin), but are mostly reflecting the varied number of searchers across queries (i.e., the extensive margin). Note that the Google Trends service excludes repeated searches from the same person over a short period (Google Trends Help, `https://support.google.com/trends/answer/4365533?hl=en&ref_topic=6248052`).

Assumption 1 may not hold in several important cases as follows. Some search activities require high-frequency updates, including seeking real-time financial investment opportunities. In this case, the SVI would be higher than the fraction of people who are interested in query q. Therefore, people's attention based on the SVI would be overestimated for the queries related to financial-sector activities (e.g., stock ticker symbols, the exchange rates), compared to slower other activities (e.g., car/home purchases, tourism). Another case is that people who are familiar with information technology may tend to submit more queries than others, and such familiarity with information technology may be correlated with some types of queries. Similarly, in the 2000s, most of Google searchers were people from colleges and universities (Stephens-Davidowitz and Varian, 2015) and they may have submitted search queries regarding their research activities (e.g., "science", "statistics") more frequently than usual people did for general search queries. In our application, people in the information technology industry may tend to be interested in queries about countries where the information technology industry is large or emerging (e.g., India). In this case, people's attention would be overestimated for these countries.

Assumption 2. Random Google search across queries
People have access to the Internet and submit queries of their interests to the Google search service at random with a constant probability that can depend on time t and location l, but not depend on query q. In other words, there is no correlation between using the Google search service and being interested in the entity represented by query q.

Assumption 2 simplifies the relationship between the SVI and the number of people interested in query q, although the assumption may be too strong. It yields:

$$\frac{G_{t,l}(q)}{\sum_{\tilde{q}} G_{t,l}(\tilde{q})} = \frac{g_{t,l}(q)N_{t,l}(q)}{\sum_{\tilde{q}} g_{t,l}(\tilde{q})N_{t,l}(\tilde{q})} = \frac{g_{t,l}N_{t,l}(q)}{g_{t,l}\sum_{\tilde{q}} N_{t,l}(\tilde{q})} = \frac{N_{t,l}(q)}{\sum_{\tilde{q}} N_{t,l}(\tilde{q})}, \tag{A.12}$$

where $g_{t,l}(q)$ denotes the probability that people who are interested in query q make a search using the Google search service and, by Assumption 2, its dependence on query q is dropped at the second equality. Therefore, combining with Assumption 1, the SVI is now proportionate to the number of people who are interested in query q. Note that this holds regardless of the improved Internet access and the increase in the use of Google search in developing countries in general during our sample period, because Assumption 2 allows the case where $g_{t,l}(q)$ can change over time and vary across locations.

Assumption 2 does not hold in the cases mostly similar to the violation of Assumption 1. As discussed for Assumption 1, the trend shift in the composition of the Google search users from people in colleges and universities to a much broader population from early 2000s to date (Stephens-Davidowitz and Varian, 2015) indicates that the probability of searching the term "science" or "statistics" was higher than other terms in the 2000s, violating Assumption 2. Also, those who are interested in information technology would be more likely to use the Google search than others. Such correlation may generate an upward bias on queries about countries where the information technology industry is large or emerging (e.g., India), as discussed for Assumption 1. Assumption 2 also implicitly requires that there must be no submission of queries by the people who are not actually interested in the entities represented by those queries. Such query submissions without interest lead to a violation of Assumption 2 and add noise in the SVI.

Assumption 3. Stable multiple interests
People may well be interested in multiple queries, but the average number per person of the interested queries is constant over time and across locations.

Assumption 3 is very useful (albeit parsimonious) in establishing a connection between the SVI and economic and social fundamentals. The SVI uses the sum of all submitted queries as its denominator, but this sum is very difficult to analyze in general. Assumption 2 simplifies the denominator to the gross headcount of people who get interested in any of submitted queries. But it is still difficult to see how much such a gross headcount would be, except for a guess that it would be much larger than the population because the sum over queries should count one person several times if that person is interested in multiple queries. Assumption 3 claims that this multiple counting occurs to everyone to the same extent on average, establishing the following simple relationship:

$$\sum_{\tilde{q}} N_{t,l}(\tilde{q}) = \sum_{i} M_{t,l}(i) = \bar{M} \times \text{Population}_{t,l}, \tag{A.13}$$

where $M_{t,l}(i)$ denotes the number of queries that person i at location l in time t is interested in, and $\bar{M}$ is the average of such a number per person, assumed to be constant by Assumption 3. The first equality holds because the sum counted over queries on the left-hand side is just recounted as the sum over persons on the right-hand side. Note that Assumption 3 has nothing to do with whether people access the Internet or how often they search on Google. Rather, Assumption 3 is about a human nature of getting interested in multiple things, which would be generic enough to justify the parsimonious assumption that the average per person would not be different across locations and over time.

Propostion 1. SVI as a measure of attention
Under Assumptions 1, 2, and 3, the SVI is proportionate to the degree of people's attention on the entity represented by a query:

$$SVI_{t,l}(q) = C_0 \frac{G_{t,l}(q)}{\sum_{\tilde{q}} G_{t,l}(\tilde{q})} = C_0 \frac{N_{t,l}(q)}{\sum_{\tilde{q}} N_{t,l}(\tilde{q})} = \left(\frac{C_0}{\bar{M}}\right) \frac{N_{t,l}(q)}{\text{Population}_{t,l}} = \left(\frac{C_0}{\bar{M}}\right) A_{t,l}(q). \tag{A.14}$$

Proposition 1 formalizes the use of the SVI to analyze people's attention in general. It justifies the use of the SVI and sets a basis to discuss possible biases that could arise in the estimates based on the SVI.

Appendix A.5. How to detect jumps in the SVIs

We employ a methodology in the finance literature to detect acute increases in the SVIs. We applyLee and Mykland (2008)'s continuous-time model of the log level of stock prices to the log level of SVIs. It uses the difference between squared percent changes and consecutive absolute percent changes (called bi-power variations) to identify huge changes within a period. While the finance model is intended to be applied at a very high frequency such as a 30-minute window, we apply it to monthly data, suffering from lower efficiency and higher bias from the remaining mean drift component, which should be negligible only if the observation frequency goes to infinity. On this account, we first regress the log of SVIs on a third-order polynomial trend and the monthly dummies. We then use its residuals for calculating squared and bi-power percent changes. The jump detection is based on a statistical inference at the 1 percent significance level. It requires estimating the time-varying instantaneous volatility without jumps, for which we use the rolling-window bi-power variation over the past 36 months, excluding the current month. To keep the observations as many as possible, the rolling-window estimation is conducted forwardly (i.e., over 36 months ahead) for the first 36 months in the sample. We only focus on positive jumps and ignore negative jumps (i.e., huge drops), because an acute decrease of people's attention is not intuitive, and thus, its detection would be erroneous. To increase accuracy, we iterate the procedure once again after removing the detected jumps. Furthermore, we conduct the same procedure at the quarterly frequency, by taking period averages and setting the size of the rolling window at 12 quarters. We then take the union of jumps detected monthly and quarterly.

Appendix B. Supplementary tables

Table B.1: Groupings of the economies

Low-income developing countries (LIDCs; 59)[1]
Afghanistan, Bangladesh, Benin, Bhutan, Burkina Faso, Burundi, Cambodia, Cameroon, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Republic of the Congo, Côte d'Ivoire, Djibouti, Eritrea, Ethiopia, The Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, Honduras, Kenya, Kiribati, Kyrgyz Republic, Lao P.D.R., Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Moldova, Mozambique, Myanmar, Nepal, Nicaragua, Niger, Nigeria, Papua New Guinea, Rwanda, So Tom and Prncipe, Senegal, Sierra Leone, Solomon Islands, Somalia, South Sudan, Sudan, Tajikistan, Tanzania, Timor-Leste, Togo, Uganda, Uzbekistan, Vietnam, Yemen, Zambia, Zimbabwe
Emerging market economies (EMEs; 95)[2]
Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Azerbaijan, The Bahamas, Bahrain, Barbados, Belarus, Belize, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Cabo Verde, Chile, China, Colombia, Costa Rica, Croatia, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eswatini, Fiji, Gabon, Georgia, Grenada, Guatemala, Guyana, Hungary, India, Indonesia, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kosovo, Kuwait, Lebanon, Libya, Malaysia, Maldives, Marshall Islands, Mauritius, Mexico, Micronesia, Mongolia, Montenegro, Morocco, Namibia, Nauru, North Macedonia, Oman, Pakistan, Palau, Panama, Paraguay, Peru, Philippines, Poland, Qatar, Romania, Russia, Samoa, Saudi Arabia, Serbia, Seychelles, South Africa, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Suriname, Syria, Thailand, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Tuvalu, Ukraine, United Arab Emirates, Uruguay, Vanuatu, Venezuela
Advanced economies (AEs; 39)
Australia, Austria, Belgium, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong S.A.R. of China, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Macao S.A.R. of China, Malta, Netherlands, New Zealand, Norway, Portugal, Puerto Rico, San Marino, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Taiwan Province of China, United Kingdom, United States

Source: World Economic Outlook (IMF, 2018e).
[1] See also IMF, 2018c, Appendix I for the update of the classification of the LIDCs.
[2] EMEs are defined as the residual group of economies that are not included in AEs nor LIDCs.

Table B.2: Groupings of the economies

Panel A: Major categories

Google Trends Main Categories (25 plus all)	
All Categories	Internet & Telecom
Arts & Entertainment	Jobs & Education
Autos & Vehicles	**Law & Government**
Beauty & Fitness	News
Books & Literature	Online Communities
Business & Industrial	People & Society
Computers & Electronics	Pets & Animals
Finance	Real Estate
Food & Drink	Reference
Games	Science
Health	Shopping
Hobbies & Leisure	Sports
Home & Garden	**Travel**

Panel B: Subcategories under selected five major categories

Finance	Business & Industrial	Law & Government	Health	Travel
Accounting & Auditing	Advertising & Marketing	Government	Aging & Geriatrics	Air Travel
Banking	Aerospace & Defense	Legal	Alternative & Natural Medicine	Bus & Rail
Credit & Lending	Agriculture & Forestry	Military	Health Conditions	Car Rental & Taxi Services
Currencies & Foreign Exchange	Automotive Industry	Public Safety	Health Education & Medical Training	Carpooling & Ridesharing
Financial Planning	Business Education	Social Services	Health Foundations & Medical Research	Cruises & Charters
Grants & Financial Assistance	Business Finance		Health News	Hotels & Accommodations
Insurance	Business News		Medical Devices & Equipment	Luggage & Travel Accessories
Investing	Business Operations		Medical Facilities & Services	Specialty Travel
Retirement & Pension	Business Services		Medical Literature & Resources	Tourist Destinations
	Chemicals Industry		Men's Health	Travel Agencies & Services
	Construction & Maintenance		Mental Health	Travel Guides & Travelogues
	Energy & Utilities		Nursing	
	Enterprise Technology		Nutrition	
	Entertainment Industry		Oral & Dental Care	
	Hospitality Industry		Pediatrics	
	Industrial Materials & Equipment		Pharmacy	
	Manufacturing		Public Health	
	Metals & Mining		Reproductive Health	
	Pharmaceuticals & Biotech		Substance Abuse	
	Printing & Publishing		Women's Health	
	Professional & Trade Associations			
	Retail Trade			
	Small Business			
	Textiles & Nonwovens			
	Transportation & Logistics			

Source: Google Trends website (`https://trends.google.com/trends/`).
Note: Queries are assigned to categories using a natural language processing algorithm, whose details are not disclosed to the public.

Table B.3: Variable definitions and data sources

Variable	Transformation	Series code	Database
Google search volume index (SVI)	Natural logarithm	Set a country name as a search topic	Google Trends
Foreign direct investment (FDI) inflows	Natural logarithm	IFDI	FFA
Non-FDI private capital inflows	Natural logarithm	ICAPFLP minus IFDI	FFA
Total private capital inflows	Natural logarithm	ICAPFLP	FFA
Total capital inflows	Natural logarithm	ICAPFL	FFA
Export price growth (export-value weighted average of import deflators in export destination countries)	Percent change	TM_D_WX001	WEO (GEE)
Import price growth (import-value weighted average of export deflators in import origination countries)	Percent change	TX_D_WM001	WEO (GEE)
Trading partners' growth (export-value weighted average of real GDP growth in export destination countries)	Percent change	NGDP_R_WX001	WEO (GEE)
Real effective exchange rate (REER)	Natural logarithm or percent change, p.a.	EREER_IX	IFS
Fiscal spending	Natural logarithm	GGX	WEO
Inflation	Percent change, p.a.	PCPI_PCH	WEO
Nominal exchange rate (local currencies to one U.S. dollar)	Percent change, p.a.	ENDA	WEO
Population	Natural logarithm	LP	WEO
Real exports	Natural logarithm	TX_R	WEO
Trade openness	Natural logarithm	(BMGS_BP6 + BXGS_BP6) / NGDPD	WEO
Age dependency ratio	None	SP.POP.DPND	WDI
GDP (constant 2010 US$)	Natural logarithm	NY.GDP.MKTP.KD	WDI
International tourism, number of arrivals	Natural logarithm	ST.INT.ARVL	WDI
International tourism, receipts	Natural logarithm	ST.INT.RCPT.CD	WDI
Internet users (per 100 people)	Natural logarithm	IT.NET.USER.P2	WDI
Capital account openness index	None	KA_OPEN	Chinn and Ito (2006), updated as of July 20, 2017.
Nighttime lights per area, HSW (2012)	Natural logarithm	lndn	Henderson et al. (2012)
Nighttime lights per area, Rnightlights	Natural logarithm	Compiled via an R package, **Rnightlights**	Njuguna (2018)

Sources: Chinn and Ito (2006); Earth Observation Group; Financial Flows Analytics (FFA, IMF, 2018a); GADM (2018); Google Trends; Henderson et al. (2012); International Financial Statistics (IFS, IMF, 2018b); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (WDI, World Bank, 2018); and World Economic Outlook (WEO, IMF, 2018e).

Note: For nighttime lights, the nighttime is at some instant during 8:30 and 10:00 pm local time, depending on the location and the light intensity is digitalized as an integer between 0 (no light) to 63 (Henderson et al., 2012). The R package **Rnightlights** (Njuguna, 2018) compiles nighttime light data for 1992-2013 based on DMSP OLS data and for 2015-2016 based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) data. The DMSP OLS data are based on the processed images provided by National Geophysical Data Center, while images are collected by U.S. Air Force Weather Agency. The VIIRS DNB data are produced by the Earth Observation Group, NOAA/NCEI. The FFA database compiled from the IMF's Balance of Payments Statistics, IFS, and WEO databases, World Bank's WDI database, Haver Analytics, CEIC Asia database, and CEIC China database. DMSP OLS: Defense Meteorological Satellite Program Operational Linescan System; GEE: Global Economic Environment; NCEI: National Centers for Environmental Information; NOAA: National Oceanic and Atmospheric Administration.

Table B.4: Summary statistics for LIDCs

Variable	Percentiles 25th	50th	75th	Mean	SD	Within SD	# of obs.	# of countries	# of obs. per country
Google search volume index (SVI): All	-1.38	-0.70	0.22	-0.60	1.16	0.24	760	59	12.9
SVI: Finance	-2.80	-1.85	-0.91	-1.78	1.39	0.31	760	59	12.9
SVI: Business and industrial	-2.11	-1.25	-0.48	-1.30	1.27	0.20	760	59	12.9
SVI: Law and government	-1.09	-0.17	0.50	-0.27	1.15	0.21	760	59	12.9
SVI: Health	-1.49	-0.63	0.27	-0.64	1.28	0.25	760	59	12.9
SVI: Travel	-2.17	-1.17	-0.13	-1.11	1.36	0.21	760	59	12.9
Domestic SVI: All	3.66	3.95	4.21	3.87	0.46	0.26	751	59	12.7
Domestic SVI: Finance	2.34	3.02	3.52	2.84	0.92	0.43	722	59	12.2
Domestic SVI: Business and industrial	2.48	3.25	3.92	3.11	0.93	0.39	741	59	12.6
Domestic SVI: Law and government	2.64	3.28	3.81	3.14	0.82	0.39	745	59	12.6
Domestic SVI: Health	2.14	2.91	3.51	2.77	0.94	0.44	715	59	12.1
Domestic SVI: Travel	2.22	2.92	3.50	2.81	0.93	0.46	719	59	12.2
Foreign direct investment (FDI) inflows	4.11	5.64	6.80	5.35	2.04	1.02	626	54	11.6
Non-FDI private capital inflows	4.74	5.82	6.70	5.60	1.75	0.94	505	53	9.5
Total private capital inflows	5.34	6.39	7.41	6.26	1.69	0.85	575	53	10.8
Total capital inflows	5.34	6.43	7.42	6.28	1.69	0.87	579	53	10.9
Export price growth	-3.27	4.31	8.26	1.85	7.83	7.80	767	59	13.0
Import price growth	-3.17	5.16	8.90	2.31	8.40	8.39	767	59	13.0
Trading partners growth	2.96	4.21	5.55	4.26	2.21	1.64	767	59	13.0
REER (log level)	4.55	4.61	4.66	4.61	0.15	0.13	717	56	12.8
REER (percent change, p.a.)	-2.40	1.12	4.90	1.55	7.04	6.58	713	56	12.7
Fiscal spending	3.40	5.73	7.54	5.42	2.99	0.58	746	58	12.9
Inflation	3.12	6.42	10.04	8.22	17.01	14.82	746	58	12.9
Nominal exchange rate	-0.65	0.79	6.65	3.52	12.82	11.54	759	59	12.9
Real exports	-0.58	0.49	1.27	0.29	1.68	0.36	702	54	13.0
Trade openness	3.87	4.16	4.54	4.20	0.44	0.18	751	59	12.7
Age dependency ratio	70.00	83.74	91.58	80.55	16.14	3.34	762	59	12.9
GDP (constant 2010 US$)	1.32	2.18	3.07	2.08	1.51	0.21	724	57	12.7
International tourism, number of arrivals	11.31	12.55	13.67	12.44	1.59	0.45	655	54	12.1
International tourism, receipts	17.43	18.77	19.97	18.56	1.86	0.64	675	57	11.8
Internet users (per 100 people)	0.34	1.37	2.31	1.26	1.37	0.94	687	59	11.6
Population	1.47	2.37	3.19	2.19	1.52	0.09	747	58	12.9
Capital account openness index	0.17	0.17	0.41	0.33	0.32	0.08	650	56	11.6
Nighttime lights per area, HSW (2012)	-2.89	-2.09	-1.05	-1.99	1.44	0.13	290	58	5.0
Nighttime lights per area, **Rnightlights**	-2.55	-1.62	-0.58	-1.61	1.48	0.41	701	59	11.9

Sources: Chinn and Ito (2006); Earth Observation Group; Financial Flows Analytics (IMF, 2018a); GADM (2018); Google Trends; Henderson et al. (2012); International Financial Statistics (IMF, 2018b); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' calculation.

Note. Sample period: 2004-2016. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SD: standard deviation.

Table B.5: Pairwise correlation coefficients for selected variables for LIDCs

Variable	Google SVI						Nighttime lights per area, **Rnightlights**
	All	Finance	Business and Industrial	Law and Government	Health	Travel	
Google search volume index (SVI): All	1	-	-	-	-	-	0.32*
SVI: Finance	0.93*	1	-	-	-	-	0.42*
SVI: Business and industrial	0.92*	0.92*	1	-	-	-	0.34*
SVI: Law and government	0.96*	0.91*	0.92*	1	-	-	0.32*
SVI: Health	0.96*	0.92*	0.91*	0.92*	1	-	0.28*
SVI: Travel	0.83*	0.83*	0.78*	0.77*	0.81*	1	0.38*
GDP (constant 2010 US$)	0.83*	0.82*	0.82*	0.82*	0.81*	0.58*	0.24*
Real exports	0.78*	0.82*	0.84*	0.79*	0.76*	0.64*	0.32*
International tourism, arrivals	0.73*	0.78*	0.76*	0.74*	0.71*	0.67*	0.35*
International tourism, receipts	0.72*	0.72*	0.69*	0.70*	0.70*	0.72*	0.30*
Inflation	0.06	0.05	0.04	0.07	0.05	0.01	-0.05
Nominal exchange rate	0.06	0.10*	0.04	0.04	0.06	-0.01	-0.04
Real effective exchange rate (percent change, p.a.)	0.06	0.01	0.02	0.05	0.04	0.03	-0.00
Foreign direct investment (FDI) inflows	0.56*	0.59*	0.59*	0.52*	0.57*	0.46*	0.14*
Non-FDI private capital inflows	0.66*	0.68*	0.68*	0.64*	0.67*	0.51*	0.17*
Total private capital inflows	0.67*	0.70*	0.69*	0.66*	0.67*	0.51*	0.15*
Total capital inflows	0.67*	0.70*	0.69*	0.66*	0.67*	0.50*	0.14*
Export price growth	-0.05	-0.09	-0.04	-0.00	-0.03	0.04	0.07
Import price growth	-0.06	-0.11*	-0.04	-0.02	-0.04	0.03	0.07
Trading partners' growth	-0.13*	-0.18*	-0.21*	-0.09	-0.13*	-0.12*	-0.13*
Fiscal spending	0.44*	0.38*	0.40*	0.41*	0.43*	0.32*	0.09
Trade openness	-0.12*	-0.08	0.01	-0.11*	-0.12*	0.08	0.18*
Age dependency ratio	-0.21*	-0.26*	-0.20*	-0.22*	-0.13*	-0.31*	-0.56*
Internet users (per 100 people)	0.21*	0.32*	0.25*	0.20*	0.17*	0.23*	0.53*
Capital account openness index	0.08	0.11*	0.07	0.12*	0.09	0.19*	0.18*
Nighttime lights per area, HSW (2012)	0.26*	0.40*	0.29*	0.27*	0.19*	0.32*	0.99*
Nighttime lights per area, **Rnightlights**	0.32*	0.42*	0.34*	0.32*	0.28*	0.38*	1

Sources: Chinn and Ito (2006); Earth Observation Group; Financial Flows Analytics (IMF, 2018a); GADM (2018); Google Trends; Henderson et al. (2012); International Financial Statistics (IMF, 2018b); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' calculation.

Note. Sample period: 2004-2016. Superscript * indicates significance at the one percent level. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate.

Table B.6: Nighttime lights (NLs) and real GDP in LIDCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables	Real GDP							
NLs from HSW (2012)	0.245***				0.084*	0.017		
	(0.058)				(0.046)	(0.018)		
NLs from HSW (2012) (lag)					-0.076**	-0.007		
					(0.035)	(0.014)		
NLs from **Rnightlights**		0.244***	0.246***	0.257***			0.067*	0.015
		(0.061)	(0.053)	(0.048)			(0.036)	(0.013)
NLs from **Rnightlights** (lag)							-0.063**	-0.007
							(0.029)	(0.009)
Lagged real GDP					0.897***	0.789***	0.925***	0.855***
					(0.039)	(0.047)	(0.022)	(0.035)
Other lagged controls included	NO	NO	NO	NO	NO	YES	NO	YES
Observations	917	917	1,196	1,306	861	382	1,194	686
Sample period	1992-2008	1992-2008	1992-2013	1992-2013, 2015-2016	1992-2008	1992-2008	1992-2013, 2015-2016	1992-2013, 2015-2016
Number of countries	56	56	57	57	56	53	56	53
Adjusted R-squared	0.697	0.695	0.776	0.801	0.932	0.936	0.967	0.973
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006); Earth Observations Group; GADM (2018); Google Trends; Henderson et al. (2012); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' estimation.

Note. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. The NLs from HSW (2012) line shows the coefficients on NL data (variable lndn) compiled by Henderson et al. (2012), available for 1992-2008. The NLs from Rnightlights line shows the coefficients on NL data compiled by R package **Rnightlights** developed by Njuguna (2018), available for 1992-2013 based on DMSP OLS data (also used by Henderson et al., 2012) and for 2015-2016 based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) data. The DMSP OLS data are based on the processed images provided by National Geophysical Data Center, while images are collected by U.S. Air Force Weather Agency. The VIIRS DNB data are produced by the Earth Observation Group, NOAA/NCEI. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. DMSP OLS: Defense Meteorological Satellite Program Operational Linescan System; LIDCs: low-income developing countries; NCEI: National Centers for Environmental Information; NOAA: National Oceanic and Atmospheric Administration; REER: real effective exchange rate; SVI: search volume index.

Table B.7: Nighttime lights (NLs) and real GDP in EMEs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables				Real GDP				
NLs from HSW (2012)	0.298***				0.035	0.032*		
	(0.058)				(0.021)	(0.016)		
NLs from HSW (2012) (lag)					-0.077***	-0.035		
					(0.021)	(0.024)		
NLs from **Rnightlights**		0.272***	0.332***	0.237***			0.049**	0.033**
		(0.065)	(0.060)	(0.045)			(0.024)	(0.013)
NLs from **Rnightlights** (lag)							-0.054***	-0.033***
							(0.020)	(0.011)
Lagged real GDP					0.892***	0.893***	0.903***	0.912***
					(0.044)	(0.063)	(0.038)	(0.022)
Other lagged controls included	NO	NO	NO	NO	NO	YES	NO	YES
Observations	1,459	1,327	1,727	1,885	1,369	498	1,726	897
Sample period	1992-2008	1992-2008	1992-2013	1992-2013, 2015-2016	1992-2008	1992-2008	1992-2013, 2015-2016	1992-2013, 2015-2016
Number of countries	88	80	80	80	88	67	80	71
Adjusted R-squared	0.777	0.770	0.798	0.795	0.952	0.955	0.962	0.975
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006); Earth Observations Group; GADM (2018); Google Trends; Henderson et al. (2012); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' estimation.

Note. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. The NLs from HSW (2012) line shows the coefficients on NL data (variable *lndn*) compiled by Henderson et al. (2012), available for 1992-2008. The NLs from Rnightlights line shows the coefficients on NL data compiled by R package **Rnightlights** developed by Njuguna (2018), available for 1992-2013 based on DMSP OLS data (also used by Henderson et al., 2012) and for 2015-2016 based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) data. The DMSP OLS data are based on the processed images provided by National Geophysical Data Center, while images are collected by U.S. Air Force Weather Agency. The VIIRS DNB data are produced by the Earth Observation Group, NOAA/NCEI. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. Among EMEs, the NL data exclude countries identified as outliers by Henderson et al. (2012, footnote 16, p. 1011; Bahrain, Equatorial Guinea, Serbia, Montenegro). For the data compiled by **Rnightlights**, several large economies are also excluded due to their heavy computational burden (Brazil, Chile, China, Indonesia, India, Mexico, Peru, Russia). DMSP OLS: Defense Meteorological Satellite Program Operational Linescan System; EMEs: emerging market economies; NCEI: National Centers for Environmental Information; NOAA: National Oceanic and Atmospheric Administration; REER: real effective exchange rate; SVI: search volume index.

Table B.8: Best specifications that minimize out-of-sample MSE for LIDCs

Dependent variable	Independent variables	Specification that minimize the MSE in out-of-sample nowcasting for 2013-2016
Real GDP	Controls only	Lagged dependent variable, population, trade openness, fiscal spending, inflation, trading partners' growth
	Controls + SVIs	Lagged dependent variable, population, Internet users, trade openness, fiscal spending, inflation, **finance SVI**, **law-government SVI**
Real exports	Controls only	Lagged dependent variable, trade openness, fiscal spending, REER, capital openness, age dependency ratio
	Controls + SVIs	Lagged dependent variable, fiscal spending, REER, capital openness, age dependency ratio, **health SVI**
Tourist arrivals	Controls only	Lagged dependent variable, trade openness, fiscal spending, REER, trading partners' growth
	Controls + SVIs	Lagged dependent variable, trade openness, fiscal spending, REER, trading partners' growth, export price growth, **health SVI**, **travel SVI**
Inflation	Controls only	Lagged dependent variable, population, trade openness, REER (percent change)
	Controls + SVIs	Lagged dependent variable, population, fiscal spending, REER (percent change), **finance SVI**, **business-industrial SVI**, **health SVI**, **travel SVI**
Nominal exchange rate	Controls only	Lagged dependent variable, Internet users, import price growth
	Controls + SVIs	Real GDP, import price growth, age dependency ratio, **finance SVI**, **business-industrial SVI**, **law-government SVI**, **travel SVI**
Private capital inflows	Controls only	Internet users, real GDP, trade openness, trading partners' growth
	Controls + SVIs	Internet users, real GDP, trade openness, export price growth, **business-industrial SVI**, **health SVI**, **travel SVI**
FDI inflows	Controls only	Lagged dependent variable, population, real GDP, REER, trading partners' growth
	Controls + SVIs	Lagged dependent variable, population, Internet users, real GDP, REER, export price growth, **finance SVI**, **business-industrial SVI**, **health SVI**

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. All control variables are one-year lagged, while SVIs are contemporaneous. See the note under Table 7 for the estimation details. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; MSE: mean square error; REER: real effective exchange rate; SVI: search volume index.

Table B.9: Best specifications that minimize out-of-sample MSE for EMEs

Dependent variable	Independent variables	Specification that minimize the MSE in out-of-sample nowcasting for 2013-2016
Real GDP	Controls only	Lagged dependent variable, Internet users, trade openness, fiscal spending, trading partners' growth, capital openness, age dependency ratio
	Controls + SVIs	Lagged dependent variable, trade openness, fiscal spending, trading partners' growth, capital openness, age dependency ratio, **finance SVI**, **business-industrial SVI**
Real exports	Controls only	Lagged dependent variable, population, inflation, trading partners' growth, age dependency ratio
	Controls + SVIs	Lagged dependent variable, Internet users, REER, inflation, trading partners' growth, age dependency ratio, **finance SVI**, **law-government SVI**, **travel SVI**
Tourist arrivals	Controls only	Lagged dependent variable, Internet users, trade openness, REER, export price growth
	Controls + SVIs	(same as above)
Inflation	Controls only	Lagged dependent variable, Internet users, import price growth, capital openness
	Controls + SVIs	Lagged dependent variable, Internet users, import price growth, capital openness, **business-industrial SVI**, **law-government SVI**, **health SVI**, **travel SVI**
Nominal exchange rate	Controls only	Lagged dependent variable, Internet users, fiscal spending, inflation, trading partners' growth, import price growth
	Controls + SVIs	Lagged dependent variable, Internet users, fiscal spending, inflation, trading partners' growth, import price growth, **finance SVI**, **law-government SVI**
Private capital inflows	Controls only	Lagged dependent variable, trade openness, REER, inflation, trading partners' growth, capital openness
	Controls + SVIs	Lagged dependent variable, real GDP, trade openness, fiscal spending, REER, inflation, trading partners' growth, export price growth, capital openness, age dependency ratio, **finance SVI**, **business-industrial SVI**, **travel SVI**
FDI inflows	Controls only	Lagged dependent variable, REER, trading partners' growth, export price growth, age dependency ratio
	Controls + SVIs	(Same as above)

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. All control variables are one-year lagged, while SVIs are contemporaneous. See the note under Table 7 for the estimation details. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. EMEs: emerging market economies; MSE: mean square error; REER: real effective exchange rate; SVI: search volume index.

Table B.10: Regressions, excluding periods with jumps in SVIs, for LIDCs

Dependent variables	(1) Real GDP	(2) Real exports	(3) Tourist arrivals	(4) Inflation	(5) Nominal exchange rate	(6) Private capital inflows	(7) FDI inflows
SVI: Finance	-0.00	0.04	-0.06	3.43**	10.07***	-0.35	0.05
	(0.01)	(0.06)	(0.11)	(1.35)	(3.19)	(0.23)	(0.33)
SVI: Business and industrial	0.11***	0.18	0.48**	-2.52	-2.14	1.71***	1.24
	(0.03)	(0.11)	(0.22)	(1.77)	(5.13)	(0.40)	(0.75)
SVI: Law and government	-0.05***	-0.27**	-0.37**	0.56	-6.51**	0.00	-0.65
	(0.02)	(0.11)	(0.15)	(1.61)	(2.59)	(0.33)	(0.41)
SVI: Health	-0.06**	-0.01	-0.34**	2.18*	0.89	-0.67**	-0.40
	(0.02)	(0.05)	(0.13)	(1.20)	(2.47)	(0.30)	(0.43)
SVI: Travel	0.02	0.04	0.28**	-4.33**	-2.45	-0.72***	-0.19
	(0.01)	(0.07)	(0.11)	(1.67)	(2.62)	(0.24)	(0.34)
Lagged dependent variable	0.78***	0.83***	0.60***	0.28***	0.16***	0.03	0.19*
	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.10)	(0.10)
Population (lag)	0.05	0.42	-0.57	-3.50	6.02	-3.25*	-0.42
	(0.07)	(0.57)	(0.50)	(17.39)	(21.27)	(1.61)	(1.73)
Internet users (lag)	-0.00	0.03	0.05*	0.84	0.86	0.03	0.00
	(0.00)	(0.02)	(0.03)	(0.50)	(0.92)	(0.09)	(0.13)
Real GDP (lag)		-0.44*	-0.48**	2.55	2.19	1.14	1.89*
		(0.22)	(0.23)	(3.36)	(5.63)	(0.80)	(1.02)
Trade openness (lag)	0.01	-0.07	-0.28**	1.51	-7.25**	0.70**	0.68*
	(0.01)	(0.10)	(0.10)	(1.56)	(2.97)	(0.27)	(0.34)
Fiscal spending (lag)	0.02**	0.02	0.19***	-1.76	-6.68***	0.44	-0.10
	(0.01)	(0.05)	(0.07)	(1.38)	(2.23)	(0.27)	(0.28)
REER, log level (lag)	-0.01	0.11	-0.33**			-0.02	-0.15
	(0.02)	(0.09)	(0.14)			(0.47)	(0.45)
REER, percent change (lag)				-0.20***			
				(0.04)			
Inflation (lag)	-0.00**	-0.00	-0.00		-0.12***	-0.00	-0.00
	(0.00)	(0.00)	(0.00)		(0.03)	(0.01)	(0.01)
Trading partners' growth (lag)	-0.00	-0.01	0.01	0.07	-0.21	0.04	0.03
	(0.00)	(0.01)	(0.01)	(0.31)	(0.39)	(0.04)	(0.03)
Export price growth (lag)	0.00	0.00	-0.00			0.02**	0.03*
	(0.00)	(0.01)	(0.01)			(0.01)	(0.01)
Import price growth (lag)				0.08	-0.13		
				(0.08)	(0.12)		
Capital account openness (lag)	-0.01	0.07	0.02	-0.57	-3.82	0.52	-0.91*
	(0.01)	(0.10)	(0.12)	(2.49)	(3.62)	(0.45)	(0.46)
Age dependency ratio (lag)	0.00	-0.00	0.01	0.03	0.27	0.00	0.00
	(0.00)	(0.01)	(0.01)	(0.13)	(0.20)	(0.03)	(0.02)
Observations	503	494	455	500	524	355	422
Number of countries	53	52	51	53	54	48	48
Adjusted R-squared	0.979	0.773	0.767	0.317	0.336	0.512	0.343
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES
Excluding periods with jumps	YES	YES	YES	YES	YES	YES	YES

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Appendix I, Section E for the methodology used to detect jumps. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; REER: real effective exchange rate; SVI: search volume index.

Table B.11: Regressions with lagged SVIs for LIDCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	Real GDP	Real exports	Tourist arrivals	Inflation	Nominal exchange rate	Private capital inflows	FDI inflows
SVI: Finance	0.00	-0.05	0.09	2.63**	6.89**	-0.14	-0.04
	(0.02)	(0.04)	(0.08)	(1.13)	(2.86)	(0.32)	(0.30)
SVI: Business and industrial	0.06**	0.15	0.28	-3.04*	-0.87	0.15	1.16*
	(0.03)	(0.11)	(0.23)	(1.79)	(4.29)	(0.58)	(0.64)
SVI: Law and government	-0.08***	-0.20***	-0.40***	0.58	-6.78**	0.17	-0.49
	(0.02)	(0.06)	(0.13)	(1.51)	(2.55)	(0.27)	(0.45)
SVI: Health	-0.02*	0.00	-0.27**	2.14*	-0.57	-0.30	-0.47
	(0.01)	(0.04)	(0.11)	(1.13)	(1.61)	(0.20)	(0.30)
SVI: Travel	0.00	0.04	0.20	-3.23**	-0.87	-0.29	0.29
	(0.03)	(0.06)	(0.15)	(1.59)	(2.41)	(0.27)	(0.33)
SVI: Finance (lag)	-0.03	0.06	-0.16	0.78	3.70	-0.21	-0.13
	(0.03)	(0.06)	(0.10)	(1.11)	(2.44)	(0.31)	(0.33)
SVI: Business and industrial (lag)	0.05*	-0.02	0.05	1.19	-2.36	1.47**	0.18
	(0.03)	(0.09)	(0.09)	(1.87)	(3.34)	(0.68)	(0.65)
SVI: Law and government (lag)	0.04**	0.01	0.10	-1.66	1.20	-0.13	0.33
	(0.02)	(0.06)	(0.10)	(1.35)	(2.58)	(0.28)	(0.41)
SVI: Health (lag)	-0.03*	-0.01	0.04	-0.08	2.07	-0.48*	0.05
	(0.02)	(0.04)	(0.09)	(1.00)	(1.99)	(0.27)	(0.37)
SVI: Travel (lag)	0.01	-0.02	0.12	0.21	-2.28	-0.24	-0.75*
	(0.02)	(0.06)	(0.14)	(1.37)	(2.25)	(0.35)	(0.40)
Control variables included	YES	YES	YES	YES	YES	YES	YES
Observations	597	587	535	595	620	423	499
Number of countries	53	53	52	54	55	49	49
Adjusted R-squared	0.960	0.799	0.731	0.317	0.323	0.361	0.277
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES
Excluding periods with jumps	NO	NO	NO	NO	NO	NO	NO

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; SVI: search volume index.

Table B.12: Regressions with domestically-made SVIs (DSVIs) for LIDCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	Real GDP	Real exports	Tourist arrivals	Inflation	Nominal exchange rate	Private capital inflows	FDI inflows
SVI: Finance	0.01	-0.03	-0.03	2.99**	7.01**	-0.30	-0.36
	(0.02)	(0.04)	(0.10)	(1.18)	(2.73)	(0.27)	(0.32)
SVI: Business and industrial	0.07**	0.10	0.30	-0.67	1.49	0.81**	1.12*
	(0.03)	(0.07)	(0.24)	(1.69)	(5.35)	(0.38)	(0.58)
SVI: Law and government	-0.06***	-0.11**	-0.46***	-0.85	-8.93***	0.38	-0.14
	(0.02)	(0.05)	(0.13)	(1.36)	(3.27)	(0.36)	(0.33)
SVI: Health	-0.05**	0.01	-0.16	2.17	2.15	-0.49*	-0.27
	(0.02)	(0.05)	(0.11)	(1.31)	(2.27)	(0.28)	(0.34)
SVI: Travel	0.01	-0.02	0.25**	-3.85**	-3.55	-0.48**	-0.12
	(0.01)	(0.05)	(0.10)	(1.55)	(3.04)	(0.21)	(0.40)
DSVI: Finance	-0.02	-0.01	0.04	0.39	2.16	0.02	0.21
	(0.01)	(0.03)	(0.05)	(0.71)	(1.36)	(0.19)	(0.23)
DSVI: Business and industrial	0.01	0.01	-0.02	-2.20*	-3.18	0.28	0.35
	(0.01)	(0.05)	(0.08)	(1.22)	(2.77)	(0.26)	(0.29)
DSVI: Law and government	0.00	-0.07*	0.08	1.76	3.88	-0.22	-0.28
	(0.01)	(0.04)	(0.08)	(1.08)	(2.38)	(0.31)	(0.33)
DSVI: Health	0.01	-0.03	-0.05	-0.83	-1.94	0.15	-0.21
	(0.01)	(0.04)	(0.05)	(0.74)	(1.46)	(0.17)	(0.18)
DSVI: Travel	0.00	0.09	-0.01	-0.01	0.40	-0.06	-0.07
	(0.01)	(0.06)	(0.07)	(1.04)	(1.27)	(0.17)	(0.18)
Control variables included	YES	YES	YES	YES	YES	YES	YES
Observations	592	581	532	590	617	435	504
Number of countries	52	52	51	53	54	49	49
Adjusted R-squared	0.966	0.828	0.728	0.319	0.316	0.462	0.343
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES
Excluding periods with jumps	NO	NO	NO	NO	NO	NO	NO

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. LIDCs: low-income developing countries; SVI: search volume index.

Table B.13: Summary statistics for EMEs

Variable	Percentiles 25th	50th	75th	Mean	SD	Within SD	# of obs.	# of countries	# of obs. per country
Google search volume index (SVI): All	-0.45	0.84	1.91	0.63	1.73	0.23	1,222	94	13.0
SVI: Finance	-1.64	0.07	1.20	-0.25	2.00	0.32	1,222	94	13.0
SVI: Business and industrial	-1.29	0.11	1.66	-0.01	2.01	0.23	1,222	94	13.0
SVI: Law and government	-0.65	1.06	2.21	0.75	1.84	0.23	1,222	94	13.0
SVI: Health	-0.67	0.70	1.90	0.57	1.80	0.26	1,222	94	13.0
SVI: Travel	-0.45	0.90	2.09	0.78	1.64	0.22	1,222	94	13.0
Domestic SVI: All	3.86	4.13	4.32	4.03	0.44	0.22	1,204	94	12.8
Domestic SVI: Finance	3.09	3.70	4.07	3.47	0.81	0.37	1,171	94	12.5
Domestic SVI: Business and industrial	3.29	3.74	4.07	3.53	0.80	0.33	1,186	94	12.6
Domestic SVI: Law and government	3.26	3.68	4.01	3.50	0.73	0.35	1,197	94	12.7
Domestic SVI: Health	2.98	3.76	4.04	3.45	0.82	0.30	1,171	94	12.5
Domestic SVI: Travel	3.21	3.72	4.05	3.52	0.76	0.33	1,199	94	12.8
Foreign direct investment (FDI) inflows	5.49	7.05	8.41	6.91	2.21	0.74	1,049	89	11.8
Non-FDI private capital inflows	5.26	6.98	8.53	6.86	2.36	0.96	831	89	9.3
Total private capital inflows	5.95	7.70	9.07	7.45	2.36	0.88	1,000	89	11.2
Total capital inflows	5.95	7.72	9.12	7.49	2.34	0.87	994	89	11.2
Export price growth	-3.11	4.52	8.11	1.71	8.07	8.06	1,222	94	13.0
Import price growth	-2.94	4.47	8.46	1.97	8.30	8.30	1,222	94	13.0
Trading partners growth	2.22	3.29	4.61	3.31	2.24	1.86	1,222	94	13.0
REER (log level)	4.54	4.60	4.64	4.59	0.16	0.14	1,188	92	12.9
REER (percent change, p.a.)	-2.04	0.66	4.21	1.19	7.62	7.07	1,184	92	12.9
Fiscal spending	1.58	4.36	7.09	4.32	3.90	0.46	1,210	94	12.9
Inflation	2.03	4.06	7.15	5.77	10.06	8.13	1,211	94	12.9
Nominal exchange rate	-1.01	0.00	4.16	2.17	10.19	9.09	1,215	94	12.9
Real exports	1.17	2.67	3.76	2.46	1.99	0.22	1,057	87	12.1
Trade openness	4.17	4.48	4.72	4.43	0.42	0.12	1,131	93	12.2
Age dependency ratio	45.04	50.90	59.47	53.22	13.21	3.19	1,144	88	13.0
GDP (constant 2010 US$)	1.64	3.54	5.08	3.27	2.42	0.16	1,194	92	13.0
International tourism, number of arrivals	12.67	14.07	15.44	13.95	1.94	0.36	1,134	92	12.3
International tourism, receipts	19.64	20.92	22.20	20.80	1.88	0.42	1,130	91	12.4
Internet users (per 100 people)	2.53	3.32	3.81	3.09	0.94	0.61	1,091	93	11.7
Population	-0.40	1.49	3.28	1.35	2.51	0.08	1,216	94	12.9
Capital account openness index	0.17	0.45	0.88	0.51	0.35	0.10	1,026	86	11.9
Nighttime lights per area, HSW (2012)	-0.30	0.64	1.26	0.49	1.40	0.11	449	90	5.0
Nighttime lights per area, **Rnightlights**	0.13	1.12	1.87	0.94	1.45	0.36	984	82	12.0

Sources: Chinn and Ito (2006); Earth Observation Group; Financial Flows Analytics (IMF, 2018a); GADM (2018); Google Trends; Henderson et al. (2012); International Financial Statistics (IMF, 2018b); National Geophysical Data Center (with U.S. Air Force Weather Agency); World Development Indicators (World Bank, 2018); World Economic Outlook (IMF, 2018e); and the authors' calculation.

Note. Sample period: 2004-2016. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. Among EMEs, the nighttime light data exclude countries identified as outliers by Henderson et al. (2012, footnote 16, p. 1011; Bahrain, Equatorial Guinea, Serbia, Montenegro). For the data compiled by **Rnightlights**, several large economies are also excluded due to their heavy computational burden (Brazil, Chile, China, Indonesia, India, Mexico, Peru, Russia). EMEs: emerging market economies; REER: real effective exchange rate; SD: standard deviation.

Table B.14: Regression Results for EMEs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	Real GDP	Real exports	Tourist arrivals	Inflation	Nominal exchange rate	Private capital inflows	FDI inflows
SVI: Finance	0.01	0.01	0.03	1.13	3.71**	0.25	0.22
	(0.01)	(0.02)	(0.05)	(1.43)	(1.72)	(0.16)	(0.14)
SVI: Business and industrial	-0.02*	-0.01	-0.01	4.65	3.23	-0.30	-0.20
	(0.01)	(0.03)	(0.06)	(4.59)	(3.63)	(0.21)	(0.22)
SVI: Law and government	-0.00	-0.02	-0.08*	-3.75*	-3.96	0.13	0.10
	(0.01)	(0.02)	(0.05)	(2.12)	(2.38)	(0.27)	(0.22)
SVI: Health	-0.01	-0.03	-0.02	-1.12	-2.26	-0.15	0.02
	(0.01)	(0.03)	(0.05)	(1.99)	(2.38)	(0.29)	(0.21)
SVI: Travel	0.01	0.04	0.20***	-1.36	1.44	-0.04	-0.12
	(0.01)	(0.03)	(0.05)	(1.31)	(2.23)	(0.21)	(0.21)
Lagged dependent variable	0.91***	0.80***	0.75***	0.43***	0.29***	0.29***	0.36***
	(0.03)	(0.03)	(0.03)	(0.09)	(0.04)	(0.07)	(0.10)
Population (lag)	0.00	-0.01	-0.37*	-3.41	-9.89**	0.10	-1.03
	(0.03)	(0.08)	(0.21)	(3.37)	(4.09)	(1.10)	(0.90)
Internet users (lag)	0.00	0.00	-0.00	-0.20	1.05	0.06	0.01
	(0.01)	(0.01)	(0.02)	(0.52)	(0.99)	(0.13)	(0.11)
Real GDP (lag)		-0.03	0.29**	-2.62	-2.22	1.53***	1.37***
		(0.07)	(0.13)	(4.56)	(4.15)	(0.51)	(0.39)
Trade openness (lag)	0.01	0.00	0.10	2.67	-11.02***	1.04**	0.52**
	(0.02)	(0.04)	(0.08)	(2.19)	(3.13)	(0.43)	(0.25)
Fiscal spending (lag)	-0.02	-0.04	0.03	4.08	2.50	-0.13	-0.08
	(0.01)	(0.03)	(0.06)	(3.70)	(2.79)	(0.17)	(0.14)
REER, log level (lag)	-0.05***	-0.05	-0.09			0.59*	0.16
	(0.02)	(0.05)	(0.11)			(0.33)	(0.33)
REER, percent change (lag)				-0.02			
				(0.08)			
Inflation (lag)	-0.00***	-0.00**	0.01		-0.09	-0.01	-0.00
	(0.00)	(0.00)	(0.00)		(0.09)	(0.01)	(0.01)
Trading partners' growth (lag)	-0.00	-0.00	0.01	0.02	0.34	0.07**	0.04*
	(0.00)	(0.00)	(0.01)	(0.16)	(0.28)	(0.03)	(0.02)
Export price growth (lag)	0.00***	0.00	-0.00			0.03*	0.01
	(0.00)	(0.00)	(0.00)			(0.02)	(0.01)
Import price growth (lag)				0.09	0.16		
				(0.11)	(0.15)		
Capital account openness (lag)	-0.02	0.01	-0.03	-3.10**	0.32	-0.88***	-0.94***
	(0.01)	(0.03)	(0.05)	(1.33)	(2.40)	(0.29)	(0.28)
Age dependency ratio (lag)	-0.00	-0.00**	-0.00	-0.07	0.18	0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.07)	(0.12)	(0.01)	(0.01)
Observations	976	938	923	975	978	753	826
Number of countries	80	77	79	80	80	75	75
Adjusted R-squared	0.966	0.829	0.790	0.271	0.329	0.285	0.316
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES	YES	YES
Excluding periods of jumps	NO	NO	NO	NO	NO	NO	NO

Sources: Chinn and Ito (2006), Financial Flows Analytics (IMF, 2018a), Google Trends, International Financial Statistics (IMF, 2018b), World Development Indicators (World Bank, 2018), World Economic Outlook (IMF, 2018e), and the authors' estimation.

Note. Sample period: 2004-2016. Cluster-robust standard errors are reported in parentheses. Superscripts *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively. See Table B.1 for country groupings and Table B.3 for variable definitions (most of variables are in natural logarithm or percent change) and data sources. EMEs: emerging market economies; REER: real effective exchange rate; SVI: search volume index.