

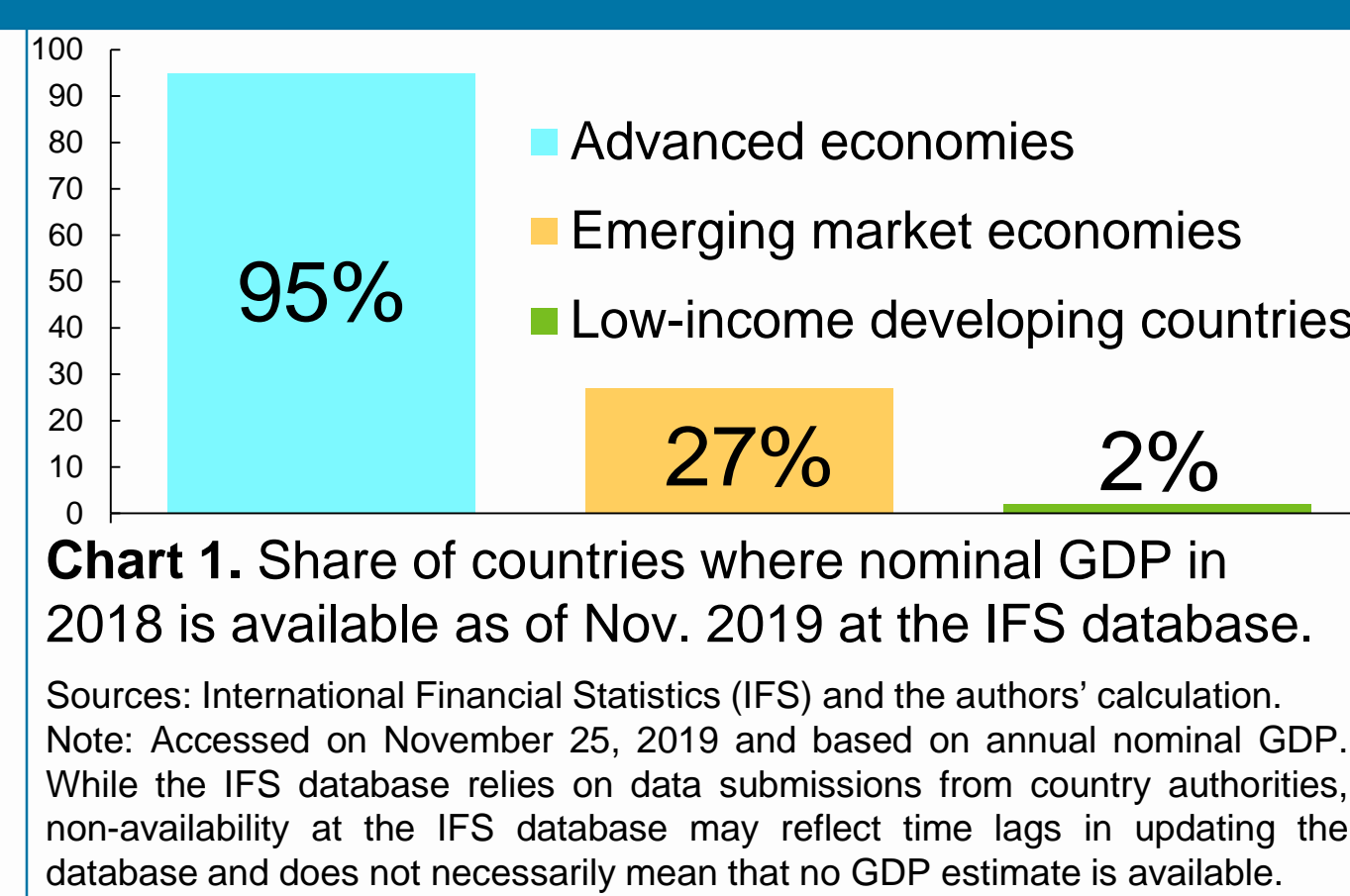
*Disclaimer:* This paper is part of a research project on macroeconomic policy in low-income countries (IATI Identifier: GB-1-202960) supported by the U.K.'s Department for International Development (DFID). The views expressed in this paper are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, IMF management, or the DFID. We thank the Google Trends support team for their assistance. <sup>†</sup> Ms. Yin is a former IMF employee.

## 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.

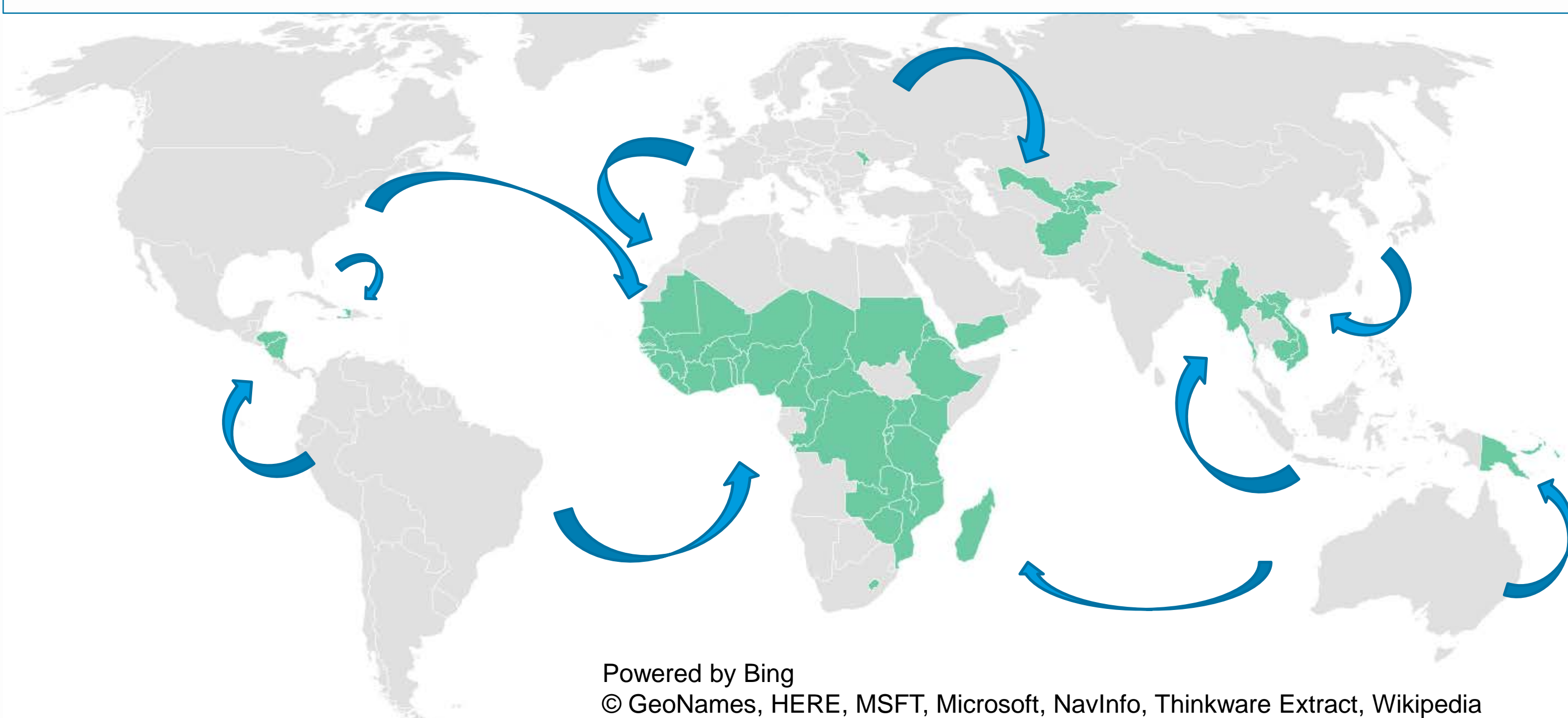
## Motivation

- **Macroeconomic data for low-income countries are available with time lags**, posing a challenge in real-time economic assessment (Chart 1).
- Can unconventional data such as **online search volume** (Choi and Varian, 2012) help address this issue?



## Online search about a country

- Use **Google Trends** to quantify **online searches about a country**.
- Leverage "big data" generated in higher-income countries to extract useful information about low-income countries (Chart 2).



**Chart 2.** Online searches about low-income countries can be generated elsewhere.

## Proposition 1. Search $\approx$ Attention

Google's **search volume index (SVI)** about a query is **proportionate** to the population share of **those who are interested** in the object represented by the query:

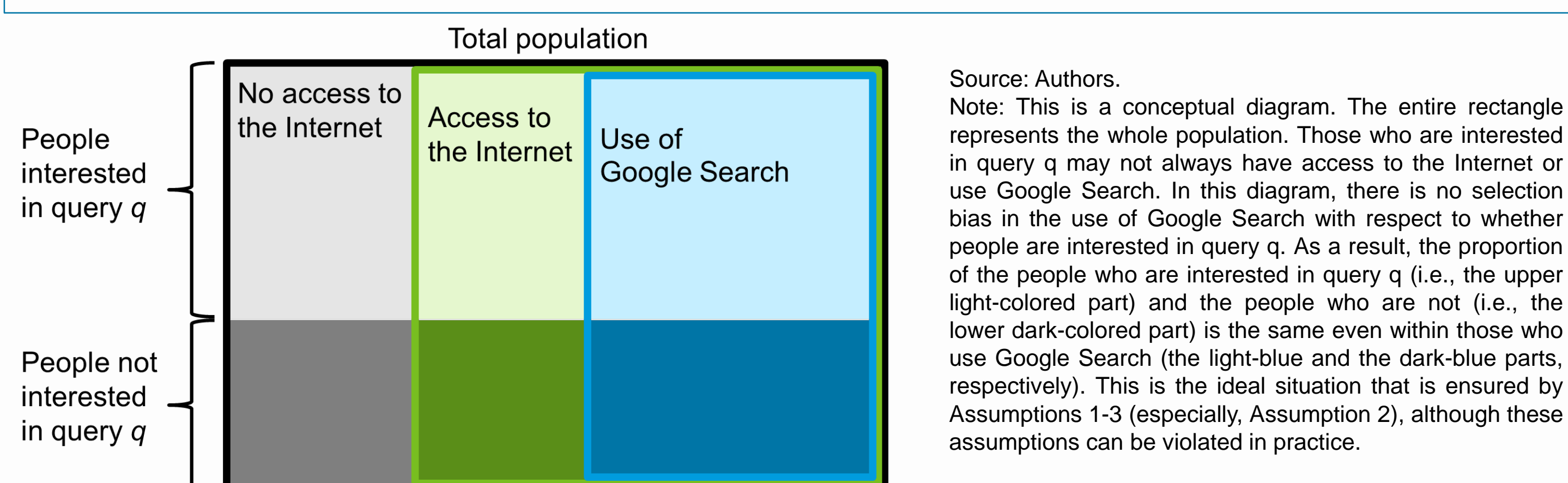
$$SVI_{t,l}(q) = Constant \times \frac{N_{t,l}(q)}{Population_{t,l}}$$

← Number of people interested in query q in location l at time t

if, across all queries, the following **three assumptions** hold (Chart 3):

1. Same average number of Google searches per person
2. Random use of Google Search (and access to the Internet)
3. Constant average queries per person

Remark: Proposition 1 contributes to the literature by **formalizing the use of SVIs to proxy people's attention** (e.g., Da, Engelberg, and Gao, 2011) and sets a basis to discuss possible biases that could arise in such analyses.



**Chart 3.** Ideal case: use of Google Search as a random sampling to infer people's attention.

## Macroeconomic nowcasting using online search indexes

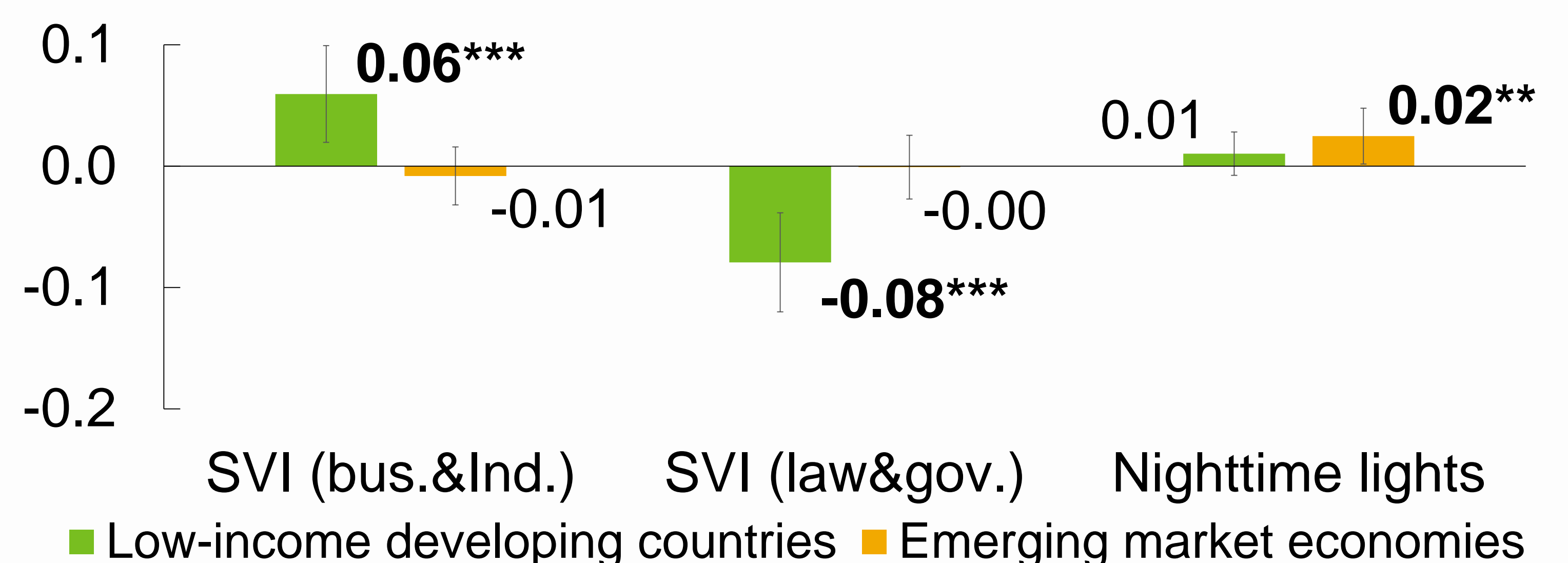
Examine how useful the **online search volume indexes** are in **nowcasting** macroeconomic variables in **low-income countries**.

$$Y_{it} = \beta SVI_{it} + \rho Y_{it-1} + \gamma X_{it-1} + \alpha_i + D_t + \varepsilon_{it} \quad (1)$$

- $Y_{it}$  = Macroeconomic variables to nowcast (e.g., real GDP, etc.)
- $SVI_{it}$  = indexes based on **online searches about country i** under **five categories** (finance, business & industry, law & government, health, travel)
- $X_{it-1}$  = control variables **with one-year lag** (e.g., population, Internet users), **reflecting time lags** in data availability
- $\alpha_i$  = country fixed effects;  $D_t$  = year effects;  $\varepsilon_{it}$  = error;  $\beta, \rho, \gamma$  = parameters

## Results – key highlights

- **Online search indexes about a country are more useful than nighttime lights** (Henderson, Storeygard, and Weil, 2012) in nowcasting real GDP for **low-income countries** (Chart 4).
- In contrast, **the opposite** is found for **emerging market economies**.
- Some indexes capture positive effects (e.g., business & industry) and others capture negative effects (e.g., law & government, health).
- Online search indexes improve out-of-sample performance, albeit to a small extent (a 2.6 percent reduction in the mean squared error).



**Chart 4.** Online searches vs. nighttime lights.

Sources: see the draft (Table 6) uploaded at the 2020 AEA conference website.  
Note: The chart shows the estimated coefficients with the 95 percent confidence intervals (\*\*\*) and \*\* indicate 1 percent and 5 percent significance, respectively) of equation (1) for real GDP, where nighttime lights (for time t and t-1) are also included as regressors. The sample period is from 2004 to 2016, excluding 2014. SVI = search volume index.

## Discussions

- The major source of information seems to be **attention from foreign locations** because online searches made domestically (e.g., searches about Malawi made in Malawi) do not generally change the results.
- Online search indexes generally **work better for low-income countries** than emerging market economies.
- **Other estimation methods** (Bayesian Model Averaging, LASSO, Ridge, Elastic-Net, Random Forests) lead to **broadly similar results**.
- **Lagged** online search indexes do not perform well, likely because the lag length of one year could be too long.
- **Jumps** (or positive outliers) in online search indexes could be used to form dummies for associated **critical events** (e.g., natural disasters).
- Online search indexes are also **correlated with other macroeconomic variables**. For example, online searches under the **finance** category are associated with **currency depreciation** and high inflation.

## Conclusions

- This paper demonstrates the **usefulness of the information** contained in Google's **online search indexes** in macroeconomic nowcasting, particularly for **low-income countries**.
- The **assumptions** required in Proposition 1 provide useful **guidance** on when **online search indexes** represent people's attention without bias.
- The **contrasting results** between low-income countries versus emerging market economies imply **some structural differences** between these groups of economies. Further investigation would be interesting.

## References

- (see the uploaded draft paper for the full list)
1. Choi, H., Varian, H., 2012. Predicting the present with Google Trends. *Economic Record*, 88, 2-9.
  2. Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. *The Journal of Finance*, 66, 1461-1499.
  3. Henderson, J. V., Storeygard, A., Weil, D. N., 2012. Measuring economic growth from outer space. *American Economic Review*, 102, 994-1028.