Boletín de la Asociación de Geógrafos Españoles N.º 72 - 2016, págs. 563-567
I.S.S.N.: 0212-9426

THE STATISTICAL PICTURE OF INTERNATIONAL MIGRATION OUTFLOWS IN SPAIN: ANALYSIS OF THE NEW STATISTICAL SOURCES REGARDING THE RELIABILITY OF THE RESIDENTIAL VARIATION STATISTICS

Carmen Ródenas
Mónica Martí
Applied Economic Analysis Department
University of Alicante
crodenas@ua.es, mmarti@ua.es

Spain, compared to other EU countries, was one of the biggest receivers of international immigrants in the years previous to the economic recession, between 2000 and 2007. This phenomenon, clearly perceived by society and the mass media, was not collected in the same way by national statistical sources available at the time. For different methodological reasons[1], the strong growth in the pace of foreign inflows since 1999 was not properly registered in the *Spanish Labour Force Survey* (*Encuesta de Población Activa*, EPA) or in the 2007 *National Immigrant Survey* (*Encuesta Nacional de Inmigrantes*, ENI) or in the 2001 *Population Census*. The only statistical source that reflected the strong increase in arrivals was the Residential Variation Statistics (*Estadística de Variaciones Residenciales*, EVR), an administrative register which contains the registrations (and de-registrations) due to a change of place of residence declared by the population in Spain. The EVR is used to update both the official *Population Database* (*Base Padronal*), centralized by the National Statistics Institute (*Instituto Nacional de Estadística*, INE), and the corresponding Municipal Population Registers (see a comparison between different sources in Figure 1).

The EVR is based on the requirement for residents in Spain to register any change of habitual residence to the local authorities, although in certain circumstances it is the municipalities who officially register these changes. The fact that the inscription on the Register is

1 Explained in Ródenas y Martí (1997) and in Martí y Ródenas (2004, 2007) for the case of the *Encuesta de Migraciones/EPA*; regarding ENI in Martí y Ródenas (2012a) and Ródenas y Martí (2013) and, finally, related to 2001 *Population Census* in Ródenas y Martí (2009a, b).

required and is necessary for the exercise of very basic rights such as public health coverage, school enrollment or to obtain residence permits, means that the EVR records of registrations in Spain are considered a good reflection of the inflow of international migration. However, direct observation of the EVR de-registrations (available since 2002) might be insufficient for proper measurement of migratory outflows from Spain. This is particularly true, in the case of foreigners, as few of them apply for deregistration from the Municipal Population Registers when they leave Spain.

In this situation, there are two good reasons for taking another look at the analysis of migration figures from the Spanish statistics. Firstly, because since 2008 the country has been undergoing a profound economic crisis and the extent to which it is impacting on the outflows of foreigners is unknown; we do not know how many immigrants have left and how many of those who came before the recession remain. And, secondly, because, in addition to the publication by the INE in December 2013 of the detailed results of the 2011 *Population Census* on the basis of a new methodology, we have new official statistical sources. In April 2014 the INE published the first year, 2013, of the *Continuous Household Survey* (*Encuesta Continua a los Hogares*, ECH); the series of the *Labor and Geographic Mobility Statistics* (*Estadística de Movilidad Laboral y Geográfica*, EMLG) was released in November 2013 and, finally, in this year the figures of the *Migration Statistics* (*Estadística de Migraciones*, ESMI) are being made available. Given that these sources are derived from new and innovative statistical procedures, our paper is even more relevant.

The methodological developments in these sources are briefly as follows. To begin, the 2011 *Census* was not built as a classical census with an exhaustive enumeration of houses, but has combined information from different administrative records (including the *Population Database*) with a specific survey of about three million households (12.3% of the population). The ECH has been designed to become the backbone of statistical operations that require subsamples for other surveys of households in Spain. The sample of the ECH is surveyed over the four quarters and provides annual information on basic demographic characteristics of the population, households and housing; similar to the census information but more frequent than the decennial censuses. The case of the EMLG is different. Its innovation is the integration of data provided by the sample of the EPA with the information recorded in the Population Database. With this source, therefore, we are able to investigate the relation between the situation in the labor market and the changes of residence of each individual. The last source, ESMI, is a statistical exploitation of the EVR. Its aim is to improve the mere observation of registered population movements, focusing on those associated with real migrations. Thus, in the ESMI, INE technicians (i) estimate the date of occurrence of emigration abroad when the de-registrations of foreign nationals are recorded by the municipalities, as is the case of the so called cancellations by expiration (*bajas por caducidad*) and cancellations for undue inscription (*bajas por inclusion indebida*). The first arise as a result of the legal requirement for non-nationals to renew their registration every two (non-EU citizens) or five (EU citizens) years. The second are due to residential variations bound for another country where the destination country is officially unknown; (ii) they try to adjust the time lag between the occurrence of migratory movement and the time of registration; (iii) they clean the database removing registrations and cancellations where less than twelve months have passed between each occurrence for the same individual in order to lead back

data to the concept of migration; and, finally, (iv) they do imputations of the country of birth, country of nationality and country of origin or destination of migratory flows that have any of such variables blank, or with invalid values.

Therefore, our paper focuses on analyzing the results of these sources about recent international migrations in Spain. In addition, we will attempt to find out if they describe a similar picture, and if not we will try to explain why. Moreover, using the new sources, we will try to specify the degree of underestimation (overestimation) of the international de-registrations (registrations) in the EVR, that is, to assess the reliability of this long-established statistic. This ultimate goal is justified in that the characteristics of the EVR (continuous updating and inclusion of all persons residing in the territory regardless of their legal situation, unlike other EU countries), gives us a huge advantage to adequately describe the intensity and the changes in the migratory phenomenon.

Our first set of results is related to the quantification of the underestimation of the de-registrations of foreigners in the EVR by annual cohort entry. Firstly, we compared the external mobility of the 2011 Census and that recorded by the EVR. Once it is verified that the 2011 Census does not have the problems related to sub-estimation of recent immigration (in the last years recorded) that the 2001 Census had, their data are used to approximate a first interval for external cancellations underestimated by the EVR relating to the group of emigrants born outside by entry cohorts in the 2002-2011 period. The sum of the annual differences between EVR registrations and immigrant *survivors* recorded by the Census in 2011 has obviously been lowered by the registered external cancellations known. However, as in the EVR the arrival year of an immigrant who is now going abroad (external de-registrations) is not available, we must establish some hypotheses. Specifically, two assumptions have been established to allocate the registrations due to external immigration to the registered cancellations in the period. The calculations are shown in Table 2, and the conclusion is that the EVR could fail to register between a minimum of 40.01% and a maximum of 64.01% of emigrations. Secondly, it has conducted a similar exercise but based on the new ECH, with a more current reference date, July 2013, which provides a second interval. As shown in Table 3, in this case the EVR underestimate from a minimum equivalent to 18.13% of the flow of emigration and a maximum of 46.54%. What is noteworthy is the drop of the interval of underestimation in relation to 2011, when the economic downturn persisted, therefore we would expect greater intensity of the exits and, thus, greater underestimation.

The third interval estimated is based on the comparison between EVR and EMLG, using information provided by the LFS sample data linking residential variations registered in the INE *Population Database*. The strategy here is not to estimate the differences for a single long period (like the Census of 2002 and 2011, or like the ECH 2002-2013), but to calculate for each of the years available in the EMLG. Thus, for each year between 2009 and 2013, we compare the number of EMLG people aged 16 and over who, coming from abroad, spend less than one year living in their current municipality with the number of net external inscriptions registered yearly by the EVR.

To obtain the correct figure of net EVR annual registrations to be compared to the EMLG data, it has been necessary to subtract two types of cancellations, registered and unregistered, from the gross EVR immigrant flow. On the one hand, we know that registered cancellations of people with entry inscription in the same year must be removed, because if they are not

in Spain, they will of course never be interviewed in the EMLG. This has been done linking the records of annual EVR micro-data files. And we have obtained an annual average of more than 5,000 records of immigration/emigration in the same year. Those people that have immigrated from abroad to Spain and during the same year have carried out a new internal migration, will also not appear on the EMLG as having come from abroad. In the EMLG the origin of this internal migration would be their last (in Spain), while in the EVR they continue to be from abroad. The above mentioned linked micro-data files have also been used to detect and remove from the EVR a minimum of between 6,000 and 10,000 immigrants from abroad, who made a subsequent internal movement in the same year. In addition, it has been necessary to estimate the number of a second type of cancellations. It is the case of the –above explained– cancellations by expiration and cancellations for undue inscription corresponding to each of the analyzed years. As these are official cancellations resulting from long administrative processes, it has been necessary to approximate their value.

After removing all of these cancellations, the difference between EVR annual net registrations and the number of annual EMLG immigrants –see Table 6–, rises a minimum of 33% and a maximum of 54% of annual net inflow registered by the EVR. However, taking into account that recent mobility is underestimated by the EPA and, in consequence, by the EMLG, and that it must reach at least 22.4% (see Table **7** estimation); then, the last interval of underestimation of the external EVR cancellations would be between 18% and 43.8% of registrations in the same year. Considering the three intervals, it seems reasonable to conclude that in the EVR around a third of the external registrations of each year may not be counted as external cancellations.

Our second set of results is related to the quality of the new statistical sources. On the one hand, it is worrying not to find that for a particular and recent year, 2013, entries from abroad of EVR are, as they should be, slightly above immigrants captured by the ECH. Specifically, the ECH establishes at 338,900 the number of persons born abroad who have come abroad under one year of residence in Spain in July 2013. However, the EVR registered 315,773 immigrations for this group between July 2012 and June 2013; that is, almost 7% less (see Table 5). This confirms that, of course, the EVR does not overestimate the flow of immigration. But the fact that this difference is not distributed uniformly or independently of the characteristics of immigrants should be a good reason to consider the possibility that the newly implemented ECH had some sort of bias. The fact that the differences focus nearly exclusively on females and the age group 25-44, reminds us of the problems of sample design in the EPA generated when the family structure, gender distribution and the situation in the labor market of the various groups, make more likely to "find" some than others at the time of the survey. And, therefore, skewed estimations might be obtained. If so, the ECH should carefully review the results of questions such as immigrants' residence period.

Finally, it is known that Demographic Accounting provides an equality that allows the connection between stocks and flows of people. In the paper we have added to the initial 2011 Census stock of foreign-born the net international flows of EVR, and, alternatively, of ESMI between January 2012 and June 2013. The result is that there would be a shortfall of 97,117 people according to the EVR and, what is worse, a shortfall of more than 330,000 according to ESMI, to match the stock in July 2013 provided by the ECH (see Table 4). Assuming there are no mistakes in the initial and final stocks, and that flows reflect reality,

then the difference would correspond to false cancellations (not real emigrations) in the EVR and the ESMI. Perhaps this is because cancellations by expiration and cancellations for undue inscription corresponding to each Municipal Council ex officio de-registrations, are not real. In other words, today there would be no underestimation in the EVR for international emigrations of those born abroad, rather the opposite. What is of concern should be the result obtained from the use of the ESMI flows, the "cleansing" of residential variations, because they almost triple the difference and that can only be a consequence of the criteria (expansion coefficients, weights, and so on) and procedures adopted by INE technicians.