

An Innovative Approach

Using Online Search Data to Predict Migration

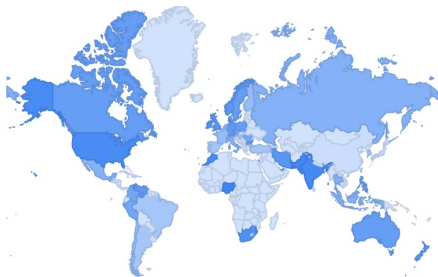
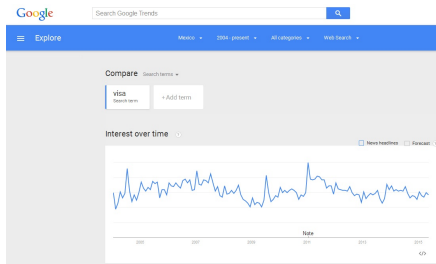
André Gröger

UAB, BGSE & MOVE

EMN conference, 29/09/20

- **Data requirements** for migration prediction: (some) flows and predictors (i.e. push- vs. pull factors)
- **Problem:** lack of migration data (outcome variable) and reliable predictors to train prediction model
- **Recent work:** use geo-located **digital trace data** (here: *Google Trends Index*) to predict bilateral migration flows

Recent work: why Google Trends?



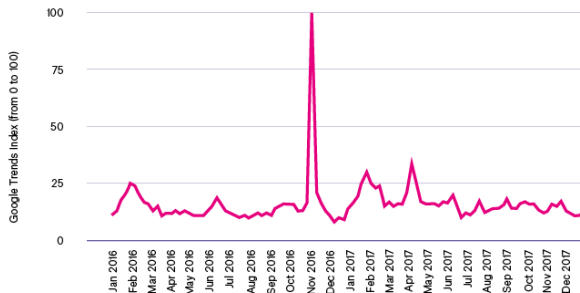
- Most common search engine worldwide (market share: 80%; 97% on mobile devices)
- Index reflects **daily** search intensities through Google for a given **keyword** and **geographical area**

Recent work: assumptions

- Migrants search for information **online** prior to departure

The election of Trump triggered Google searches in Mexico about migrating to the United States

Evolution of search intensity for the terms *immigration* and *USA* in Mexico between 2016 and 2017



Source: Compiled by the author based on data from Google Trends



ELSEVIER

Contents lists available at [ScienceDirect](#)

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec



Searching for a better life: Predicting international migration with online search keywords[☆]

Marcus H. Böhme^{a,1}, André Gröger^{b,c,*}, Tobias Stöhr^{d,e}

^a German Federal Ministry of Finance, Germany

^b Universitat Autònoma de Barcelona (UAB), Spain

^c Barcelona Graduate School of Economics (BGSE), Spain

^d Kiel Institute for the World Economy (IfW), Germany

^e IZA, Germany

ARTICLE INFO

JEL classification:

F22

C82

J61

O15

Keywords:

International migration

Migration intention

Google trends

ABSTRACT

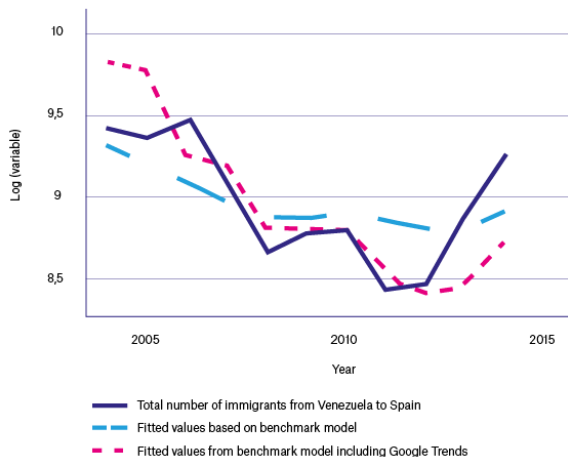
Migration data remains scarce, particularly in the context of developing countries. We demonstrate how geo-referenced online search data can be used to measure migration intentions in origin countries and to predict bilateral migration flows. Our approach provides strong additional predictive power for international migration flows when compared to reference models from the migration and trade literature. We provide evidence, based on survey data, that our measures partly reflect genuine migration intentions and that they outperform any of the established predictors of migration flows in terms of predictive power, especially in the bilateral within dimension. Our findings contribute to the literature by (1) providing a novel way for the measurement of migration intentions, (2) allowing real-time predictions of current migration flows ahead of official statistics, and (3) improving the performance of conventional models of migration flows.

Recent work: summary findings

- **Result:** strong predictive performance compared to (gravity) benchmark models
- **Interpretation:** proof of concept, not practical forecasting implementation (yet)
- **Promise:** measure migration intentions and allow [nowcasting](#) and short-term forecasts, e.g. during humanitarian crises

Recent work: example VEN-ESP

Evolution of immigration from Venezuela to Spain and predictions based on a classical migration flow model as well as an augmented one including the Google Trends Index

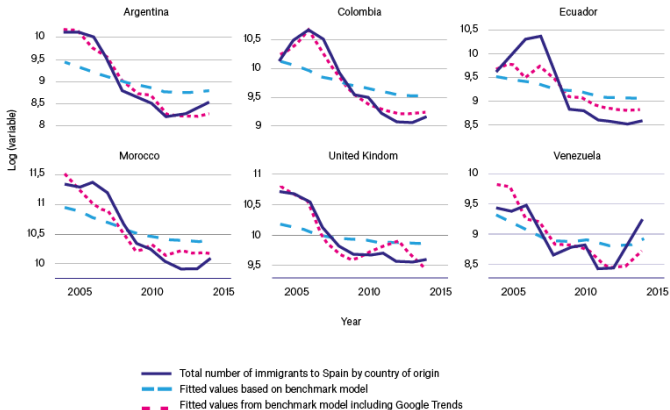


Source: Compiled by the author based on data from Google Trends, World Bank World Development Indicators, and OECD International Migration Database.

Recent work: multiple origins-ESP

Google searches help predict migration flows more accurately

Evolution of the arrival of migrants in Spain and predictions based on the keywords searched for in each country of origin

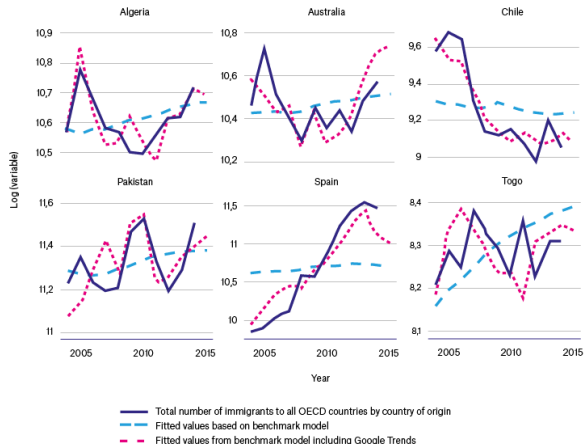


Source: Compiled by the author based on data from Google Trends, World Bank World Development Indicators, and OECD International Migration Database.

Recent work: multiple origins-OECD

Towards a tool for global migration flow prediction

Evolution and prediction of migration from six countries of origin towards the 35 states of the OECD



Source: Compiled by the author based on data from Google Trends, World Bank World Development Indicators, and OECD International Migration Database.



Horizon 2020

- Collaborative project with 14 European nodes (2020-23) lead by Universitat Autònoma de Barcelona (UAB)
- Practical implementation of migration forecasts for the EU, including **digital trace data** (incl. Google Trends, Twitter, etc.)
- Policy objective: improve management capacity of migration flows during reception, relocation, and settlement

- **Data limitations:** *availability* and *frequency* of training data
- **Keyword selection:** semantic vs. crowd-sourcing or alternatives
- **Time horizon:** nowcasting vs. forecasting

Thank you for your attention!

andre.groeger@uab.es

<https://sites.google.com/site/andregroeger/>

www.itflows.eu

Twitter: @ITFLOWS1