SORT 36 (1) January-June 2012, 103-114

Introducing migratory flows in life table construction

Jose M. Pavía[1], Francisco Morillas[2] and Josep Lledó[3]

Abstract

The purpose of life tables is to describe the mortality behaviour of particular groups. The construction of general life tables is based on death statistics and census figures of resident populations under the hypothesis of closed demographic system. Among other assumptions, this hypothesis implicitly assumes that entries (immigrants) and exits (emigrants) of the population are usually not significant (being almost of the same magnitude for each age compensating each other). This paper theoretically extends the classical solution to open demographic systems and studies the impact of this hypothesis in constructing a life table. In particular, using the data of residential variations made available to the public by the Spanish National Statistical Office (INE, Instituto Nacional de Estadística) to approximate migratory flows, we introduce in the process of constructing a life table these flows and compare, before and after graduation, the crude mortality rates and the adjusted death probabilities obtained when migratory flows are, and are not, taken into account.

MSC: 91D20, 62Q05

Keywords: General population, Lexis diagram, open demographic system.

1. Introduction

In the demographic and actuarial fields, the analysis of mortality in a population has particular relevance for their applications. Life tables, or mortality tables, are used to recreate an observed mortality situation or to present future values of the evolution of

[1] Department of Applied Economics, Facultat d'Economía, Universitat de Valencia, Campus Els Tarongers, 46022-Valencia (Spain), e-mail:pavia@uv.es

[2] Department of Applied Economics, Facultat d'Economía, Universitat de Valencia, Campus Els Tarongers, 46022-Valencia (Spain), e-mail:francisco.morillas@uv.es

[3] Actuarial Services Consultant, Ernst & Young, Torre Picasso, 1, 28020-Madrid (Spain), e-mail:jonabe@alumni.uv.es
Received: November 2011
Accepted: February 2012

mortality in certain groups, making it possible to generate demographic forecasts or to calculate premiums and/or income for life insurance and pension benefits. Medicine is another area where mortality analysis is also frequently used.

Life tables are usually drawn from the study and analysis of the intensity and rate at which mortality affects each age group in question. In general populations it is generated using information mainly from population censuses and lists of deceased, where individual records from insurance policies are the prime source of information in insured populations. To be specific, and once the sample period has been decided, the comparison between the numbers at risk and the number of deaths allows the actuary (demographer) to obtain initial (crude) estimates for the probability of death in each age group q_x. These probabilities are subjected to the corresponding graduation or adjustment processes (see, for example, Copers-Haberman, 1983; Forfar et al. 1988: or Ayuso et al., 2007) with a view to smoothing the profile of the associated stochastic process and to ultimately develop appropriate tables from a fictitious starting population of size ℓ_0.

In the construction of mortality tables for general populations (which is the subject of this paper) it is not usual to specifically consider migratory flows, making the hypothesis of closed demographic system (HCDS), which implicitly entails the assumption of certain limitations, the main ones being: (i) that migration flows (inputs and outputs) of the population by age and sex are considered not to be significant; (ii) that for each age group migration flows are random and show similar entry and exit figures; and, (iii) that immigrants acquire the same risk of death as the resident population.

These limitations, which are not always reasonable, should be checked because of their potential impact on, for example, the calculation of life expectancy, of premiums for life insurance or of estimates for the calculation of pensions. The aim of this paper is twofold, firstly, to introduce an estimator for an open demographic system and, secondly, to show the incidence of HCDS through a real case. More specifically, given the immense pressure of migration endured by Europe, and in particular Spain, in recent years, the analysis will be based on the comparison of mortality tables (by gender) obtained for Spain under HCDS and under the hypothesis of open demographic system (HODS), in which migratory flows are explicitly considered. The comparison is carried out in two ways. Firstly, the differences between the estimated crude probabilities obtained under each hypothesis are compared and, secondly, the comparison is again carried out after having graduated the crude data.

The rest of the paper is structured as follows. Section 2 explains the methodology used to obtain the mortality tables, both for the closed demographic system and the open demographic system. In Section 3, different comparisons are carried out between the HCDS and HODS tables. The last section presents the conclusions reached and indicates several issues for future research.

2. Methodology

The techniques and formulae used to estimate life tables are heavily influenced by the type of information available. When working with official statistics, the relevant data are usually offered, by and large, in aggregate form. Hence, in this research we have opted to work with aggregated figures – which come from information that the Spanish National Statistical Office (INE) has made available on its website (http://www.ine.es) – despite it being possible in the Spanish case to use some detailed (anonymized) microdata[1]. In particular, we will consider that, for each gender, aggregated figures of migrants, deaths and population are available by age and calendar year. Under these circumstances, the representation and analysis of the information in a Lexis diagram (named after the German statistician, economist and social scientist Wihelm Lexis, who adopted it in the nineteenth century to illustrate the procedures for calculating a mortality table) often greatly facilitates the reasoning and makes more manageable, after the application of a number of reasonable hypotheses, handling the flaws of detailed information which are present in aggregate data.[2]

The Lexis diagram is a two-dimensional diagram of lifelines with two-temporal dimension: calendar time and age. Calendar time is represented on the horizontal axis, while age is represented on the vertical axis. This diagram permits representation of the life events of a population from the personal history of the individuals within it. Each personal story is represented by a line segment forming an angle of 45 degrees to the horizontal axis. The classical approach (of closed demographic system) states that each personal story begins at birth, which is represented on the baseline, and ends at some point on the graph with the individual's death (Livi Bacci, 2000). In this paper, however, personal history is not determined solely by the events of birth and death. The introduction of migratory flows makes it necessary to modify the classical interpretation of the Lexis diagram since, in this case, the history of an individual could start from birth or from immigration and, likewise, it could end by death or by emigration.

Figure 1 shows a small section of a Lexis diagram which, as usual, comes divided into cells of dimensions of 1×1 so that between each pair of oblique lines are the lifelines that make up a generation of individuals and where each cell represents an observation period of one year (in which the age of the individuals also has a variation of a year).

For example, in Figure 1 (left) various individual life lines (thin lines) have been represented. Lifelines of individuals entering in the system due to immigration can originate anywhere in the diagram and are differentiated from the rest by having a $\circ$ at its origin. When an individual leaves the system, the cause of departure is differentiated graphically: if the reason is death, the tail end of the lifeline is marked with an $\times$; if the reason is emigration the end-lifeline is marked with a $\square$.

1. Such is the case of death statistics and of residential variation figures (directly available on the INE website). Indeed, the INE, unlike many other national and international statistical agencies, is characterized by its willingness to make available, in anomymized form, the detailed data generating the vast majority of its statistical operations.
2. Obviously, some of the hypotheses to be considered could be relaxed using more detailed data (see, for example, INE, 2009).

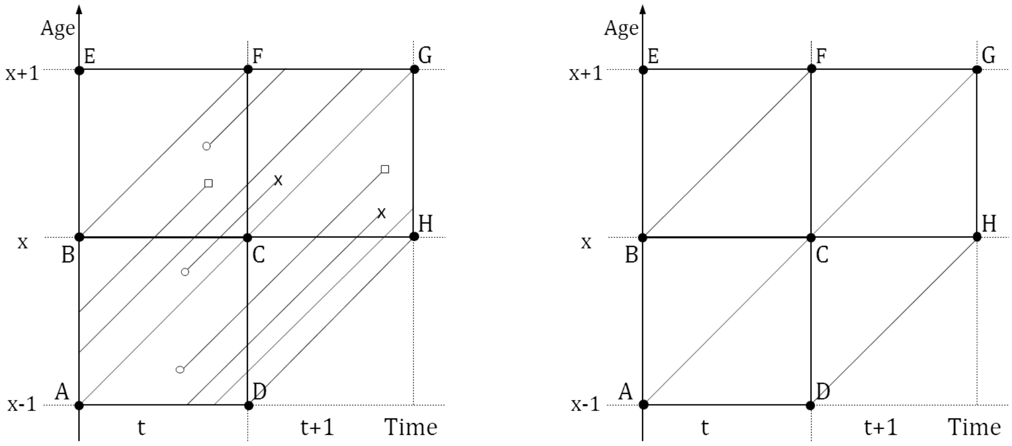

Figure 1: *Detail* (2×2) *of Lexis diagram with lifelines (left) and schematic (right).*

The aggregate information, however, does not allow for an accurate location of the lifelines of the individuals who comprise the study population. To make use of this geometric representation, therefore, the usual convention of the Lexis scheme of assigning values to segments and surfaces must be adopted. In this paper, the value of a segment is always identified by the number of lifelines that cross it, while the value(s) to be associated with each area will depend on the hypothesis under consideration.

2.1. Closed demographic system

Under HSDC, each surface is identified with a single value: the number of lifelines that end in it due to death. So, assuming for simplicity that (as in our case) for any given age x there are available the number of residents counted in January 1 of year t, C_x^t, and the number of residents who died in each age x for years t and $t+1$ (D_x^t and D_x^{t+1} respectively), we have that, as shown in Figure 1 (right), it is straightforward to draw the information. To be specific, on the one hand, the quadrilaterals $ABCD$, $BCEF$ and $CFGH$ will be identifiable respectively with D_{x-1}^t, D_x^t and D_x^{t+1}, and, on the other hand, the segment AB will be equal to C_{x-1}^t.

At this point, it is now easy to obtain an initial estimate of q_x exploiting the geometric properties of the representation and assuming uniform distribution of birth dates and deaths within each age group and year. In particular, noting that the segment BC represents the number of people reaching age x in year t, it follows that under HSDC the number of them who die before reaching age $x+1$ will come represented by the quadrilateral $BCFG$ and that therefore an estimate for q_x is obtained from:[3]

3. Given that $q_x = \frac{d_x}{\ell_x}$ is defined as the quotient of the number of deaths between ages x and $x+1$, d_x, and the number of survivors at age x, ℓ_x.

$$\widehat{q}_x = \frac{BCFG}{BC} \tag{1}$$

And from this, using the geometry of the scheme, one arrives at $BCFG = BCF + CFG$ and $BC = AB - ABC$; from which, using the hypothesis of uniform distribution, one deduces $BCFG = \frac{BCEF}{2} + \frac{CFGH}{2}$ and hence:[4,5,6]

$$\widehat{q}_x = \frac{\frac{1}{2}\left(D_x^t + D_x^{t+1}\right)}{C_{x-1}^t - \frac{1}{2}D_{x-1}^t} \tag{2}$$

2.2. Introducing migratory flows

Assuming HODS, the expression (2) becomes invalid and another approach is required to take into account the entries and exits that happen in the study group during the analysis period. At this point, it will be useful to refer to the type of reasoning usually employed for insured groups (see, for example, Benjamin and Pollard, 1992), where the number of deaths observed is separated depending on the risks of death and the time exposed to risk. That is, under HODS, the number of deaths observed, $BCFG$, will be approximately equal to the number of people that reach age x, BC, by the probability of any of them dying before reaching the age $x+1$, q_x, plus the number of people that immigrate with age $x+k$ (where $0 < k < 1$), whose lifeline starting point would be located in the surface $BCFG$, by the probability that a person of age $x+k$ dies before reaching age $x+1$, ${}_{1-k}q_x$,[7] minus the number of people that emigrate with age $x+k$ (where $0 < k < 1$), whose lifeline end point would be located in surface $BCFG$, by the probability that a person of age $x+k$ dies before reaching the age $x+1$, ${}_{1-k}q_x$.

---

4. This general expression, nevertheless, would not be appropriate for ages zero and one, since as it is well-known the deaths of children under one year old are concentrated in their first weeks of life. The assumption of uniform distribution cannot therefore be maintained for deaths counted with zero years: the greater part of these deaths will be located in the corresponding lower triangle. Thus, the formulae used for ages zero and one have been, respectively, $\widehat{q}_0 = \frac{0.7D_0^t + 0.3D_0^{t+1}}{B^t}$ (where B^t denotes the births in year t) and $\widehat{q}_1 = \frac{\frac{1}{2}(D_1^t + D_1^{t+1})}{C_0^t - 0.3D_0^t}$; which can be obtained assuming that the number of deaths occurring during the first half of age zero is approximately four times the number of deaths registered during the second half. Obviously, if the deaths by generation (also available on the INE website; INE, 2010a) were used, no hypothesis about how to distribute the deaths between the triangles would be necessary because of the values of these would be known exactly.

5. Unlike the expression used to estimate q_x in this paper, until recently the INE started with $BC = CF + BCF$ and arrived at a different equation, $\widehat{q}_x = \frac{\frac{1}{2}\left(D_x^t + D_x^{t+1}\right)}{C_x^{t+1} + \frac{1}{2}D_x^t}$, although equivalent under HCDS (INE, 2007). Since 2009, the INE estimates central age-specific death rates, m_x, using the detailed information available on death microdata to obtain in each age the exact time spent by those who die during the year of study (see, INE, 2009). The use of those detailed data makes it unnecessary to assume any hypotheses about the distribution of deaths within each age and calendar year.

6. A general formula to estimate q_x when the census of the population is located at any instant of the year and not necessarily at the start can be found in, for example, Pavía (2011, Ex. 91).

7. Note that using this expression entails the implicit assumption that immigrants acquire the same risk of death as the population in which they integrate.

The problem is that, unlike what happens in insured populations, the dates and specific ages at which a person immigrates or emigrates are not usually known, so to obtain a useful expression of the decomposition of $BCFG$ it is essential to extend the classic convention and assign additional variables to each surface of the Lexis diagram. To be specific, we propose to link to each area two new variables: the number of lifelines that start in the surface (immigrants) and the number of lifelines that finish in the surface for reasons other than death (emigrants). With this extension it will then be possible to obtain an operative expression for $BCFG$ from which an estimator for q_x can be derived by simply adding hypotheses (i) on the distribution of entries and exits in each surface (which, in the same way as death distribution, are assumed to be uniform, since it is reasonable for the level of information available) and (ii) on the risk of death throughout each age x (which as a rule is assumed proportional to the period of exposure to risk, that is, ${}_{1-k}q_x = (1-k)q_x$).

As usual when handling official statistics, it is assumed that the total number of people that immigrate and emigrate in any year t and with age x is only known in aggregate terms, I_x^t and E_x^t, respectively. Obviously, other more informative situations in which more precise data about the age distribution of migrants in each year were available, for example from microdata of residential variations, would also be perfectly treated with this strategy.

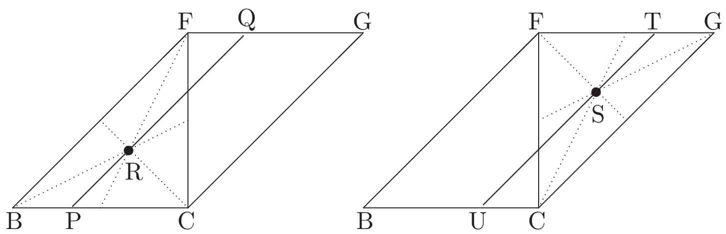

Figure 2: *Barycentres of surfaces of migratory movements.*

Under the conditions above, denoting by $N_x^t = I_x^t - E_x^t$ the net migration recorded in year t at age x, it follows that the number of people who reach age x during year t, BC, would be equal to $BC = C_{x-1}^t - \frac{1}{2}D_{x-1}^t + \frac{1}{2}N_{x-1}^t$, and that the number of entries and exits that would be registered in each of the triangles BCF and CFG would be, respectively, $\frac{I_x^t}{2}$, $\frac{E_x^t}{2}$, and $\frac{I_x^{t+1}}{2}$, $\frac{E_x^{t+1}}{2}$, and likewise, under the same hypotheses, each exit/entry produced in each triangle would be located, in average terms, in the centroid (barycentre) of the corresponding triangle (see Figure 2).[8]

8. Alternatively, it can be demonstrated that the average of the distances (across the lifelines) of each point of the corresponding triangle to segment FG is equal to the distance of the corresponding barycentre to segment FG. For example, considering the triangle BCF and, inside it, an arbitrary point K with coordinates (x,y), it is not difficult to prove that the lifeline of K intersects FG in a point, K', with coordinates $(1+x-y,1)$ —where B has been taken as the origin of the corresponding Cartesian coordinate system. Hence, the Euclidean distance from K to K' would be $(1-y)\sqrt{2}$, from which it follows that the *sum* of all the distances is $\int_0^1\int_0^x \sqrt{2}(1-y)dydx =$

As can be seen in the representation on the left of Figure 2, the point R is in the barycentre or centroid of the triangle BCF, which is easy to prove to be at a distance of $\frac{2\sqrt{2}}{3}$ from point Q, in the same way that point S (which can be taken as representative of all points in which an entry/exit occurs in triangle CFG) is at a distance $\frac{\sqrt{2}}{3}$ from T.[9] From here, taking into account that the distances of P to Q and of U to T are equal to $\sqrt{2}$, equivalent to a year, it follows that on average the exposure to risk of each immigrant/emigrant would be, respectively, $\frac{2}{3}$ and $\frac{1}{3}$. Hence bearing in mind the previous arguments, we have that the number of deaths observed in the parallelogram $BCFG$, $\frac{1}{2}\left(D_x^t + D_x^{t+1}\right)$, could be broken down in the following way:

$$\frac{1}{2}\left(D_x^t + D_x^{t+1}\right) \approx \left(C_{x-1}^t - \frac{1}{2}D_{x-1}^t + \frac{1}{2}N_{x-1}^t\right) q_x + \frac{1}{2}N_x^t \frac{2}{3} q_x + \frac{1}{2}N_x^{t+1}\frac{1}{3}q_x$$

And consequently, an estimator for q_x, under HODS, would be obtained by way of the following expression:[10]

$$\widetilde{q}_x = \frac{\frac{1}{2}\left(D_x^t + D_x^{t+1}\right)}{C_{x-1}^t - \frac{1}{2}D_{x-1}^t + \frac{1}{2}N_{x-1}^t + \frac{1}{3}N_x^t + \frac{1}{6}N_x^{t+1}} \tag{3}$$

3. Comparative analysis

In order to analyse the impact of considering, or not, migration on the estimates of the probability of survival or death at each age x, we have constructed, using the deaths of two adjacent years, the life tables of the years 2006, 2007 and 2008 (for ages 0 to 99 years). The information handled comes from the data that INE offers directly to the public on its website. sex and age (January, 1) Population Now Cast (ePOBa) estimates (for years 2006, 2007 and 2008) have been used as population data (INE, 2012). Death

---

$\sqrt{2}\int_0^1 \left(x - \frac{x^2}{2}\right) dx = \sqrt{2}\left(\frac{1}{2} - \frac{1}{6}\right) = \frac{\sqrt{2}}{3}$, which coincides with the product of the area of the triangle BCF (the *number* of points in BCF), $\frac{1}{2}$, and the length of RQ, the segment of lifeline that goes from the barycentre of BCF to FG, $\frac{2\sqrt{2}}{3}$ (see next footnote).

9. Indeed, taking B as the origin of a Cartesian coordinate system and using that the Lexis squares have unit sides, we have that the coordinates of the points C, F and G are, respectively, $(1,0)$, $(1,1)$ and $(2,1)$ and that, as a consequence, the coordinates of the barycentres R and S are $\left(\frac{2}{3},\frac{1}{3}\right)$ and $\left(\frac{5}{3},\frac{2}{3}\right)$, respectively. Hence, it is not difficult to see that the distance from R to Q is equal to the length of the hypotenuse of a right-angled triangle with right sides both of length $\frac{2}{3}$ and that the segment ST is the hypotenuse of a right-angled triangle of right sides $\frac{1}{3}$.

10. In this case, the formulae used for ages zero and one have been, respectively, $\widetilde{q}_0 = \frac{0.7D_0^t + 0.3D_0^{t+1}}{B^t + \frac{4}{10}N_0^t + \frac{1}{10}N_0^{t+1}}$ and $\widetilde{q}_1 = \frac{\frac{1}{2}(D_1^t + D_1^{t+1})}{C_0^t - 0.3D_0^t + \frac{1}{2}N_0^t + \frac{1}{3}N_1^t + \frac{1}{6}N_1^{t+1}}$; where for the estimation of $\widetilde{q}_0$ it has been assumed that on average the probability of death of a migrant of age zero located in the lower triangle of the period t is four times the probability of death of a migrant of the same age located in the upper triangle of the period $t+1$.

statistics come from sex and age vital statistics (INE, 2010a). And, approximations to sex and age annual immigrant and emigrant figures (for years 2006 to 2009) have been obtained from the data of residential variations (INE, 2010b)[11]. This section shows the differences obtained for the 2007 tables, before and after adjusting estimated crude probabilities. The adjustment has been carried out using nonparametric estimation; in particular, through a Gaussian kernel graduation (see, e.g., Ayuso et al., 2007, pp. 217-22).

Comparisons between values obtained for each age x with HCDS and HODS were carried out by use of two indicators of dissimilarity widely employed in the literature:

- Absolute relative error (ARE)

$$\frac{|\widehat{q}_x - \widetilde{q}_x|}{\widehat{q}_x}, x = 0,1,\ldots,99.$$

- Square relative error (SRE)

$$\frac{(\widehat{q}_x - \widetilde{q}_x)^2}{\widehat{q}_x}, x = 0,1,\ldots,99.$$

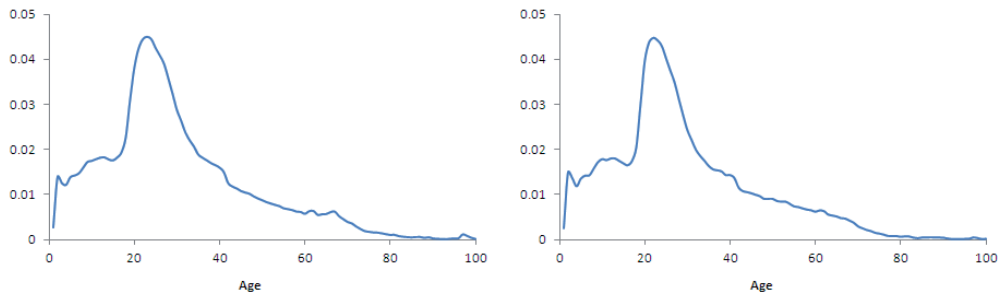

Figure 3: *Differences in crude probabilities with and without migration flows: men (left) and women (right).*

Figure 3 shows, in graphic form, ARE values obtained by comparing the estimates of crude probabilities achieved for men (left) and women (right), after applying equations (2) and (3), with and without migratory flows. As can be seen, the differences of considering HCDS or HODS are significant, reaching the greatest dissimilarities in the range of 14 to 36 years, with the maximum in both cases being reached at age 22, where the difference is close to 4.5%. Similar results are reached when *SRE* is used as measure of dissimilarity: the age range with greater differences remains the same.

11. It should be noted that the statistic of residential variations cannot be observed as a completely true source for migration flows given that this is just an account of the entrants and exits registered on the lists of the municipalities. This has been used, nevertheless, because it represents the only public source that can be used as a proxy of the migration flows occurring in Spain during the period analysed.

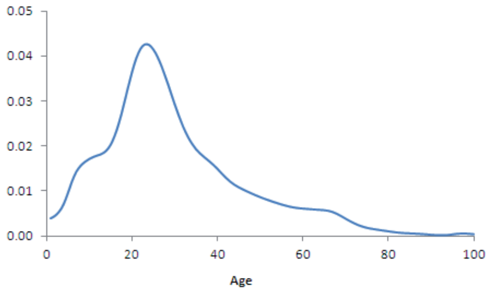

Figure 4: *Differences in graduated probabilities with and without migratory flows: (left) men and (right) women.*

The discrepancies observed reveal, at least in this case, the usual assumption of HCDS being inadequate. This provokes a systematic, not uniform overestimation in the probabilities of death for all ages; which, as can be seen below (Table 1 and Figure 5), has asymmetrical effects on the results of different actuarial calculations.

In actuarial calculations, however, the crude death estimates obtained directly from observed data, $\widehat{q}_x$ or $\widetilde{q}_x$, are not usually used without being graduated first. The objective of graduation is to soften the crude estimates in a way that eliminates (or mitigates) the random fluctuations present in empirical data. In this study, graduation has been carried out using a kernel graduation (see, for example, Ayuso et al., 2007). In particular, the kernel estimation carried out uses a Gaussian function as a kernel with a window parameter, or bandwidth, equal to *1*.[12]

Once the initial values $\widehat{q}_x$ and $\widetilde{q}_x$, were graduated, the indicators of dissimilarity ARE and SRE introduced previously were again calculated. The results obtained for men and women with the ARE measurement are shown in Figure 4. The comparison with the graduated probabilities does not change in any way the conclusions reached previously; in fact they serve to reinforce the results already obtained.

Finally, in line with Pavía and Escuder (2003), some specific probabilities have been obtained with the aim of illustrating the differences that could be derived by using one or other hypothesis on the demographic system: Table 1 shows the results. As was expected of a demographic situation such as that lived in Spain, where in recent years entries have been significantly greater than exits, the non-inclusion of migratory flows underestimates the survival probabilities and overestimates the death probabilities. Differences in every case are evident to the third significant digit in men and (at most) the fourth digit in women. The impact, therefore, is different depending on the gender. In spite of the repercussions on the individual probabilities being similar for both sexes (see Figures 3 and 4), the inclusion of migratory flows has a significantly greater impact on men than on women, at least for the range of ages and periods considered.

12. Similar results were obtained with alternative bandwidth parameters.

Table 1: *Examples of probabilities.*

	Men		Women	
	HCDS	HODS	HCDS	HODS
${}_{25}q_{40}$	.1380204	.1371103	.0584317	.0580243
${}_{15\|10}q_{50}$	.2345362	.2338640	.1112529	.1109601
${}_{85}p_{0}$	.3442274	.3452564	.5672970	.5679357
${}_{20}p_{15}$	.9872419	.9876434	.9951881	.9953219

This asymmetric impact is also clearly visible in Figure 5, where the differences in life expectancy when either migration flows have been or have not been taken into account are shown. As can be observed, the underestimation in life expectancy that entails the non-inclusion of migratory flows is, almost for all ages, double in men than in women.

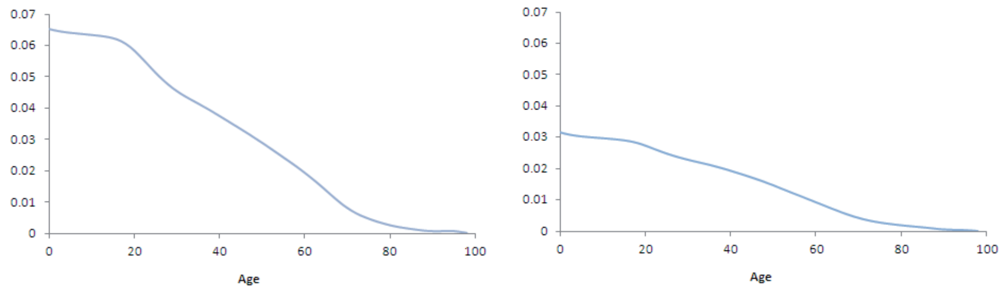

Figure 5: *Differences in life expectancy (in years) with and without migration flows: (left) men and (right) women.*

4. Conclusions

When working with general populations, the usual practice in the construction of life tables consists in ignoring the entry and exit flows that occur in the study group during the analysis period, under the assumption that these usually have little value compared to the size of the population. The use of a closed demographic system hypothesis has been consequently the norm among analysts.

In this paper, (i) the techniques used for estimating death probabilities have been extended to open demographic systems with aggregate data; and, (ii) the resulting estimator has been used, along with the classic HCDS estimator, to obtain (from 0 to 99 years) life tables of the resident population in Spain from 2006 to 2008.

Comparison of crude and graduated probabilities obtained with and without the inclusion of migratory flows shows the impact that entry and exit movements can have on actuarial and demographic calculations. In the examples considered the repercussion has been asymmetric by age and sex. On one hand, the greatest discrepancies, in relative terms, are concentrated in the range of ages from 14 to 36 years, where the intensity of flows has been stronger in Spain in the recent years. On the other hand, by gender, it is clear that the impact on women is less, in contrast to that in men. The well-known lower probabilities of death that women suffer in the range of ages where greater probabilities of migration occur may be behind this result.

The results obtained in this paper point to the need to explicitly consider migratory flows in the estimation of life tables for general populations. The cost to introduce this information is minimal but the potential danger could be significant, especially in situations where entry movements are well above exit movements. In this type of situation, to omit migration flows would lead to an overestimation of probabilities of death and hence to an underestimation of life expectancy, with the danger that this could entail for a correct inter-generational planning that would ensure an adequate stability of the social security pension programmes characteristic of the Western welfare systems.

Certainly, beyond the possible influence that migratory flows and other relevant information might have on results, the great quantity of data provided by modern statistical systems offers an opportunity to seek new ways to exploit the available data in more efficient fashions. So, developing new methodological approaches or implementing proper analyses that help to assess the soundness of broadly used hypotheses should be included early on the statistical demographic research agenda. Along this line, in order to assess the cumulative impact of migration in mortality in Spain, it would be interesting to compare the probabilities of death and life expectancies of born-in-Spain and total (Spaniards and foreigners) populations. Likewise, death and migrant microdata should be analysed in order to ascertain the suitability of the uniform distribution hypotheses required when handling aggregate data.

Acknowledgements

The authors wish to thank M. Hodkinson for translating into English the text of the paper and the anonymous referees their valuable comments and suggestions. The authors acknowledge the support of the Spanish Ministry of Science and Innovation (MICINN) through the project CSO2009-11246.

References

Ayuso, M., Corrales, H., Guillén, M. Pérez-Marín, A. M. and Rojo, J. L. (2007). *Estadística Actuarial Vida*. Barcelona: Publicacions Universitat de Barcelona.

Benjamin, B. and Pollard, J. (1992). *The Analysis of Mortality and Other Actuarial Statistics*. London: Butterworth-Heinemann.

Copas, J. B. and Haberman, S. (1983). Non parametric graduation using kernel methods. *Journal of the Institute of Actuaries*, 110, 135-156.

Forfar, D. O, MacCutcheon, J. J. and Wilkie A. D. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, 115, 1-149.

INE (2007). *Metodología empleada en el cálculo de las tablas de mortalidad de la población de España 1992-2005*. Madrid: INE. Instituto Nacional de Estadística. http://www.ine.es/daco/daco42/mortalidad/metodo_9205.pdf

INE (2009). *Tablas de mortalidad. Metodología*. Madrid: INE. Instituto Nacional de Estadística. http://www.ine.es/daco/daco42/mortalidad/metodo_9107.pdf

INE (2010a). *Movimiento Natural de Población. Defunciones (cifras anuales)*. http://www.ine.es/jaxiBD/tabla.do?per=12&type=db&divi=MNP&idtab=49&L=0

INE (2010b). *Estadística de Variaciones Residenciales (01/02/2011)*. http://www.ine.es/prodyser/micro_varires.htm

INE (2012). *Estimaciones de la Población Actual de España*. http://www.ine.es/jaxiBD/menu.do?L=1&divi=EPOB&his=0&type=db

Livi Bacci, M. (2000). *Introducción a la Demografía*. Barcelona: Ariel.

Pavía, J. M. (2011). *101 Ejercicios Resueltos de Estadística Actuarial Vida*. Madrid: Garceta.

Pavía, J. M. and Escuder, R. (2003). El proceso estocástico de muerte. Diferentes estrategias para la elaboración de tablas recargadas. Análisis de sensibilidad. *Estadística Española*, 45, 243-274.