



How well can we estimate immigration trends using Google data?

Philippe Wanner¹

Accepted: 25 September 2020 / Published online: 8 October 2020
© The Author(s) 2020

Abstract

For a country to efficiently monitor international migration, quick access to information on migration flows is helpful. However, traditional data sources fail to provide immediate information on migration flows and do not facilitate the correct anticipation of these flows in the short term. To tackle this issue, this paper evaluates the predictive capacity of big data to estimate the current level or to predict short-term flows. The results show that Google Trends can provide information that reflects the attractiveness of Switzerland for immigrants from different countries and predict, to some extent, current and future (short-term) migration flows of adults arriving from Spain or Italy. However, the predictions appear not to be satisfactory for other flows (from France and Germany). Additional studies based on alternative approaches are needed to validate or overturn our study results.

Keywords Big data · Forecasting · Immigration trends · Migration estimates · Switzerland

1 Introduction

The emergence of alternative data derived from the operations of internet companies or communication providers as well as individual data (or microdata) from scientific collections has introduced a new era of quantitative research in the social sciences. Researchers have progressively gained access to a large amount of data, allowing them to think outside the box, transcend traditional approaches and develop new tools. In demography, the issue of data availability refers to two main areas. First, individual data (or microdata) gathered using scientific standards are increasingly available (for instance, through the IPUMS project; see Ruggles 2014), which has facilitated the rapid development of (spatial or temporal) comparative studies. Second, researchers have begun to use other non-traditional or alternative data (mobile phone records, social media, satellite maps, internet-based platforms, also commonly called “big data”, IOM 2015), particularly to understand migration and mobility in light of new methodological approaches. However, scepticism and uncertainty remain regarding the feasibility of using alternative data

✉ Philippe Wanner
Philippe.Wanner@unige.ch

¹ Institute of Demography and Socioeconomics, University of Geneva, Pont d'Arve 40, 1211 Genève 4, Switzerland

sources to measure migration-related dimensions (Rango and Vespe 2017), and more research is necessary to test the value of using big data for such research. This paper is an attempt to measure the extent to which internet activities can predict people's intentions to migrate and, consequently, future migration trends. In contrast to most comparable studies that address long-distance flows, this study focuses on migration flows from industrialized countries close to Switzerland.

To efficiently monitor migration flows, defined as the arrival in the country of persons of a foreign origin, it is helpful to have quick access to information on the trends and characteristics of current flows (Willekens et al. 2016). From an economic perspective, it is also important to be informed about the level of attractiveness of a national labour market among potential migrants abroad. Unfortunately, traditional data sources, which are based either on surveys or registers, generally fail to quickly provide statistical information on migration flows and do not facilitate correct anticipation of these flows in the short term (Wladyka 2017).

This well-known limitation is one of the reasons underlying the development of new methods based on alternative sources, which are generally combined under the common heading of big data (Facebook users—Zagheni et al. 2017; geolocalized e-mail addresses—Zagheni and Weber 2012; State et al. 2013; tweets—Zagheni et al. 2014; Hawelka et al. 2014; LinkedIn profiles—State et al. 2014). However, according to the UN inventory of big data (United Nations 2018), such applications remain unusual with regard to international migration. Until now, most case studies based on big data dealing with mobility have focused on the description of (daily) mobility. In particular, CDRs—call detail records from mobile phones – can be used to map the mobility of, for instance, commuters or transnational migrants (see Chen et al. 2016 for the discussion of the advantages and limitations of such data). Other recent studies have estimated the potential of such data to monitor undocumented and asylum-related flows (Connor 2017), with promising results. However, new experiments are necessary to assess the possibility of using big data in the field of migration.

In Switzerland, a country with approximately 30% of its population being foreign-born and 25% being of foreign citizenship at the end of 2018¹, migration depends heavily on the economy and the needs of the labour market. Migration is partially managed by different quota systems for non-EU/EFTA nationals, who represent 40% of migrants (D'Amato et al. 2019). At the same time, EU/EFTA citizens are subject to the regime of the free movement of persons and can easily move to Switzerland. According to the Swiss government, “this is conditional, however, on possession by the individuals concerned of a valid employment contract, being self-employed, or in the case of their not being in gainful employment, proof of financial independence and full health insurance coverage”².

The size of the quotas for non-EU/EFTA citizens is decided at the end of each year for the next year. To make appropriate decisions regarding size, it is important for the administration to have information not only on the economic outlook but also on recent flows and the availability of potential migrant workers within EU/EFTA countries. Even though administrative registers are able to rapidly provide data on flows and stocks (after approximately 3 months), additional information on the attractiveness of Switzerland for potential migrants and on short-term trends is useful, if not necessary. In this context, we

¹ <https://www.bfs.admin.ch/bfs/fr/home/statistiques/population/migration-integration/selon-lieu-naissance.html>. Consulted on November 13, 2019.

² <https://www.eda.admin.ch/missions/mission-eu-brussels/en/home/key-issues/free-movement-persons.html>. Consulted on April 15, 2019.

developed one very simple indicator to assess the level of attractiveness of the Swiss labour market and short-term migration trends. This indicator uses Google search statistics available through the Google Trends webpage, a tool that allows us to follow the number of searches conducted every week or every month.

We aim first to estimate the extent to which the number of searches performed on Google can provide information that measures one country's attractiveness for specific groups of potential migrants defined by country of origin. The ability of Google searches to measure attractiveness is studied by analysing the number of searches made by the inhabitants of one country of departure concerning potential countries of destination and by comparing this number with migration flows.

Second, we aim to test the capacity of our approach to predict short-term immigration to Switzerland from surrounding countries. That is, we measure the predictive capacity of Google Trends in estimating the current level of ("nowcasting", cf. Giannone et al. 2008) and short-term trends in immigration (short-term "forecasting").

Finally, we discuss the limits of the approach and the extent to which this tool is useful for a country of immigration to document the interest of potential migrants abroad (attractiveness) and update migration policies using the most current information available.

2 State of the art

Human migration is a global phenomenon that challenges modern nations and has a wide range of implications for the economy, social structures and politics. It is a phenomenon that is difficult to measure due to the increased potential for mobility among human populations, the demise of national borders within Europe and the lack of efficient statistical systems in numerous countries. Traditional statistics, based on registers, censuses or surveys, fail to measure the diversity and complexity of international flows. The increased use of administrative registers offers potential for improvement but faces issues relating to the non-exhaustive tracking of international movements.

At the same time, increased interest has been observed among official statistics producers regarding the potential of alternative data. In March 2018, the Economic Social Council of the United Nations released a report for statistics producers from the UN Global Working Group on Big Data (United Nations 2018), which was established in 2014. One of the proposals of this group was to create a knowledge platform to share experiences and best practices. During the last 10 years, some statistical administrations in industrialized countries have developed plans to implement big data applications as part of their activities. Currently, however, most activities involving the use of big data to document social and economic dimensions remain the purview of universities and research scientists, particularly in the fields of economics and medicine. Internal mobility and commuting have also been extensively examined using big data sources such as CDRs and mobile phone tracking (see, for instance, Blumenstock 2012; Gonzalez et al. 2008). These data are better adapted for mobility rather than long-term migration, but some applications have been developed in the field of transnational behaviours among migrants (i.e., persons sharing their time between two or more countries, cf. Ahas et al. 2018).

The emergence of big data as inputs to statistical tools occurred in a period characterized by great changes in the way people engage in migration. With the development of technologies and means of communication, international migration cannot be viewed as a unique move to a new country followed by a settlement or permanent stay until re-emigration. Currently, migration must be understood as a complex phenomenon that takes

multiple forms. Some migrants settle permanently, some spend time in two or more different countries, and others organize their migration trajectory around different stages, crossing different countries in circular (Constant and Zimmermann 2011), stepwise (Paul 2011) or serial migration (Ossman 2004) patterns. In this context, migration becomes a specific form of mobility (the so-called migration-mobility nexus, D'Amato et al. 2019), a wider concept that refers not only to persons but also to goods, knowledge or money. Based on this new paradigm, the measurement of movement is an issue, and alternative methods are welcomed.

It has been demonstrated that big data can provide useful information regarding migration stocks (or foreign-born populations) not only in a context of poor statistical infrastructure, in other words, in the absence of surveys, censuses or registers, but also when mobility and migration increase (IOM 2018; Zagheni et al., 2017). Among the advantages of big data use, one can mention the immediate availability and low cost of these data. This availability and easy access have led, for instance, to the generation of estimates of the size of migrant populations throughout the world using the Facebook Adverts Manager platform (Zagheni et al. 2017, see also Spyrtatos et al. 2018, for 17 European countries). Based on the number of Facebook users classified as “expats” (i.e., foreigners living in a country other than the country of origin), the two studies estimated the migrant population in different countries. The authors first had to consider the penetration rates of this social network by age group and gender for every country of origin and destination; once the fluctuation in penetration rates was taken into account, estimates of migrants could be easily generated. Zagheni et al. 2017 observed a relatively high level of coherence between their results and those derived from “traditional” data. The authors conceded, however, that numerous methodological issues may have resulted in biases, particularly in their comparisons of the share or number of migrants moving from one country to another (with different penetration rates and different behaviours concerning social networks). The same social media was also used to describe the frequency of transnational networks (Barnett and al. 2017).

Zagheni et al. 2014 also used another social networking service to estimate migration flows. Based on geolocated data for approximately 500,000 Twitter users living in OECD countries during the period from 2011 to 2013, the authors evaluated movements within and between countries for a period of four months. Using a statistical model to control for selection bias, they obtained trends in out-migration that were consistent with the results derived from traditional statistical data. The same method was used by Hawelka et al. 2014, who extracted a dataset of nearly 1 billion tweets sent in 2012. After taking into account the respective national penetration rates, they described patterns of mobility during the period under study for different countries of origin. However, the authors were unable to distinguish among different forms of mobility, particularly between seasonal and temporary mobility and migration.

State et al. 2013 used IP geolocation with Yahoo! Mail users to identify temporary and long-term migration. The authors distinguished between “mobile persons”, who used their e-mail accounts abroad for less than 3 months, and “international migrants”, who used their e-mail accounts abroad for a period of three months or more. Similarly, Ahas et al. 2018 used roaming data from mobile operators to map transnationalism from Estonia. Their approach aimed to provide a typology of migrants based on the number of phone records generated from abroad and within Estonia.

Big data, such as those generated from social networks, has also been used to measure integration-related issues. For instance, Dubois et al. 2018, mentioned in Syrtatos et al. 2018, estimated the level of assimilation of Arabic-speaking migrants in Germany based

on Facebook data. Herdagdelen et al. 2016 used the same source to describe the composition of immigrants' social networks in the United States. Studying a sample of people living in the United States and registered on Facebook who specified a hometown outside the United States in their profile, the authors examined the distribution of the subjects' friendship ties with natives, compatriots living in the United States and immigrants from other countries. Their results showed that the subjects' proportion of American friends varied from 90% among migrants from Germany, Canada, Great Britain, Australia and South Africa to 42% for migrants from China and 29% for migrants from India. The authors also demonstrated an association between the size of the group in the United States territory and compatriot affinity.

Google searches have been used less frequently to measure migration trends. Researchers from the Office for National Statistics made a first attempt to measure migration stock in the United Kingdom, a country with weak data on migration flows (Williams and Ralphs 2013)³. Traditional data referring to stocks of migrants were derived from the Labour Force Survey, but the researchers assessed the capacity of Google searches to estimate the number of migrants living in the country. Based on searches in foreign languages, they were able to estimate the number of people from Poland living in the United Kingdom from 2004 to 2010. They obtained patterns similar to those derived from the Labour Force Survey, but they were unable to replicate the results with other migrant groups, such as migrants from India. Another paper based on Google searches was published by Boehme et al. 2018), who used a combination of economic and migration-related keywords to predict the levels of migration between groups of countries, with a rather good predictive power. In contrast to other studies, Boehme et al. 2018 used an algorithm based on dozens of keywords. Wladyka 2017 attempted to estimate flows from Latin America to Spain and observed that the predictive power of Google varies from one country to another, being more powerful in Colombia and Argentina and less powerful in Peru. In 2017, the Pew Research Center released a study comparing the Google Trends data with migration data from Middle East refugees to Europe, with a high degree of correlation⁴. The UNFPA published in an exploratory study correlating the number of searches referring to the key phrase "work in Australia" in two countries (India and Italy) with the number of migrants. They observed a rather good correlation between both series, particularly for Italy ($r^2 = 0.74$).

Almost all the literature refers to intercontinental migrations. As far as we know, until now, no study was released referring to intra-continental flows.

3 Hypotheses

In this paper, we use the Google Search engine to study requests for information regarding work in Switzerland made in different countries of origin. Our hypothesis is that such searches are generally performed by possible candidates for immigration to Switzerland and reflect potential intentions. The trends regarding these searches can provide information on the attractiveness of the labour market of the destination country compared to

³ This paper was presented in the 18th GSS Methodological Symposium on June 26, 2013, but never published. A discussion of the paper is available at <https://www.unglobalpulse.org/big-data-migration>. Consulted on October 30, 2019.

⁴ <https://www.pewglobal.org/2017/06/08/online-searches-eu-refugees-methodology/>. Consulted on April 15, 2019.

other countries and help in estimating the current and short-term (1–2 years) trends regarding the level of migration.

In recent decades, academic researchers have shown a growing interest in studying intentions to migrate and their consequences (Migali and Scipioni 2018). Theories and empirical research have indicated that *intentions* to migrate internationally (either to leave one's country of origin or return to it after a stay abroad) affect *actual* migration behaviours, as concrete opportunities translate that desire into an actual decision event (Docquier et al. 2014).

Intentions to migrate can therefore serve as a predictor of future behaviour (e.g., Van Dalen and Henkens 2008; Armitage and Conner 2001; De Jong 2000; Carling and Pettersen 2014; Carling 2014; Creighton 2013). However, the theoretical proposals that link intentions to behaviours generally take into account alternative dimensions that affect the relationship between the two phenomena, such as investments in social contacts, skills (Carling and Pettersen 2014), actual opportunities to migrate, aspirations and information (de Haas 2014). The factors, constraints or incentives that influence the decision to migrate are diverse, and the number of these factors varies from one candidate to another. There is no single path from intentions to movement but a multitude of different schemes. Even if a relationship between intentions and behaviours is anticipated, this relationship is not expected to be linear, which makes it difficult to establish a universal theoretical framework. According to Steiner 2017, however, “exposure to information plays a crucial role in providing people with the necessary information and resources for their decision to stay or migrate”. This can be done by different ways, including the contact with persons who previously migrated, but also by searching information available on the web.

Statistical information on migration intentions or aspirations is very limited to specific migrant surveys, such as the Gallup World Poll (Migali and Scipioni 2018), which constrains the analysis of the role of intentions in realizing migration. The empirical relationship between intentions and actual flows has increasingly been investigated during the last decade in the context of developing countries (Van Dalen and Henkens 2008; Docquier et al. 2014; Laczko et al. 2017; Carling 2017; Tjaden et al. 2019). Carling 2017 attempted to theorize the process between the so-called “root-causes” of migration (conditions of states and prospects for improvement), the desire to change, the migration aspirations and, finally, the migration outcomes. According to this author, the link between the desire for change and migration is influenced by the migration infrastructure, which reflects the human and non-human elements that enable and shape migration.

We adapted Carling scheme by adding an intermediary step between the migration aspiration and the migration outcome based on the fact that the decision to ultimately migrate is influenced by the information gathered through the internet or using alternative sources (such as the personal network—Fig. 1). Even if the focus of this paper is the link between the search for information and actual migration, we know that the entire process is influenced by the situation in the country of origin, which can lead to the desire for change, as well as the “migration infrastructure” suggested by Carling (2017).

Due to limitations in terms of sampling, poll data, such as the Gallup World Poll, cannot be used in order to precisely describe the relationship between migration intentions and realized migration for pairs of countries. They also fail to reflect rapid changes in the intentions (Tjaden and al. 2019), which is why alternative data reflecting intentions to migrate has to be investigated. Thus, we contend that the level of Google searches for keywords or key phrases associated with migration is a marker of such intentions. In other words, by searching for specific information on immigration towards a specific country, a potential migrant will evaluate the extent to which migration towards the potential country

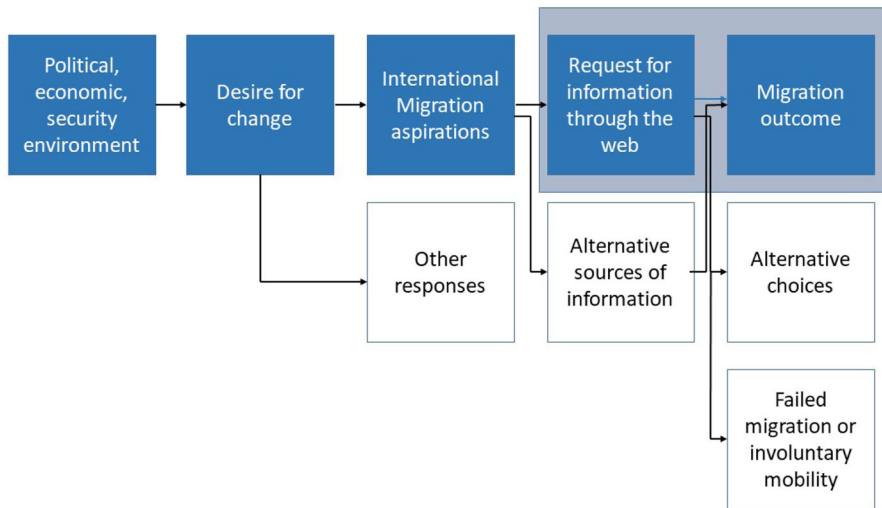


Fig. 1 Process leading to migration, adapted from Carling 2017

of destination is feasible and fits with his/her own expectations. Changes in the frequency of Google searches can therefore indicate a change in the intentions or aspirations of the population under study. The level of searches can also be seen as an indicator of the attractiveness of a country for potential migrants.

We are aware that other methods for gathering information on the living and working conditions in Switzerland are available to candidates. One such alternative is to contact friends or family members who are already living in Switzerland, with another one being simply to not look for specific information before migration. Nevertheless, we believe that Google searches represent a good method that candidates can use to gather or refine information on Switzerland. Therefore, a relationship between intentions and behaviours is anticipated, although the relationship is not expected to be linear.

In our study, we will first test the extent to which trends in Google searches predict the attractiveness of Switzerland as an immigration country. After demonstrating the accuracy of our tool for measuring the attractiveness of Switzerland, in the second part of our paper, we aim to measure whether it is possible to anticipate the level (number) of immigrations to Switzerland based on our indicator. Our hypothesis is that a decrease (or increase) in the number of Google searches will, following a delay (representing the duration between the expression of intentions to migrate to actual migration, if any), translate to a corresponding change in the number of immigrants.

To test this prediction, we must consider a period of time during which the relationship between intention and behaviour (realization of migration) remains constant. In other words, it is necessary to avoid periods characterized by changes in immigration policies or strong transformations in the social or economic contexts in the country of destination, which are difficult if not impossible to predict. This is an a priori condition that enables the exclusion of the impact of external factors that could influence the relationship to be tested. Since the beginning of the 21st century, Switzerland has been characterized by regular economic growth and a low and constant unemployment rate relative to European standards. The economic climate has been rather good in Switzerland during the last 20 years. However, based on a quarterly survey among private businesses, the KOF Swiss Economic

Institute, which is responsible for the computation of employment indicators, reported a cyclical trend regarding the employment climate from 2006 to 2012 (with a negative period from the 3rd quarter of 2008 to the 1st quarter of 2010), followed by a stable situation⁵.

Moreover, for the surrounding countries (belonging to the 'old' (EU-15) group of EU countries), the admission policy has remained exactly the same during the period from June 2007 to now, characterized by the regime of free movement of persons. From June 2002 to May 2007, these countries were basically under the same regime⁶ but were subject to some transitory measures that were implemented to avoid abuses in terms of salaries and working conditions for EU citizens. On the whole, the contextual situation of immigration to Switzerland can be considered stable for the countries that are the traditional suppliers of its immigrants (France, Germany, Italy, Spain and Portugal).

4 Methods

We used Google Trends to estimate the intensity of searches referring to Switzerland. For privacy reasons, this engine does not provide the absolute number of searches but an index ranging from 0 to 100 based on the proportion of searches with the term under study (a keyword or a key phrase). A value of 100 indicates the highest proportion of searches during the period considered. Other values indicate the ratio between the current proportion divided by the highest proportion observed during the period considered.

We identified different keywords and key phrases that can indicate intentions to migrate to Switzerland. These keywords related to the labour market and to practical aspects of migration (for instance, housing and school). We focused on France, Germany, Spain and Italy, as other European countries were often characterized by a small number of searches, and the results were not systematically displayed by Google Trends due to Google's privacy policy. The selection of countries under study was based on recent migratory flows to Switzerland, which are dominated by Italy, Germany, France, Portugal and Spain. However, the migration of Portuguese persons to Switzerland involves low-qualified migration, which differs from the migration flows from the other countries (Wanner and Steiner 2018). This type of migration is organized mainly through family and friend networks (chain migration), and the number of Google searches on the topic was too low to be analysed.

First, we defined terms that could indicate intentions to migrate to Switzerland. Strategies for selecting keywords in Google Trends generally differ according to the study purpose, varying from using one or a limited number of specific terms (Williams and Ralphs 2013; Connor 2017; Wladyka 2017) to using a large number of terms (Böhme et al. 2018 used 68 generic keywords in 68 languages). After different attempts (using key phrases such as "life in Switzerland", "migration to Switzerland", "expat(s) in Switzerland", "Swiss labour market", and "housing in Switzerland"), we selected the most popular key phrase. This phrase was the equivalent of "working in Switzerland" in the different languages considered (German, French, Italian and Spanish). This choice is

⁵ <https://www.kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-employment-indicator.html>. Consulted on December 28, 2018.

⁶ The agreement regarding the free movement of persons, which went into effect on June 1, 2002, for the countries belonging to the EU-15 and the EFTA, was extended to states that adhered to EU policy in 2004. This extension went into effect April 1, 2006, and involved Poland, the Czech Republic, Hungary, Slovakia, Slovenia, Lithuania, Latvia and Estonia. In June 2009, this agreement was extended to include Bulgaria and Romania, and Croatia was included in 2017.

similar to the one of Wladyka 2017 when estimating migration flows from Latin America to Spain. It was also chosen in the study published by the UNFPA 2014. This strategy was chosen to facilitate our focus on a specific flow (from one country to Switzerland) and ensure our results would be as accurate as possible. We found that more generic terms, such as “Switzerland” or “visiting Switzerland”, which were tested in other contexts, had limited predictive values in the Swiss context.

We observed a relationship between the intensity of research based on the phrase “working in Switzerland” and that of research based on “living in Switzerland”. Migration to Switzerland is essentially labour force migration. Seventy per cent of French migrants who arrived in Switzerland during the period from 2006–2016 had a job position or a contract at the time of migration. This was also the case for 60% of Spanish, 59% of German and 50% of Italian immigrants (Wanner and Steiner 2018). Therefore, it was not surprising that “working in Switzerland” was a frequent search in the various languages. As migration to Switzerland is mainly characterized by a desire to participate in the labour market, generally considered gainful among citizens from other countries, one could presume a relatively strong relationship between the search for information on work in Switzerland and intentions to migrate. It is probable, however, that working as a frontier worker or engaging in cross-border commuting may represent an alternative to migration for citizens from Italy, Germany and France who live close to the border.

Google was established at the end of the 1990s and rapidly became popular due to the efficiency of its search engine. In 2002, Google replaced Yahoo! as the most popular search engine. After a stable period in the mid-2000s with a market share of less than 60%, Google quickly consolidated its position with more than 70% and even 80% of market share⁷. We did not consider the years from 2004 to 2006 in order to retain only the years during which Google was in a quasi-monopolistic position. Using Google Trends, we extracted monthly indexes in each country for the following translations:

- France: Travailler en Suisse
- Italy: Lavorare in Svizzera
- Germany: Arbeiten in der Schweiz
- Spain: Trabajar en Suiza

In parallel, we obtained access to different official immigration statistics. Since 1981, international flows of foreign persons (excluding asylum seekers) have been recorded by the Swiss administration and are available through a so-called Central Aliens Register (for the period from 1981–2008), which was replaced by another register (including asylum seekers) called Symic. These registers allowed us to compute the number of monthly immigrations to Switzerland. As it was not possible to identify country of residence prior to migration, citizenship was used as a proxy. Then, monthly immigration, defined as the arrival in Switzerland of persons who have been granted a residence permit⁸, were computed for the four countries under study (France, Italy, Spain and Germany). In the case of multiple migrations during the same year (which is not uncommon for persons holding a

⁷ <https://www.distilled.net/blog/your-seo-approach-is-obsolete-evolve-in-2014/>. Google was founded in 1997, and although it achieved considerable success early in its history, it was only in 2002 that it becomes the most popular search engine.

⁸ Access to Swiss territory for nationals of these countries is guaranteed by the free movement of persons. Upon request, a residence permit is issued upon entry. This permit authorizes work and residence in Switzerland for a period of three months or more, as well as access to various social rights. Illegal migration, defined as the stay in Switzerland of persons without a permit, is considered extremely rare for EU/EFTA nationals.

short-term permit), we considered only their first arrival in Switzerland during the year under study. All types of immigrants, including short-term immigrants (from 3 months to 1 year), were taken into account, regardless of their status of stay. Monthly data were available up to December 2016. As it is assumed that children and retirees have no reasons to search for information on work in Switzerland, we retained immigration from working-age foreigners (18–64 years).

Data on emigration were also obtained from the statistical administrations of the countries under study. These data allowed us to test, using a descriptive approach, the extent to which the intensity of Google searches for potential host countries was associated with the actual number of emigrations. In this preliminary analysis, we took cross-border workers into account. For Italy, we used 2015 figures from the A.I.R.E.⁹ register, which is a unique source of information on emigration. These figures are considered to underestimate the number of Italian emigrants. In fact, by comparing information on the number of Italians emigrating to Switzerland according to the A.I.R.E register and information from the Swiss population register (immigration of Italians), we found a coverage rate of 48%, indicating that more than half of Italians did not declare their emigration. A recent study¹⁰ obtained the same figure for the United Kingdom and stated that “the number of Italians obtaining Social Security numbers in the UK last year [in 2016] was twice the number of those officially registering with the Italian authorities as living in Britain”. Thus, we acknowledge that the figures on emigration are underestimated.

For Germany, data were available through the Federal Statistical Office (DESTATIS 2017). Both Italy and Germany have cross-border employment, and the data on emigration were corrected by taking into account the number of Italians and Germans working in a border country as frontier workers. Unfortunately, data for France and Spain were not available or accessible.

The assumption of our paper is that persons who are planning to migrate to Switzerland would use the internet to access specific information on the Swiss labour market. By extension, the number of such internet searches can provide an estimate of the future migrations to Switzerland. Then, there is a time lag between the search and the migration that is unknown and that can differ according to each personal situation. For Latin American citizens aiming to immigrate to Spain, Wladyka 2017 estimated an approximately seven-month time lag between the searches on specific terms (“working in Spain”, “visa for Spain”) and actual migration. As we have no information on this time lag that can vary according to the country, we selected the one that provides the highest coefficient of determination (r^2).

To estimate our model in the simplest possible way, a linear regression was used in order to measure the relationship between the number of searches (x) and the number of moves (y) that finally occurred. The linear model ($y = ax + by$) was estimated using the ordinary least squares (OLS) method for different time lags. Table 1 provides the coefficients and the r^2 for the linear models for the period from 2007 to 2016 according to different time lags between information search and immigration. One can observe that the parameters of the equation are gradually modified according to the time lag, but the slope remains positive and increases parallel to r^2 . The model fits the best when considering a time lag of 3 months for Germany, 12 months for Italy and Spain and 24 months for France. The time lag represents the time necessary to finalize the migration movement between the expression of the intention to migrate, translated by internet searches and the

⁹ <https://www.settimananews.it/reportage-interviste/la-migrazione-italiana/>.

¹⁰ <https://www.ft.com/content/cb9bd2ee-c07d-11e7-9836-b25f8adaa111>.

Table 1 Parameters of the OLS models according to the country of origin and different time lags

2007–2016	France			Germany			Italy			Spain		
	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²
No time lag	485.1	11.0	0.23	1832.8	23.9	0.61	349.1	20.3	0.61	167.1	8.8	0.45
3 months	473.4	11.4	0.26	1797.0	24.5	0.62	326.8	21.2	0.69	146.4	9.6	0.57
6 months	468.2	11.6	0.29	1775.1	24.6	0.61	318.7	21.7	0.76	135.1	10.2	0.67
9 months	454.5	12.0	0.33	1788.1	23.7	0.57	326.0	21.8	0.80	133.4	10.5	0.73
12 months	429.1	12.7	0.38	1824.8	21.9	0.51	351.6	21.4	0.81	142.5	10.3	0.75
15 months	410.7	13.2	0.43	1884.9	19.5	0.43	394.7	20.7	0.80	161.6	9.9	0.73
18 months	390.5	13.7	0.48	1959.5	16.7	0.34	449.7	19.7	0.76	189.2	9.2	0.67
21 months	367.9	14.3	0.53	2039.2	13.7	0.26	515.1	18.4	0.70	226.1	8.2	0.58
24 months	351.2	14.8	0.58	2115.6	10.9	0.18	587.6	17.1	0.63	269.4	7.1	0.47
27 months	371.0	14.5	0.58	2146.7	9.6	0.14	666.3	15.6	0.56	318.2	5.9	0.35
30 months	418.9	13.7	0.57	2114.0	10.2	0.15	752.8	14.0	0.49	370.1	4.5	0.24

potential migration movement. This period is longer for the French population, which may be related to the fact that French workers often start their professional experience in Switzerland as cross-border workers before possibly migrating¹¹.

To minimize the impact of statistical and seasonal fluctuations, immigration flows and Google searches were smoothed using a mobile average over twelve months. However, the raw data are displayed in the first part of the results.

To test the stability of the model, we also computed OLS for different periods of time (Table 2). The comparison of the models confirms first that taking into account the 2004–2006 period makes the model perform worse, particularly for the German and French estimates, which is explained by the fact that Google was not in a quasi-monopolistic situation on that period. Second, one can observe relatively stable estimates of the slope, at least for Italy, France and Spain, regardless of the period considered.

5 Results

5.1 Number of immigrations

The monthly immigration numbers are presented in Fig. 2. The trends and the fluctuations are explained either by external factors (such as the consequences of the global financial crisis, especially in Southern European countries) or by seasonal effects related to the activities of immigrants, with an increase after the summer break and a decrease during the first months of each year. On the whole, migration to Switzerland during the period from 2004 to 2016 increased relative to earlier periods. The immigration of foreigners aged 18–64 increased from 107,000 to 224,600 between 2004 and 2013 but slightly decreased after to reach 168,700 in 2017. During the whole period under study, the four countries of origin under study represented approximately 40% of these immigrants. From 2004 to

¹¹ According to the Migration-Mobility Survey, the proportion of French immigrants living in Switzerland who worked as cross-border workers before moving to Switzerland is 17%, which is higher than that for the other groups (Italy, 9%; Germany, 14%).

Table 2 Results of the OLS models for different periods of time

	France (24 months)			Germany (3 months)			Italy (12 months)			Spain (12 months)		
	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>a</i>	<i>b</i>	<i>r</i> ²
2008–2016	471.3	12.7	0.59	1610.1	30.9	0.61	414.3	20.3	0.79	189.4	9.3	0.73
2007–2016	351.2	14.8	0.58	1797.0	24.5	0.62	351.6	21.4	0.81	142.5	10.3	0.75
2006–2016	340.5	15.4	0.48	2112.8	19.5	0.40	375.5	21.3	0.77	145.2	10.3	0.69
2005–2016	333.7	15.0	0.41	2268.8	14.3	0.26	290.7	22.5	0.75	112.5	10.8	0.69
2004–2016	484.7	11.4	0.23	2266.8	13.4	0.21	234.8	23.2	0.72	92.4	11.1	0.69

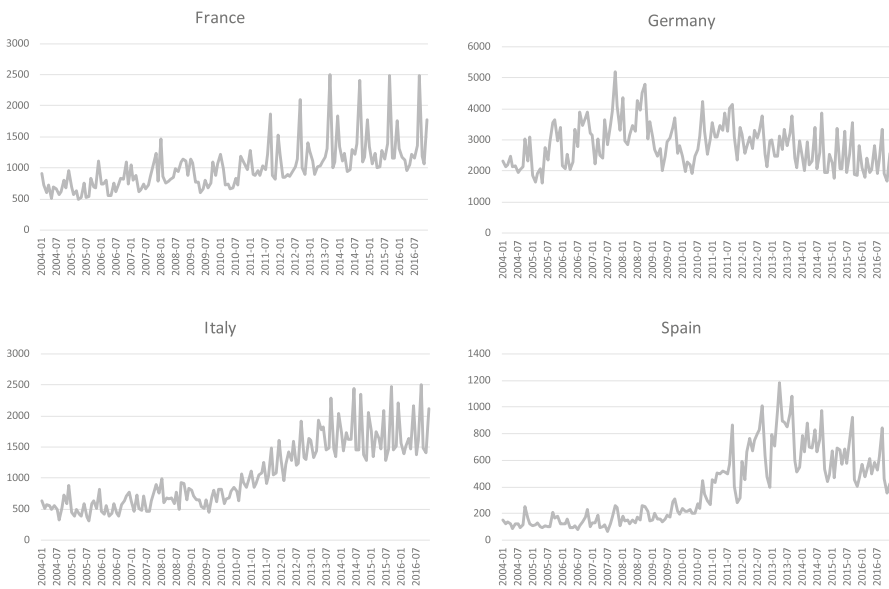


Fig. 2 Monthly immigration flows to Switzerland, 2004–2016, 18–64 years. *Source:* Swiss Federal Statistical Office. STATPOP Register. Vertical scales differ between each chart

2017, Germany, France, Italy and Spain provided an annual average of 33,300, 12,400, 13,100 and 4700 immigrants, respectively.

Figure 2 shows a slight decrease in German immigration following the end of the first decade of this century, and an increase in Italian and French flows for the period 2004–2014, followed by a stabilization, while Spanish immigration has decreased since 2014. Thus, we found different trends that we then examined based on the intensity of Google searches.

5.2 Google trends indexes

Figure 3 shows the Google Trends indexes for the 4 countries from January 2004 to July 2018. There were important variations, especially before 2007, during the period when Google searches were less successful and subject to competition from other engines. Even

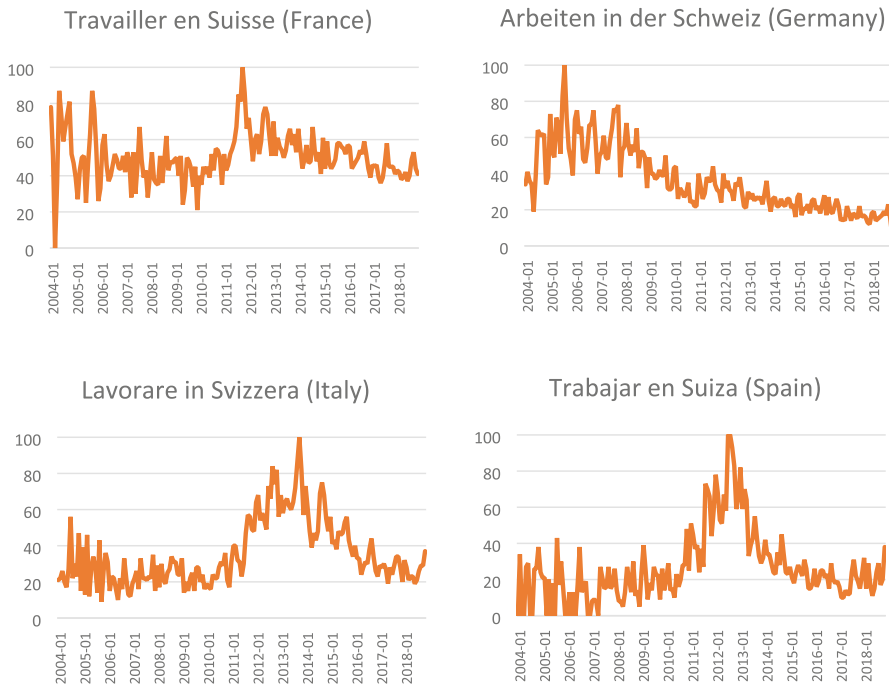


Fig. 3 Google Trends Index, 2004–2018. Source: Google Trends (www.trends.google.com)

though fluctuations were less frequent after 2007, they remained relevant, which called for a smoothing of the data. As mentioned before, we limited the analysis to searches conducted during 2007 and later.

5.3 Can the attractiveness of Switzerland for potential immigrants be measured using google trends?

To validate our approach, we assessed the extent to which the Google Trends index can concretely measure the attractiveness of Switzerland to potential migrants.

To accomplish this, we first used a very descriptive approach and compared the Google Trends index for Switzerland with indexes for other destination countries and then described the association between the index and the actual number of emigrants from those countries for the most recent year of data availability (2015). As mentioned before, due to the lack of available emigration data for Spain and France, this comparison was performed only for Germany and Italy. To take into account the fact that some Germans and Italians search for information about Switzerland with regard to cross-border activity, we added to the number of immigrants the estimated number of new contracts for cross-border workers.

In 2015, 71,435 German persons living in Germany decided to leave their country of citizenship (DESTATIS, 2017). The main European destinations were Switzerland (12,064), Austria (6,832), Spain (6,088), and the United Kingdom (6,043)¹². For Switzerland and Austria, where cross-border employment is an alternative to migration, we

¹² The United States (10,159) was the second country of destination for German people but is not considered here.

included the numbers of new authorizations of frontier workers (Switzerland, 11,006; Austria, 6,800¹³). Figure 4 presents the average monthly index of Google Trends in 2014 and 2015 compared with 2015 emigration data for the four countries. As seen in the figure, the index is more than two times higher for Switzerland than for Austria and approximately five times higher for Switzerland than for Spain and the United Kingdom. These differences translate rather well to the levels of emigration and cross-border employment from the different countries.

The same exercise was performed for Italian emigrants (Fig. 5). Emigration numbers, corrected for frontier workers, were compared to the Google Trends index for the main European countries of emigration (Switzerland, Germany, Spain and France). There was a somewhat strong relationship between the index and the number of emigrations from the countries. However, the ratio of the index to the number of emigrations was weaker for France and, to a smaller extent, Germany than for Switzerland and Spain. The index was even higher for Spain than for France, even though the number of migrants was much smaller. One explanation may be that migration in France and, to a smaller extent, Germany involved labour force migration less frequently than migration in Switzerland and Spain. This can be explained by the fact that migration to Switzerland, from the point of view of a Spanish resident, is probably more closely linked to professional activity, whereas for residents of Germany or France, other reasons (such as family reasons) may also explain the arrival in Switzerland. Neoclassical economic theories highlight the fact that large differences in wage conditions can increase flows (Arango 2000). However, according to the OECD statistics¹⁴, Spain is characterised by a lower average wage than Germany and France.

Overall, the comparison of the emigration numbers with the Google Trends index revealed a relationship between both figures. However, the data available did not allow in-depth analysis of the causality between the indicators; only descriptive analyses were possible. Based on these descriptive analyses, current evidence is not strong enough to confirm the relationship between Google Trends and the emigration numbers, and therefore other analyses are required, which is done in the next chapter.

Another aspect supporting the validity of the Google Trends index was the situation observed in approximately February of 2014. At that time, a popular vote was organized in Switzerland with the aim of strongly limiting migration. This controversial anti-immigration initiative unexpectedly won the support of the electorate. The results of this vote were widely reported in the international press, and as anticipated, the vote was followed by a decrease of 15% (Germany) to 44% (Italy) in the Google Trends index during the next 5 months compared to the 5 months preceding the vote. It is probable that Swiss citizens, by agreeing with this popular initiative, sent a negative message to candidates for migration regarding the openness of Switzerland. Therefore, we expected to observe lower interest in jobs in Switzerland, and this was reflected in the Google Trends index.

The results demonstrate that the Google Trends index is a potential indicator of the attractiveness of a country (at a global level) and probably translates to intentions to emigrate (at the individual level). For candidates for emigration, searches more frequently reference countries that are relatively attractive or accessible. We also observed that a singular event, possibly impacting the attractiveness of a country, can modify the number

¹³ We did not find any data on the number of new border workers in Austria coming from Germany, only an estimate of 34,000 border workers working in Austria. This suggests that if the rate is the same in Austria and Switzerland, 6800 new workers were registered in 2015 in Switzerland.

¹⁴ https://stats.oecd.org/Index.aspx?DataSetCode=AV_AN_WAGE. Consulted on August 15, 2020.

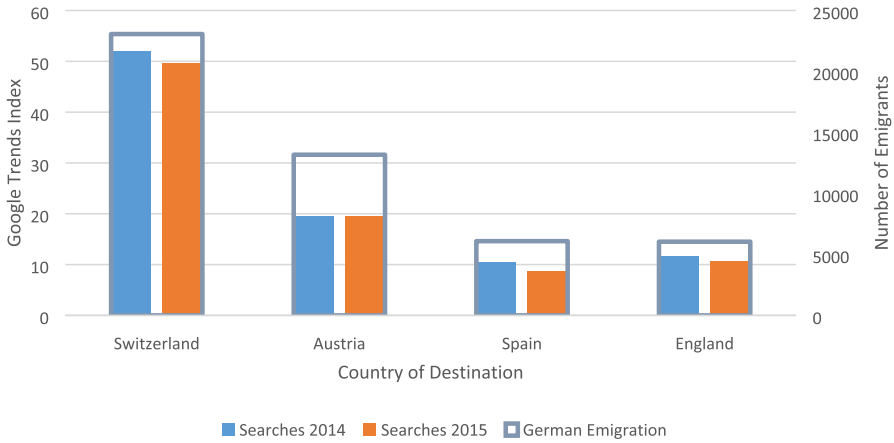


Fig. 4 Comparison between the number of German emigrants in 2015 (grey box) and the Google Trends index obtained for the phrase “Arbeiten in Switzerland/Austria/Spain /England” in 2014 and 2015. Sources: Google Trends (www.trends.google.com), DESTATIS,2017 , Swiss Federal Statistical Office (STATPOP Register) and Statistics Austria. The number of new authorizations for border workers living in Germany but working in Switzerland and Austria were added to the number of German immigrants in the two countries

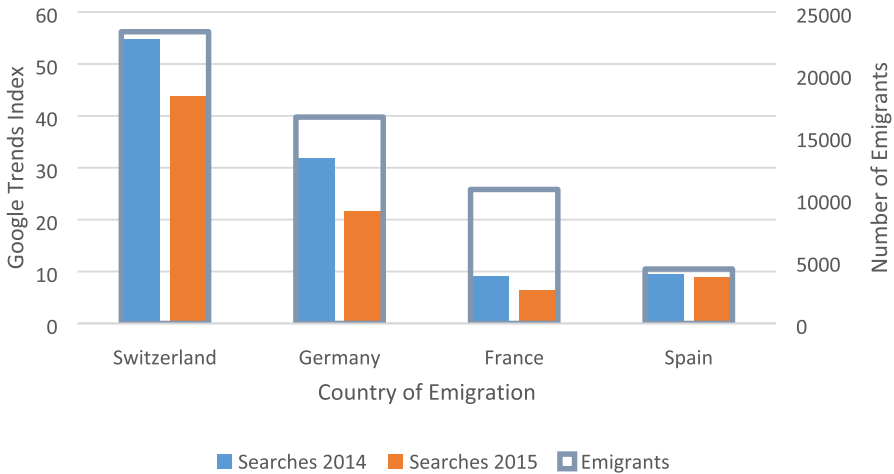


Fig. 5 Comparison between the number of Italian emigrants in 2015 (grey box) and the Google Trends index obtained for the phrase “lavorare in Switzerland/Germany/Spain/ France” in 2014 and 2015. Sources: Google Trends (www.trends.google.com), A.I.R.E, and Swiss Federal Statistical Office (STATPOP register). The number of new authorizations for border workers living in Italy but working in Switzerland was added to the number of Italian immigrants in Switzerland. The number of cross-border Italians working in France is not available but is considered to be very limited

of searches. The results provide support for an in-depth study of the relationships between the index and actual migration from a longitudinal perspective.

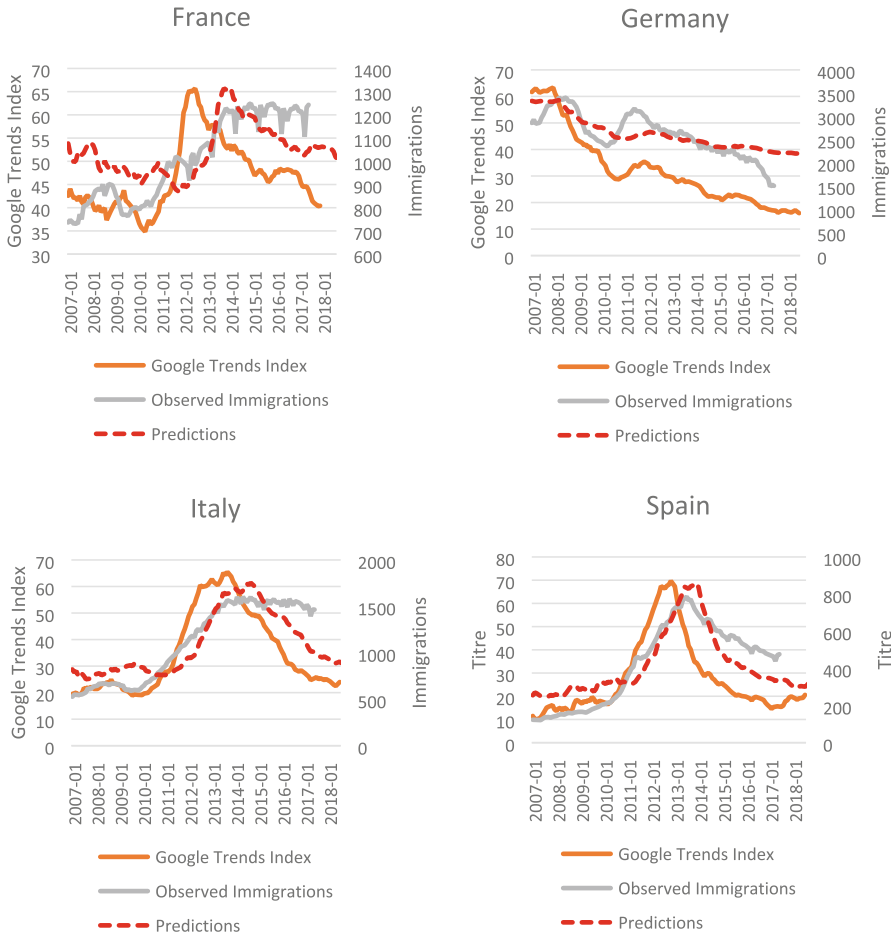


Fig. 6 Google Trends index and the number of immigrants from France, Germany, Italy and Spain. Left scale: Index of searches; Numbers are smoothed using a moving average. Forecasts are estimated using a linear regression for the period from 2006 to 2016

5.4 Google searches as a predictor of migration trends

Figure 6 plots the monthly indexes and monthly immigration flows for the four countries under study, beginning in January 2007. Both series were smoothed over 12 months in order to avoid seasonal fluctuations. The time lags that correspond to the highest r^2 , ranging from three months for Germany to 24 months for France, were taken into account. The values from July 2017 (dashed lines) were predicted and resulted from the linear model based on the level of the indexes.

Overall, the quality of the association between the searches and observed immigration was better for Southern European immigrants in terms of the r^2 value. For the French population, the predictive capacity of the index is low compared to the Southern European countries ($r^2 = 0.58$), which is explained by the decrease in the index following the end of 2011, contradicting the regular increase in the number of immigrants until 2015. There may be different explanations for these differences in the trends, including that migration

from France to Switzerland partially involves cross-border workers who are active in the Swiss labour market for some years before settling in Switzerland. According to the Swiss Migration-Mobility Survey 2016, 21% of French citizens who arrived as adults in Switzerland between 2006 and 2016 were former cross-border workers¹⁵. The proportion was lower for other groups: 11% among Italians, 10% among Germans and less than 1% among Spanish. Cross-border workers do not need to gather information on the labour force conditions in Switzerland, as they are already working in the country. Thus, the link between the trends is less explicit. Another explanation may be the social and political uncertainty in France, which has in some periods increased interest in emigration, particularly among elites and the wealthy. In this context, it is not surprising that we observed a peak in searches just after the presidential election of François Hollande, whose fiscal policy was politically oriented against those with high incomes. However, the growing interest for the Swiss labour market was not followed by a significant increase in the migration flows, even if it may be one of the reasons why the number of immigrants to Switzerland continuously grew from 2011 to 2014 and remains stable at a high level, as the Google Trends index began to diminish. Another reason for this growth could be work opportunities in the French part of Switzerland, which has a rather positive climate as measured by the expectations of employers in terms of job creation¹⁶. Overall, the Google Trends index is not an adequate indicator of future migration trends from France to Switzerland.

Compared to the situation for France, the intensity of searches with the phrase “arbeiten in der Schweiz” in Germany was a slightly better predictor of migratory flows (with a delay of 3 months between the search and immigration), with a r^2 score of 0.62. As the figure shows, the Google Trends index failed to predict some of the fluctuations observed in migration. It was unable to predict an unanticipated decrease in the number of immigrations from mid-2008 to mid-2010, which coincided with the aforementioned negative employment dynamic, or the spectacular increase in the level of immigration during the year 2011. According to the index, our model predicts a stabilization of the immigration flows, which is not in line with the most recent data available.

The search “lavorare in Svizzera” predicted the migratory flows from Italy to Switzerland rather well until the end of 2013. After 2013, the Google Trends index declined quickly, while migration flows remained approximately stable, at least until the end of 2016. The difference between both trends may be explained by the existence of work opportunities in Switzerland for two groups of Italians: highly skilled migrants in managerial positions and low- and middle-skilled migrants who are willing to accept relatively insecure jobs and low wages. The first group was probably more likely than the second group to search for information using the internet, particularly during the financial crisis that created trouble for the Italian economy. The second group is now slightly more represented in the flow of Italian immigrants (Wanner and Steiner 2018). We presume that members of the second group searched for information about the Swiss labour market on the internet less frequently than members of the first group. Overall, by considering a delay of 12 months between the searches and immigration, we observed a rather strong capacity for prediction ($r^2 = 0.82$). The decrease in the number of searches since 2014 could provide strong arguments in favour of a future decline in Italian immigration to Switzerland.

¹⁵ Own computations.

¹⁶ <https://www.kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-employment-indicator.html>.

The same profile was observed for Spain, with an important increase in the Google Trends index after 2009, which preceded a substantial increase in immigration to Switzerland. The index decreased after 2010, slightly more 12 months before a decrease in the immigration flows was observed. The predictive capacity of the index was therefore good ($r^2 = 0.75$). However, as was observed for Italy, and probably for the same reasons, the decline in the Google Trends index was more substantial than the decline in migration.

Overall, the study results were mixed but not unsatisfactory. There was a clear correspondence between the Google Trends index and migration flows for Italian and Spanish candidates and a smaller correspondence for French and German candidates. The quality of the correlation in terms of r^2 was rather high.

6 Conclusions

The objective of this study was to test the usefulness of Google Trends indexes to “nowcast” or forecast (short-term) migration using a simple model in a context of (relatively) regular migration. According to Choi and Varian (2012); Google Trends represent a way to predict current trends in different domains, such as automobile sales, unemployment claims, travel destination planning, housing prices (Wu and Brynjolfsson 2015), home refinancing rate (Mohebbi 2011) and consumer confidence. It also allows surveying the diffusion of diseases, such as influenza or salmonella (Mohebbi 2011; Ginsberg et al. 2009).

Before summarizing the results of this study, it is important to note the limits of our approach. First, Google Trends provides an index of searches for information on the Swiss labour market, which may differ from actual intention to migrate to/work in Switzerland. The index does not take into account the actual possibility of obtaining access to a job, which can also depend on candidates’ individual characteristics and their knowledge of opportunities in Switzerland. Second, candidates for migration can also access information on the Swiss labour market using other channels, such as friends, family members, and employers (Windzio 2018, Delhey and al. 2019). The so-called chain migration is well documented (see for instance Durand and Massey 2004) and can provide all the useful information to migrants. However, in the case of Switzerland, the phenomenon of chain migration appears to be rather low in international comparison (Crettaz and Dahinden, 2019). Therefore, the use of Google Trends indexes is far from comprehensive. Moreover, we focused here on the *immigration* flows to Switzerland, which we try to explain by the number of Google Trends index. An analysis of the link between this index in a country and *emigration* flows would have strengthened our results. However, such an analysis is difficult to conduct due to the lack of reliable data on emigration. Finally, we did not control for duplicate or repeated searches performed during the same period and presumed that the behaviour of candidates was constant during the time under study. Even if the Google Trends index accurately assesses intentions to migrate to/work in Switzerland, there is a difference between the intentions and actual behaviour of candidates.

The existing literature provides us with two different methodological approaches. Some studies input many keywords or key phrases in different algorithms, while others focus specifically on one keyword or a small number of specific keywords. Our approach was to choose one key phrase that has been previously used in other contexts (UNFPA 2014; Wladyka 2017) rather than to include many keywords in a complex algorithm. This choice is justified by the fact that the search for information regarding Switzerland among people living in EU/EFTA countries can also be related to other forms of movements, such as

tourism or business trips. Therefore, focusing on one specific key phrase specifically oriented to migration appears to better fit with our objectives. The main objective of the approach also advocates a simplified model based on a limited number of parameters and a simple regression. Indeed, to ensure the best predictability and quality, model parameters should ideally be validated on a regular basis, as soon as updated data become available. A system that is too complex would then incur costs that exceed the benefits. In this context, our method provides only a sentinel indicator, which does not aim to replace official statistics but to provide some preliminary trends on future immigration flows.

Our model is based on a linear regression estimated through an ordinary least squares method that assumes that all observations are independent. Alternative approaches exist, such as including other parameters in the model translating the existence of alternative sources of information (social networks and other search engines such as lilo.org in France) or alternative choices (cross-border work and failed migration). However, such approaches will probably not significantly improve the results. Too complex techniques would also increase the difficulty of implementing the model to the monitoring of future flows.

The relationship between the Google Trends index and official immigration estimates can vary over time. Our models are estimated for the 2006–2016 period, but we also tested different periods. As a whole, the choice of the period does not modify the slope of the regression. This finding indicates that the proposed modeling is not very influenced by time and is therefore rather stable.

In a context where statistical offices tend to limit access to individual data (or charge for it), alternative data that are easily accessible can be useful for scientists and planners. Our results show mixed results. Notwithstanding these limits and possible developments, our study demonstrates, on the one hand, that Google Trends can provide information that translates the level of attractiveness of different countries to the number of emigrants from other countries. On the other hand, Google Trends is not able to systematically predict, for some bilateral migration flows, current and future (short-term) trends. Regarding Google Trends' capacity to anticipate migratory flows, it appears that the platform was a satisfying predictor of increases in flows of adult immigrants from countries belonging to the same continent and the same economic area: for Italy and Spain, increases in the index were closely related to increases in migration to Switzerland. For the same countries, however, decreases in the index during the mid-2010s did not precisely reflect migratory flows. It is interesting to observe that Google's capacity to predict future flows is better during the first half of the period under study, which was characterized by an increase in migration flows, than during the last 5 years. This finding can be explained by the fact that because migration was at a high level during the last 20 years, an increasing number of immigrants may have obtained information from their friends and family members who had already emigrated to Switzerland.

The availability of current data on migration is central to efficiently managing migration policies, particularly when contingent policies are regularly updated and, more generally, when immigration impacts society. Therefore, there is a compelling need for current figures and data on short-term trends. To date, the use of Google Trends to predict migrant flows or intentions has been applied to long-distance (intercontinental) migration flows and for specific groups of migrants, such as forced and irregular migrants. It is not surprising that governments and international organizations, in search of monitoring systems that can help them track immigrant flows, are increasingly looking at Google searches in order to control generally undesirable migrant flows. The research released in 2017 by the Pew Research Center on Syrian Middle Eastern migrants to Europe demonstrated the capacity of this approach to “nowcast” the flows and eventually to intervene in them (Connor 2017

). In this context, the approach based on big data may rapidly have a significant impact on the monitoring of forced and irregular migrations. Our paper addresses labour migration and free movement migration, which are generally considered as difficult to predict (OECD 2018). It also addresses the form of immigration that is generally welcomed by the destination country, which is certainly why efforts to determine future migration trends using new approaches is lower for these kinds of flows. However, due to the importance of correcting trends to monitor labour force migration, testing alternative approaches is necessary.

In conclusion, combining traditional and innovative sources is the only way to adequately describe migration and mobility in a context in which traditional statistics are unable to precisely measure flows. In this context, the current study tests one of the many approaches based on big data, with mixed results. Compared to other approaches, such as those using social networks, the advantage of Google Trends is that limitations related to penetration rates, variable-level social network use and fake accounts are not prevalent. Our results show that the approach tested is able to predict future increase in the migration flows, but fail to predict decreases or rapid changes in the trends. Based on the overall results obtained, further efforts are required before such data can be fully used for administrative purposes.

Funding Open access funding provided by University of Geneva. This research was supported by the National Center of Competence in Research ncr—on the move funded by the Swiss National Science Foundation.

Compliance with ethical standards

Conflict of interest The author has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahas, R., Silm, S., & Tiru, M.: Measuring Transnational Migration with Roaming Datasets. Zurich, 14th Conference on Location Based Services. doi: <https://doi.org/10.3929/ethz-b-000225599>. (2018)
- Arango, J.: Explaining migration: a critical view. *Int Soc Sci J* **52**, 283–296 (2000). <https://doi.org/10.1111/1468-2451.00259>
- Armitage, C.J., Conner, M.: Efficacy of the theory of planned behaviour: a meta-analytic review. *Br J Soc Psychol* **40**(4), 471–499 (2001). <https://doi.org/10.1348/014466601164939>
- Barnett, G.A., Xu, W.W., Chu, J., Jiang, K., Huh, C., Park, J.Y., Park, H.W.: Measuring international relations in social media conversations. *Gov Inf Q* **34**, 37–44 (2017). <https://doi.org/10.1016/j.giq.2016.12.004>
- Blumenstock, J.E.: Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda. *Inf Technol Dev* **18**(2), 107–125 (2012). <https://doi.org/10.1080/02681102.2011.643209>

- Böhme, M.B., Gröger, A., & Stöhr, T.: Searching for a Better Life: Predicting International Migration with Online Search Keywords, Berlin: The IZA World Labor Conference. Retrieved from: https://conference.iza.org/conference_files/WoLabConf_2018/22651.pdf. (2018)
- Carling, J. (2014). The role of aspirations in migration. University of Oxford, Oxford, Determinants of International Migration. International Migration Institute pp. 23–25. John Wiley & Sons Ltd
- Carling, J.: How does migration arise? In: McAuliffe, M., Klein Solomon, M. (eds.) Ideas to inform international cooperation on safe orderly and regular migration. IOM, Geneva (2017)
- Carling, J., Vatne Pettersen, S.: Return migration intentions in the integration–transnationalism matrix. *Int Migr* **52**(6), 13–30 (2014). <https://doi.org/10.1111/imig.12161>
- Chen, C., Ma, J., Susilo, Y., Lu, Y., Wang, M.: The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transp Res Part C Emerg Technol* **68**, 285–299 (2016). <https://doi.org/10.1016/j.trc.2016.04.005>
- Choi, H., Varian, H.: Predicting the present with google trends. *Econ Record* **88**(s1), 2–9 (2012). <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
- Connor, P.: ‘The digital footprint of Europe’s refugees. Methodology’. Pew Research Center. Available on <https://www.pewglobal.org/2017/06/08/online-searches-eu-refugees-methodology/>. (2017)
- Constant, A.F., Zimmermann, K.F.: Circular and repeat migration: counts of exits and years away from the host country. *Popul Res Policy Rev* **30**(4), 495–515 (2011). <https://doi.org/10.1007/s11113-010-9198-6>
- Creighton, M.J.: The role of aspirations in domestic and international migration. *Soc Sci J* **50**(1), 79–88 (2013). <https://doi.org/10.1016/j.soscij.2012.07.006>
- Crettaz, E., Dahinden, J.: How transnational are migrants in Switzerland? an analysis of the migration-mobility-transnationality nexus. In: Steiner, I., Wanner, P. (eds.) *Migrants and expats: the Swiss migration and mobility nexus*. IMISCOE Research Series Springer, London (2019). https://doi.org/10.1007/978-3-030-05671-1_11
- D’Amato, G., Wanner, P., Steiner, I.: Today’s Migration-Mobility Nexus in Switzerland. In: Steiner, I., Wanner, P. (eds.) *Migrants and expats: the swiss migration and mobility nexus IMISCOE research series*. Springer, London (2019)
- De Haas, H.: ‘Migration Theory. Quo Vadis?’, DEMIG project paper 24. Oxford. Retrieved from <https://www.imi.ox.ac.uk/publications/wp-100-14>. (2014)
- Jong, De, Gordon, F.: Expectations, gender, and norms in migration decision-making. *Popul Stud* **54**(3), 307–319 (2000). <https://doi.org/10.1080/713779089>
- Delhey, J., Verbalyte, M., Aplowski, A., Deutschmann, E.: Free to move: the evolution of the European migration network, 1960–2017. In: Heidenreich, M. (ed.) *Horizontal europeanisation: the transnationalisation of daily life and social fields in Europe*, pp. 63–88. Routledge, New York (2019)
- DESTATIS.: *Bevölkerung und Erwerbstätigkeit. Wanderungen 2015*’. Wiesbaden, Statistisches Bundesamt, 118p. (2017)
- Docquier, F., Peri, G., Ruysen, I.: The Cross-country Determinants of Potential and Actual Migration. *Int Migr Rev* **48**, S37–S99 (2014)
- Dubois, A., Zagheni, E., Garimella, K., and Weber, I.: ‘Studying Migrant Assimilation Through Facebook Interests’. In: Staab, S., Koltsova, O., Ignatov, D. I., (eds.) *Social Informatics: 10th International Conference, SocInfo 2018, St. Petersburg, Russia, September 25-28, 2018, Proceedings, Part II*, 51–60. Springer, Cham (2018)
- Durand, J., Massey, D.S.: *Crossing the border: research from the Mexican migration project*. Russell Sage Foundation, New York (2004)
- Giannone, D., Reichlin, L., Small, D.: Nowcasting: The realtime informational content of macroeconomic data. *J Monetary Econ* **55**(4), 665–676 (2008)
- Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., et al.: Detecting influenza epidemics using search engine query data. *Nature* **457**, 1012–1014 (2009). <https://doi.org/10.1038/nature07634>
- Gonzales, M.C., Hidalgo, C.A., Barabasi, A.L.: Understanding individual human mobility patterns. *Nature* **453**(7196), 779–782 (2008). <https://doi.org/10.1038/nature06958>
- Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Ratti, P.K.C.: Geo-located twitter as proxy for global mobility patterns. *Cartogr Geogr Inf Sci* **41**, 260–271 (2014). <https://doi.org/10.1080/15230406.2014.890072>
- Herdagdelen, A., State B., Adamic L., & Mason W.: ‘The Social Ties of Immigrant Communities in the United’, Proceedings of the 8th ACM Conference on Web Science, ACM, 78–84 (2016). doi: <https://doi.org/10.1145/2908131.2908163>(2016)
- IOM - International Organization for Migration ‘Big Data and Migration’. *Data Bulletin 5: Informing a Global Compact for Migration*. Geneva. 4p. Retrieved from <https://www.gmdac.iom.int/data-bulletin-5-big-data-and-migration> (2018)

- Laczko, F., Tjaden, J.D., Auer, D.: 'Measuring Global Migration Potential, 2010–2015' International Organisation for Migration. *Glob Migr Data Anal Cent Data Brief Ser* **9**(1), 1–14 (2017)
- Migali, S., Scipioni, M.: A global analysis of intentions to migrate. *Eur Comm (JRC111207)*, 60p (2018)
- Mohebbi, M., Vanderkam, D., Kodysh, J., Schonberger, R., Choi, H. & Kumar, S.: (2011 Google correlate whitepaper'. Available on www.google.com/trends/correlate/whitepaper.pdf.
- OECD: Can we anticipate future migration flows? *Migration Policy Debates* **16**, 9p (2018)
- Ossman, S.: Studies in serial migration. *Int Migr* **42**(4), 111–121 (2004). <https://doi.org/10.1111/j.0020-7985.2004.00297.x>
- Paul, A.M.: Stepwise international migration: a multistage migration pattern for the aspiring migrant. *Am J Sociol* **116**(6), 1842–1886 (2011). <https://doi.org/10.1086/659641>
- Rango, M., & Vespe, M.: 'Big Data and alternative data sources on migration: From case-studies to policy support'. European Commission – Joint Research Centre (JRC) Ispra Retrieved from [https://gm-dac.iom.int/data-bulletin-big-data-and-migration\(2017\)](https://gm-dac.iom.int/data-bulletin-big-data-and-migration(2017))
- Ruggles, S.: Big microdata for population research. *Demography* **51**(1), 287–293 (2014). <https://doi.org/10.1007/s13524-013-0240-2>
- Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., Rango, M.: 'Migration Data using Social Media. *Eur Perspect JRC Tech Rep Eur Comm* (2018). <https://doi.org/10.2760/964282>
- State, B., Weber, I., Zagheni, E.: Studying inter-national mobility through IP geolocation. *ACM Int Conf Web Search Data Min* (2013). <https://doi.org/10.1145/2433396.2433432>
- State, B., Rodriguez, M., Helbing, D. & Zagheni, E.: 'Migration of Professionals to the U.S. Evidence from LinkedIn data', The 6th International Conference on Social Informatics (SocInfo) Retrieved from: <https://engineering.linkedin.com/research/2014/migration-of-professionals-to-the-us-evidence-from-linkedin-data> (2014)
- Steiner, I.: Immigration and settlement? german migration flows to and from Switzerland under the provision of free movement of persons. Université de Genève, Genève (2017). <https://doi.org/10.13097/archive-ouverte/unige:96470>
- Tjaden, J., Auer, D., Laczko, F.: Linking migration intentions with flows: evidence and potential use. *Int Migr* **57**(1), 36–57 (2019). <https://doi.org/10.1111/imig.12502>
- UNFPA: 'Estimating migration flows using online search data' www.unglobalpulse.org. Retrieved from <https://www.unglobalpulse.org/projects/migration-search-data> (2014)
- United Nations.: 'Report of the Global Working Group on Big Data for Official Statistics'. New York: United Nations Economic and Social Council, 17p. Retrieved from <https://unstats.un.org/bigdata/bureau/documents/reports/GWG%2520report%2520-%25202018-8-BigData-E.pdf>. (2018)
- Van Dalen, H. P., & Henkens, K.: 'Emigration intentions: mere words or true plans? Explaining international migration intentions and behavior'. *Discussion Paper CentER*. Retrieved from <https://pure.uvt.nl/portal/files/1001393/2008-60.pdf> (2008)
- Wanner, P., Steiner, I.: Une augmentation spectaculaire de la migration hautement qualifiée en Suisse. *Soc Change Switzerland* **16**, 267–295 (2018). <https://doi.org/10.22019/SC-2018-00007>
- Willekens, F., Massey, D., Raymer, J., Beauchemin, C.: International migration under the microscope. *Science* **352**(6288), 897–899 (2016). <https://doi.org/10.1126/science.aaf6545>
- Williams, S., & Ralphs, M.: 'Preliminary Research into Internet Data Sources', 18th GSS Methodology Symposium, Retrieved from unglobalpulse.org/sites/default/files/williamsralphs.pdf (2013)
- Windzio, M.: The network of global migration 1990–2013: Using ERGMs to test theories of migration between countries. *Soc Net* **53**, 20–29 (2018)
- Wladyka, D.K.: Queries to google search as predictors of migration flows from Latin America to Spain. *J Popul Soc Studies* **25**(4), 312–327 (2017). <https://doi.org/10.25133/JPSSv25n4.002>
- Wu, L., Brynjolfsson, E.: The future of prediction: how Google searches foreshadow housing prices and sales. *Econ Anal Dig Econ* **89**, 118 (2015). <https://doi.org/10.7208/chicago/9780226206981.003.0003>
- Zagheni, E., Garimella, V. R. K., Weber I., & State, B.: 'Inferring international and internal migration patterns from Twitter data', Proceedings of the companion publication of the 23rd international conference on World Wide Web companion, doi: [https://doi.org/10.1145/2567948.2576930\(2014\)](https://doi.org/10.1145/2567948.2576930(2014))
- Zagheni, E., Weber, I.: You are where you e-mail: using e-mail data to estimate international migration rates. *Web Sci* **348**, 351 (2012). <https://doi.org/10.1145/2380718.2380764>
- Zagheni, E., Weber, I., & Gummadi, K.: 'Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants', Retrieved from http://www.zagheni.net/uploads/3/4/4/7/34477700/zagheni_weber_gummadi_2017_accepted_version.pdf. (2017)