

# DEMOGRAPHIC RESEARCH

*A peer-reviewed, open-access journal of population sciences*

---

## ***DEMOGRAPHIC RESEARCH***

**VOLUME 35, ARTICLE 14, PAGES 381–398  
PUBLISHED 19 AUGUST 2016**

<http://www.demographic-research.org/Volumes/Vol35/14/>

DOI: 10.4054/DemRes.2016.35.14

*Formal Relationship 26*

### **Symmetries between life lived and left in finite stationary populations**

**Francisco Villavicencio**

**Tim Riffe**

© 2016 *Francisco Villavicencio & Tim Riffe.*

*This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See <http://creativecommons.org/licenses/by-nc/2.0/de/>*

## Table of Contents

1	Introduction	382
2	Rao and Carey's Theorem 1	383
3	Invalidity of the Brouard-Carey equality for finite stationary populations under continuous time	385
3.1	Counterexample to Rao and Carey's Theorem 1	385
3.2	The discrete-time framework	388
4	The Brouard-Carey equality in finite stationary populations	389
4.1	Relationship	389
4.2	Proof	389
5	Discussion	394
6	Acknowledgements	395
	References	396

## **Symmetries between life lived and left in finite stationary populations**

**Francisco Villavicencio<sup>1</sup>**

**Tim Riffe<sup>2</sup>**

### **Abstract**

#### **BACKGROUND**

The Brouard-Carey equality describes the symmetries between the distribution of life lived and life left in stationary populations. This result was formally proved for populations of infinite size and continuous time, and a subsequent attempt to prove it for populations of finite size is invalid.

#### **OBJECTIVE**

We attempt to provide a formal mathematical proof of the Brouard-Carey equality for finite stationary populations.

#### **CONCLUSIONS**

The symmetries between life lived and life left in finite stationary populations can only be proved if time is explicitly discretized. The proof is more complex than in a continuous-time framework, but it conforms with the kinds of data usually available to researchers.

#### **CONTRIBUTION**

The main contribution of this paper is to offer a complete and formal proof of the symmetries between life lived and life left for finite stationary populations in a discrete-time framework. This result is a useful tool for the study of human and non-human populations when the assumption of stationarity is acceptable, especially when subject ages are unknown, but individuals are followed-up until death.

---

<sup>1</sup> Department of Mathematics and Computer Science, Max-Planck Odense Center on the Biodemography of Aging, University of Southern Denmark, Odense, Denmark. E-Mail: villavicencio@imada.sdu.dk.

<sup>2</sup> Max Planck Institute for Demographic Research, Rostock, Germany.

## 1. Introduction

The symmetries between life lived and life left in stationary populations have drawn the attention of several scholars in past years. Examples are found in the fields of demography and population biology (Goldstein 2009, 2012; Kim and Aron 1989; Müller et al. 2004, 2007; Riffe 2015; Vaupel 2009), and reliability engineering (Cha and Finkelstein 2016; Finkelstein and Vaupel 2015). One of the main results that has recently emerged is “Carey’s equality,” as first coined by Vaupel (2009), which establishes that the age composition equals the distribution of remaining lifetimes in stationary populations. To our knowledge, this relationship was first described by Brouard (1989), and was later and independently noticed by James Carey in the study of the survival patterns of captive and follow-up cohorts of med-flies (Müller et al. 2004, 2007). Hence, we choose to refer to this relationship as the ‘Brouard-Carey equality.’<sup>3</sup>

Vaupel (2009) proved that in stationary populations of infinite size and continuous time, the probability that a randomly selected individual is age  $x$  equals the probability that the individual has exactly  $x$  time left until death. Formally,

$$(1) \quad c(x) = g(y) \quad \text{for } x = y,$$

where the probability density function  $c(x)$  describes the age composition of the population, and the probability density function  $g(y)$  gives the distribution of remaining lifetimes. This result has several applications in the study of human and non-human populations with unknown ages. For instance, in capture-recapture studies in which individuals are captured and then followed until death, assuming strict and deterministic stationarity and the absence of other biases, it can be inferred that the unobserved distribution of ages at capture is equal to the observed distribution of the follow-up durations.

Rao and Carey (2015) claim to have an alternative and innovative proof of the Brouard-Carey equality that “does not depend on any of the classical works on renewal equation and renewal theory,” and that is inspired by experimental observations of captive individuals (Rao and Carey 2015: 587–588). Hence, their approach is an attempt to prove the Brouard-Carey equality for empirical and finite stationary populations, rather than theoretical infinite populations, as is the case in Vaupel (2009). However, the use of continuous time in finite populations represents a subtle flaw in Rao and Carey’s work: In order to validate the Brouard-Carey equality for finite stationary populations, time measures need to be explicitly discretized.

In the following section we analyze Theorem 1 by Rao and Carey (2015). Next, we discuss why the symmetries between life lived and life left in finite stationary populations

<sup>3</sup> The relationship was mentioned in Brouard (1986), with a citation to a 1985 version of the 1989 manuscript. We have not been able to locate the 1985 version.

are not valid under continuous time, and that the adoption of a discrete-time framework is imperative. Finally, we suggest a reformulation of the Brouard-Carey equality for finite stationary populations.

## 2. Rao and Carey's Theorem 1

The motivation underlying Rao and Carey's approach to prove the Brouard-Carey equality is the observation of survival patterns in captive med-flies. As stated by the authors, their "theory and method of proof uses sequentially arranged data of captive individuals" (Rao and Carey 2015: 588), which implies empirical – and consequently finite – stationary populations.

If the Brouard-Carey equality were true for finite stationary populations, one would expect the number of individuals of any age  $x$  to equal the number of individuals with  $x$  life left in the same population at any time point. Rao and Carey (2015) adopt a slightly different approach, suggesting the existence of a graphical symmetry between the sets of capture ages and follow-up durations of a finite stationary population.<sup>4</sup> The following is the statement of Theorem 1 by Rao and Carey (2015: 584–585):

**Theorem 1** (Rao and Carey). *Suppose  $(X, Y, Z)$  is a triplet of column vectors, where  $X = [x_1, x_2, \dots, x_k]^T$ ,  $Y = [y_1, y_2, \dots, y_k]^T$ ,  $Z = [z_1, z_2, \dots, z_k]^T$  representing capture ages, follow-up durations, and lengths of lives for  $k$ -subjects, respectively. Suppose,  $F(Z)$ , the distribution function of  $Z$  is known and follows a stationary population. Let  $G_1$  be the graph connecting the co-ordinates of  $S_Y$ , the survival function whose domain is  $N(k) = \{1, 2, 3, \dots, k\}$  i.e. the set of first  $k$  positive integers and  $S_Y(j) = y_j$  for  $j = 1, 2, \dots, k$ . Let  $G_2$  be the graph connecting the co-ordinates of  $C_X$ , the function of capture ages whose domain is  $N(k)$  and  $C_X(j) = x_j$  for all  $j = 1, 2, \dots, k$ . Suppose  $C_X^*(-j) = x_j$  for all  $j = 1, 2, \dots, k$ . Let  $\mathcal{H}$  be the family of graphs constructed using the co-ordinates of  $C_X^*$  consisting of each of the  $k!$  permutations of graphs. Then one of the members of  $\mathcal{H}$  (say,  $H_g$ ) is a vertical mirror image of  $G_1$ .*

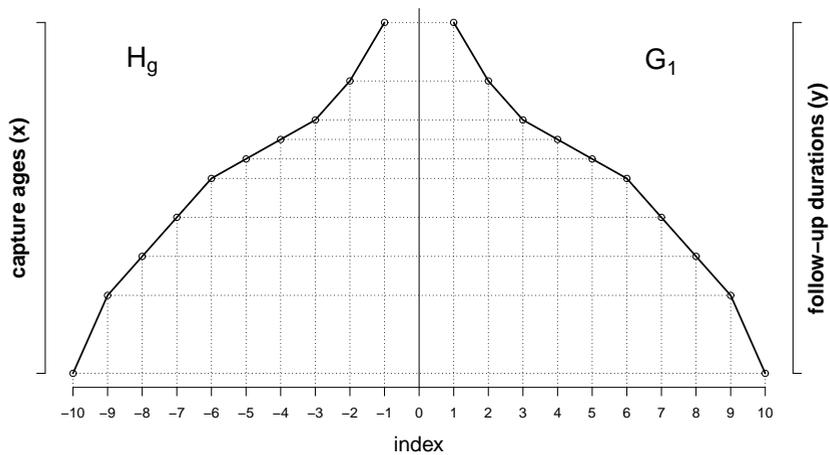
From our reading, Theorem 1 claims that if  $X = (x_1, \dots, x_k)$  are the capture ages of a stationary population of  $k$  individuals, and  $Y = (y_1, \dots, y_k)$  the corresponding follow-up durations, then the elements of  $X$  and  $Y$  can somehow be reordered such that the graph  $G_1$  connecting the coordinates  $\{(j, y_j) : y_j \in Y\}$  and the graph  $H_g$  connecting the coordi-

---

<sup>4</sup> We consider the terms 'age,' 'life lived,' and 'capture age' as synonyms, and the same applies to 'remaining lifetime,' 'life left,' and 'follow-up duration.' For instance, in capture-recapture studies, the terms 'capture ages' and 'follow-up durations' may be more appropriate. The terms 'time to death,' 'remaining lifespan,' or 'residual life' are also frequently used to refer to life left.

nates  $\{(-j, x_j) : x_j \in X\}$  are vertical mirrors of each other. Figure 1 shows an example of graphs  $G_1$  and  $H_g$  as defined by Rao and Carey (2015), for a population of ten individuals.

**Figure 1: Diagram of the graphs  $G_1$  and  $H_g$  as defined by Rao and Carey (2015) for a population of ten individuals**



Note: The right curve in the first quadrant corresponds to  $G_1$ , and it is a decreasing graph representing follow-up durations. The left curve in the second quadrant corresponds to  $H_g$ , and it is an increasing graph representing capture ages.  $G_1$  and  $H_g$  are vertical mirrors of each other with respect to the y-axis.

Unfortunately, Theorem 1 as stated is invalid at this time. The main conceptual flaw of Rao and Carey’s approach is that they do not explicitly account for discrete-time measures, which we discuss in depth in Section 3. Moreover, their paper does not provide a complete proof of the theorem because the existence of a graph  $H_g$  symmetrical to  $G_1$  is assumed, but not proved.

In the last lines of the proof of Theorem 1, the following can be read (Rao and Carey 2015: 586–587):

*Thus, by previous arguments, a family of graphs  $\mathcal{H}$  is constructed. One of these combinations, for example, the  $g^{\text{th}}$  combination, is used to construct a graph which we denote with  $H_g$ , which satisfies the following equalities:*

$$\begin{aligned} |t_1 - |-1|^{(g)}| = 0 \text{ and } |y_{t_1} - x_1^{(g)}| = 0 \\ |t_2 - |-2|^{(g)}| = 0 \text{ and } |y_{t_2} - x_2^{(g)}| = 0 \\ \dots \\ |t_k - |-k|^{(g)}| = 0 \text{ and } |y_{t_k} - x_k^{(g)}| = 0 \end{aligned}$$

here  $\{(|-1|^{(g)}, x_1^{(g)}), (|-2|^{(g)}, x_2^{(g)}), \dots, (|-k|^{(g)}, x_k^{(g)})\}$  are the  $g^{\text{th}}$  combination of cells such that  $H_g$  is a vertical mirror image of  $G_1$ . Image of  $G_1$  is visualized as if it is seen from the mirror kept on  $y$ -axis. Note that we have generated  $k!$  graphs by our construction, and one of such graphs, which we called as  $H_g$  is shown to have a vertical mirror image of  $G_1$ .

In the construction of  $\mathcal{H}$  several graphs are generated, but the authors do not prove the existence of a graph  $H_g \in \mathcal{H}$  that satisfies all the equalities. Though it is not explicitly mentioned, Rao and Carey assume that for all  $y_i \in Y$  there exists  $x_j \in X$ , such that  $x_j = y_i$ . In other words, they assume that each of the values of follow-up durations is identical to exactly one of the values of capture ages. Without this assumption it is not obvious that even for a single element, let's say  $y_{i_1}$ , there exist  $x_i$  and  $g$  such that  $|y_{i_1} - x_i^{(g)}| = 0$  (see counterexample in Section 3.1). However, from Figure 1 it is clear that having a set of follow-up durations, each equal to exactly one capture age, is a necessary and sufficient condition for symmetry between  $H_g$  and  $G_1$ . Statements of set equality and symmetry between sorted sets are equivalent, and the assumption of one cannot be used to prove the other. Hence, the existence of  $H_g$  is assumed but not proved, and the demonstration is therefore incomplete.

### 3. Invalidity of the Brouard-Carey equality for finite stationary populations under continuous time

To demonstrate the invalidity of the Brouard-Carey equality for finite stationary populations under continuous time, we concoct a trivial and reproducible counterexample. We show that under continuous time it is not true that resorting the captures ages and the follow-up durations of a finite stationary population yields two symmetrical graphs,  $G_1$  and  $H_g$ . Therefore,  $H_g$  as constructed by Rao and Carey (2015) does not necessarily exist.

#### 3.1 Counterexample to Rao and Carey's Theorem 1

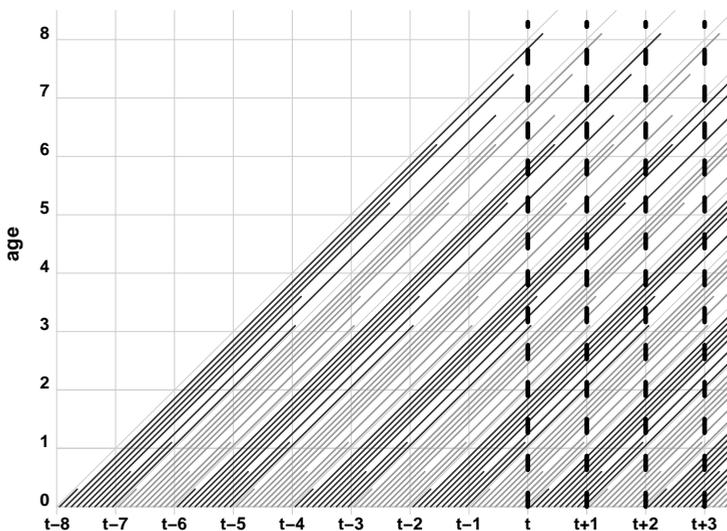
Imagine a long series of identical birth cohorts, each with ten individuals. Every year, exact clones of the same ten individuals are born in the same order, in equal and uniform intervals throughout the year. Individuals within cohorts have varying lifespans, but the exact same pairing of lifespans with individuals is repeated in each year. The order of birth, birth time, and lifespan of cohort members for this example are given in Table 1.

**Table 1: Description of ten individuals in the repeated birth cohort. Birth times are expressed as proportions of a year. Each individual lifespan is identical to an exact age at death**

Individual	1	2	3	4	5	6	7	8	9	10
Birth time	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
Lifespan	0.30	8.10	6.20	7.40	5.20	3.60	0.60	6.70	1.10	3.10

Figure 2 depicts a long series of birth cohorts composed as in Table 1. We refer to the beginning of a year or time-period as the ‘exact time  $t$ ,’ whereas the year itself would be the ‘time-period  $[t, t + 1)$ .’ From Figure 2 we note that the population would achieve a stationary size of 43 at exact time  $t$ , with ten births and ten deaths within every subsequent year or time-period  $[t, t + 1)$ ,  $[t + 1, t + 2)$ , etc. If we take a census (a complete capture) in the stationary part of the series at exact times  $t$ ,  $t + 1$ , etc. (dashed vertical lines in Figure 2), it is possible to record the exact life lived and life left of each of the 43 individuals in the population. One can easily verify from Table 2 that at exact time  $t$ , while there is an individual with age 7.85, there are no individuals in the population with 7.85 years of life left. Analogously, there is an individual ‘captured’ at age 6.85, but no individuals with 6.85 years of life left, and other similar cases can be found. Therefore, in this example – which abides by the conditions set out in Theorem 1 by Rao and Carey (2015) and is even more generous since there is no stochasticity in births and deaths – it is not the case that for each individual with a follow-up duration (life left) of  $y_i$  there is an individual  $j$  whose capture age (life lived) is exactly  $x_j = y_i$ . Consequently, Theorem 1 by Rao and Carey (2015) is invalid, because without this assumption the graphs of  $G_1$  and  $H_g$  would never be symmetrical, irrespective of the sorting process of the vectors of capture ages ( $X$ ) and follow-up durations ( $Y$ ).

**Figure 2: Lexis diagram of a long series of birth cohorts composed as in Table 1**



Note: At exact time  $t$  the population would achieve a stationary size of 43, with ten births and ten deaths within every subsequent time-period. The dashed vertical lines represent population censuses.

**Table 2: Life lived ( $x_j$ ) and life left ( $y_j$ ) data from any census carried out in the stationary part of the series at exact times  $t, t + 1$ , etc., of the stationary population described in Table 1 and Figure 2**

Individual	1	2	3	4	5	6	7	8	9	10	11
$x_j$	7.85	6.85	5.85	4.85	3.85	2.85	1.85	0.85	5.75	4.75	3.75
$y_j$	0.25	1.25	2.25	3.25	4.25	5.25	6.25	7.25	0.45	1.45	2.45
Individual	12	13	14	15	16	17	18	19	20	21	22
$x_j$	2.75	1.75	0.75	6.65	5.65	4.65	3.65	2.65	1.65	0.65	4.55
$y_j$	3.45	4.45	5.45	0.75	1.75	2.75	3.75	4.75	5.75	6.75	0.65
Individual	23	24	25	26	27	28	29	30	31	32	33
$x_j$	3.55	2.55	1.55	0.55	3.45	2.45	1.45	0.45	0.35	6.25	5.25
$y_j$	1.65	2.65	3.65	4.65	0.15	1.15	2.15	3.15	0.25	0.45	1.45
Individual	34	35	36	37	38	39	40	41	42	43	
$x_j$	4.25	3.25	2.25	1.25	0.25	0.15	3.05	2.05	1.05	0.05	
$y_j$	2.45	3.45	4.45	5.45	6.45	0.95	0.05	1.05	2.05	3.05	

### 3.2 The discrete-time framework

In a theoretical population model, one can assume time to be continuous regardless of whether the population is finite or infinite. Hence, it may be possible to compute vital rates and the instantaneous, though theoretical, amount of births and deaths at any time  $t$ . If the population is assumed to be stationary with constant size, then the instantaneous amount of births and deaths are also constant over time. On the contrary, in an empirical and finite population in which individuals are indivisible, the instantaneous amount of births and deaths can possibly be zero, and one would rather compute the number of births and deaths in a time-interval between observations.

These remarks have an important effect on how stationarity is defined in empirical populations. Let's consider the previous example from Section 3.1. If a census is carried out in the middle of a year, let's say at time  $t + 0.3$ , one would observe that four individuals died since time  $t$  (individuals 1, 27, 31, and 40 from Table 2) but only three were born over the same interval (see Table 1). Hence, the population size has not remained invariant. This shows that in an empirical set-up as the one suggested by Rao and Carey (2015), rigorous stationarity can only be detected if the population is observed at regular time points, because the population size may undulate between observations. Therefore, time measures need to be explicitly discretized.

Instead of referring to age and time as continuous variables, let's consider age and time intervals or categories, all of the same width. For instance, all individuals with ages between 0 and 1 are assigned to the age group  $[0, 1)$ , individuals with ages between 1 and 2 are assigned to the age group  $[1, 2)$ , etc.<sup>5</sup> Analogously, individuals with life left between 0 and 1 are assigned to the  $[0, 1)$  life left group, and so on. If we proceed this way with the life lived and life left values of Table 2 we observe that the same eight categories are used for age and life left: from  $[0, 1)$  to  $[7, 8)$ . Moreover, as shown in Table 3, the number of individuals in each age group equals the number of individuals in the corresponding life left group. Certainly, the same values from Table 3 would be obtained in any census carried out in the stationary part of the series at exact times  $t, t + 1$ , etc.

**Table 3: Age and life left data from Table 2 grouped in categories**

Age / life left category	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)	[5,6)	[6,7)	[7,8)
No. individuals in age	9	7	7	7	5	4	3	1
No. individuals in life left	9	7	7	7	5	4	3	1

These observations lead to the following conclusion: If a finite stationary population is observed at regular time-intervals, and age and life left are grouped at intervals of

<sup>5</sup> The same approach is used in matrix population models, in which one also considers an age-classified population, and defines the projection matrix ('Leslie matrix') from exact time  $t$  to exact time  $t + 1$  for each age class (Caswell 2001; Leslie 1945)

the same width, then the Brouard-Carey equality may hold. A more formal statement is presented in the following section.

## 4. The Brouard-Carey equality in finite stationary populations

In keeping with our treatment, we present a reformulation of the Brouard-Carey equality for finite stationary populations in a discrete-time framework that can be understood as complementary to Brouard's and Vaupel's approach to the Brouard-Carey equality (Brouard 1989; Vaupel 2009). However, it must be clearly stated that this result only generalizes to empirical populations in the form of an expectation, even if the underlying birth and death rates are invariant over time. In general, the exact symmetry of theoretical stationary populations does not hold for captured populations, where both the finite wild population and the captured subset of it (the sample) are subject to stochastic variation.

### 4.1 Relationship

Suppose there is a finite stationary population that is observed at regular time points. Assume age, life left, and time are discretely measured, and the same time unit is used for all of them. Let's define,

$N_x(t)$ : number of individuals in age  $[x, x + 1)$  at exact time  $t$ , and  
 $\Omega_y(t)$ : number of individuals with  $[y, y + 1)$  life left at exact time  $t$ .

Then,

$$(2) \quad N_x(t) = \Omega_y(t) \quad \text{for } x = y \text{ and } \forall t.$$

Equation (2) is the Brouard-Carey equality for finite stationary populations.

### 4.2 Proof

The assumption of perfect stationarity implies that the population is closed to migration, and that the population size, the observed age-specific death rates, and the number of newborns within every time-interval ( $B$ ) are constant over time (Preston, Heuveline, and Guillot 2001, Chap. 3). Let's introduce the following terms that are commonly used in demographic analysis, especially when working with the Lexis diagram:

- $B_x(t)$ : number of  $x^{th}$  birthdays in the time-interval  $[t, t + 1)$ . This is analogous to the survival function of the continuous case;

- ${}_L D_x(t)$ : number of deaths in age  $[x, x + 1)$  in the time-interval  $[t, t + 1)$  to individuals who reached age  $x$  in the same interval. This is usually referred to in the literature as the ‘lower Lexis triangle’;
- ${}_U D_x(t)$ : number of deaths in age  $[x, x + 1)$  in the time-interval  $[t, t + 1)$  to individuals who reached age  $x$  in the previous interval. This is usually referred to in the literature as the ‘upper Lexis triangle’; and
- $q_x$ : probability of dying between ages  $x$  and  $x + 1$  given survival to exact age  $x$ .

Death probabilities are given by

$$(3) \quad q_x = 1 - \frac{B_{x+1}(t+1)}{B_x(t)} = \frac{{}_L D_x(t) + {}_U D_x(t+1)}{B_x(t)} \quad \forall x \geq 0 \text{ and } \forall t,$$

and refer to the age-cohort Lexis parallelogram.<sup>6</sup> The conversion from death rates to death probabilities only depends on the age-specific average number of person-years lived in the interval by those dying in the interval, which are usually denoted by  $a_x$  (Chiang 1984; Preston, Heuveline, and Guillot 2001). It is generally accepted in stationary population and life table models that the  $a_x$  values are constant over time, in addition to the age-specific death rates. Therefore, the age-specific death probabilities  $q_x$  can also be assumed to be independent of time.<sup>7</sup>

Since  $B_0(t) = B$  for all  $t$  because the birth flow is constant over time,

$$(4) \quad B_{x+1}(t+1) = B_x(t) (1 - q_x) = B \prod_{n=0}^x (1 - q_n) \quad \forall x \geq 0 \text{ and } \forall t,$$

and therefore the  $B_x(t)$  values do not depend on time either, and can be simply denoted by  $B_x$ . Using this result and (3), one finds that for all  $x$  and at any time  $t$ ,

$$q_x = 1 - \frac{B_{x+1}(t+1)}{B_x(t)} = 1 - \frac{B_{x+1}(t)}{B_x(t)} \implies \frac{{}_L D_x(t) + {}_U D_x(t+1)}{B_x(t)} = \frac{{}_L D_x(t) + {}_U D_x(t)}{B_x(t)}$$

$$(5) \quad \implies {}_U D_x(t+1) = {}_U D_x(t)$$

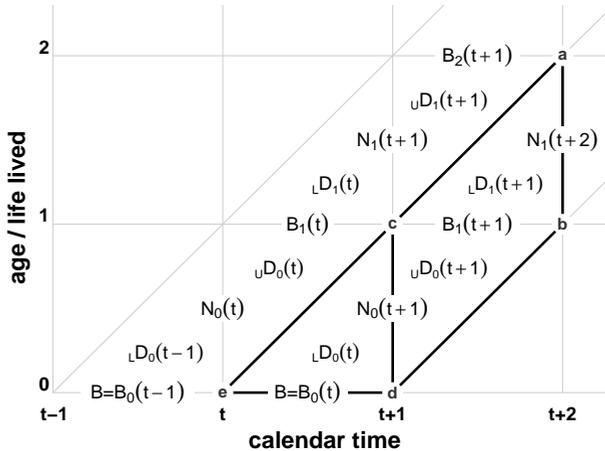
$$(6) \quad \implies {}_L D_x(t+1) = {}_L D_x(t).$$

<sup>6</sup> This line of thought is in accordance with the general acceptance that in the discrete case, the probabilities of death in the interval of discretization are used instead of death rates. Here, we follow the same approach suggested by Preston, Heuveline, and Guillot (2001) in the definition of  $q_x$ , which is identical to the one used by the Human Mortality Database for cohort life tables (Wilmoth et al. 2007). However, in matrix population models, death probabilities – or the analogous survival probabilities – usually refer to the period-cohort Lexis parallelogram, instead of the age-cohort parallelogram. For instance, the elements of the sub-diagonal of the Leslie matrix are the probabilities that an individual with age  $[x, x + 1)$  at exact time  $t$  will be alive in age group  $[x + 1, x + 2)$  at exact time  $t + 1$  (Caswell 2001; Leslie 1945).

<sup>7</sup> For time-intervals of length  $n = 1$ , the conversion from death rates  $m_x$  to death probabilities is given by  $q_x = m_x / (1 - (1 - a_x) \cdot m_x)$ , assuming  $m_x$  is represented by a step function. Thus, if  $m_x$  and  $a_x$  are constant over time (usually  $a_x$  is assumed to be a half time-interval, except for the first age group), so are the  $q_x$  values. See Preston, Heuveline, and Guillot (2001) or Wilmoth et al. (2007) for a more detailed discussion.

Equations (5) and (6) imply that the number of deaths in each age-specific lower and upper Lexis triangle are independent of time, and one can simply write  ${}_L D_x$  and  ${}_U D_x$ .<sup>8</sup> The Lexis diagram in Figure 3 may help visualize all these concepts.

**Figure 3: A Lexis diagram depicting the quantities  $N_x(t)$ ,  $B_x(t)$ ,  ${}_L D_x(t)$ , and  ${}_U D_x(t)$**



Note:  $N_x(t)$  refers to vertical segments in the diagram, as of a population census.  $B_x(t)$  refers to horizontal population segments, representing birthdays, which are typically only directly observed for the birth series.  ${}_L D_x(t)$  refers to the lower Lexis triangle of deaths for a given age, and  ${}_U D_x(t)$  to the upper Lexis triangle. The **abcd** parallelogram depicts the deaths of individuals born in  $[t, t+1)$  that die in  $[t+1, t+2)$  (period-cohort Lexis shape). The **cde** triangle depicts the deaths in  $[t, t+1)$  of individuals born in that same time-period.

From Figure 3, note that  $N_x(t) = B_x(t-1) - {}_L D_x(t-1)$  for all  $x \geq 0$  and for all  $t$ . Since all these terms do not depend on time, one finds that

$$(7) \quad N_x(t) = B_x - {}_L D_x \quad \forall x \geq 0 \text{ and } \forall t.$$

Thus, if we were able to prove that

$$(8) \quad \Omega_y(t) = B_y - {}_L D_y \quad \forall y \geq 0 \text{ and } \forall t,$$

that would demonstrate the Brouard-Carey equality in (2). To do so we proceed by mathematical induction, which is a common proof technique that requires of two steps: 1)

<sup>8</sup> As shown later in the proof, having constant values of  ${}_L D_x$  and  ${}_U D_x$  is a necessary condition for validating the Brouard-Carey equality for finite stationary populations. In this case, this is achieved by assuming that the  $m_x$  and  $a_x$  values are constant over time, but one could have obtained the same result by using other assumptions.

prove that (8) holds for the base case ( $y = 0$ ), and 2) prove that if (8) holds for  $y$  (inductive hypothesis), then it also holds for  $y + 1$ .

1. It is clear from (7) that the number of individuals in age  $[0, 1)$  at exact time  $t$  is  $N_0(t) = B - {}_L D_0$ , that is, the difference between the births occurred in the time-period  $[t - 1, t)$  and the newborns who were born and died in that time-period. On the other hand, the number of individuals with  $[0, 1)$  life left at exact time  $t$  is  $\Omega_0(t) = D(t) - \kappa_0(t)$ , where  $D(t)$  are all the deaths of individuals of all ages that will occur in the time-interval  $[t, t + 1)$ , and  $\kappa_0(t)$  represents the new individuals that will join the population in  $[t, t + 1)$  and die before  $t + 1$ , and hence are not observed at exact time  $t$ . As long as the population size is constant,  $D(t) = B_0(t) = B$ , and since the population is closed, new individuals can only come from births. Therefore,  $\kappa_0(t) = {}_L D_0(t) = {}_L D_0$ , that is, newborns from  $[t, t + 1)$  that die in  $[t, t + 1)$ . As a result,

$$(9) \quad \Omega_0(t) = D(t) - \kappa_0(t) = B - {}_L D_0 = N_0(t) \quad \forall t.$$

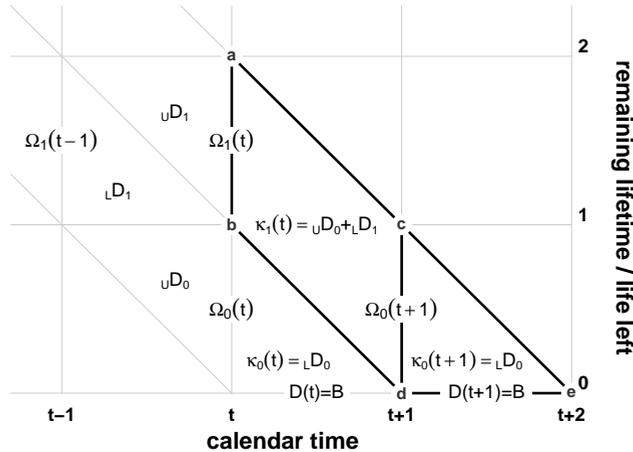
2. Second, let's assume that (8) holds for  $y$ . It needs to be proved that it also holds for  $y + 1$ .

The individuals with  $[y + 1, y + 2)$  life left at exact time  $t$  will have  $[y, y + 1)$  life left at exact time  $t + 1$ . Therefore,  $\Omega_y(t + 1) = \Omega_{y+1}(t) + \kappa_{y+1}(t)$ , where  $\kappa_{y+1}(t)$  are the individuals that join the population in  $[t, t + 1)$  and at exact time  $t + 1$  have  $[y, y + 1)$  life left. Hence, those individuals later die in the time-interval  $[t + y + 1, t + y + 2)$ . As new individuals can only enter the population via birth flow, analogous to the previous case,

$$(10) \quad \kappa_{y+1}(t) = {}_U D_y(t + y + 1) + {}_L D_{y+1}(t + y + 1) = {}_U D_y + {}_L D_{y+1},$$

i.e., newborns from  $[t, t + 1)$  that die in  $[t + y + 1, t + y + 2)$  before reaching exact age  $y + 1$  ( ${}_U D_y$ ), or after reaching exact age  $y + 1$  ( ${}_L D_{y+1}$ ). In general,  $\kappa_y(t)$  is equal to the period-cohort Lexis parallelogram of deaths of the birth cohort born in  $[t, t + 1)$  and dying in  $[t + y, t + y + 1)$ . Figure 4 shows a life-left diagram that may provide some intuition to all these concepts.

**Figure 4:** A life-left diagram, depicting two complete calendar time-periods on the x-axis and the final two remaining-lifetime periods on the y-axis



Note: Diagonal lines delineating death cohorts descend downward. Interval segments on the x-axis represent the death flow, which in this case is equal to the birth flow. Vertical segments represent discrete values of  $\Omega_y$ . The increment that accrues to a death cohort passing from  $\Omega_1(t)$  to  $\Omega_0(t+1)$ ,  $\kappa_1(t)$ , is marked by the parallelogram **abdc** (equivalent to the parallelogram **abdc** in Figure 3), while the increment that accrues in the final time-period of life,  $\kappa_0(t+1)$ , is marked by the triangle **cde**.

Note that at any time  $t$  and for any  $y \geq 0$ ,  $B_y = B_{y+1} + UD_y + LD_y$  (see Figure 3). Besides, the inductive hypothesis implies that  $\Omega_y(t+1) = B_y - LD_y$ . Hence,

$$\begin{aligned}
 \Omega_{y+1}(t) &= \Omega_y(t+1) - \kappa_{y+1}(t) = B_y - LD_y - UD_y - LD_{y+1} \\
 (11) \quad &= B_{y+1} + UD_y + LD_y - LD_y - UD_y - LD_{y+1} \\
 &= B_{y+1} - LD_{y+1},
 \end{aligned}$$

which proves (8). Given that the right-hand terms of (11) do not depend on time, this result stands for all time  $t$ .

Equations (7) and (8) imply that  $N_x(t) = \Omega_y(t)$  for all  $x = y$  and for all time  $t$ , thereby proving (2). Q.E.D.

## 5. Discussion

Rao and Carey (2015) introduce a novel approach to the Brouard-Carey equality for finite stationary populations. As stated by the authors, their work does not depend on classical renewal theory, and it is inspired by experimental observations. Unfortunately, the proof of their Theorem 1 stands invalid at this time, which jeopardizes results that depend on that theorem. In Section 3 we have provided a counterexample that invalidates the theorem as stated.

Our suggested reformulation of the Brouard-Carey equality is intuitively simpler, and takes into consideration that stationarity in finite and empirical populations can only be observed if time measures are explicitly discretized. Using discrete life table functions, we have proved that in non-stochastic stationary populations and in a discrete-time framework, at any time  $t$  the number of individuals in age interval  $[x, x + 1)$  equals the number of individuals in the life left interval  $[x, x + 1)$ . Note that this set-up allows for seasonality of the distribution of births and deaths, as long as the death counts in each Lexis triangle are constant over time. Thus, the Brouard-Carey equality in Eq. (2) applies both for populations that experience reproductive periods at concentrated points in the time-interval (birth pulses) – as is the case in many non-human populations – as well as populations that do not have seasonal reproduction, and whose birth flow is approximately uniformly distributed.

Brouard (1989) and Vaupel (2009) showed that in continuous time and with a stationary population of infinite size, the probability that a random individual is  $x$  years old equals the probability the individual has  $x$  years of life left. In this sense, both results – the continuous and the discrete cases – are complementary and prove the existence of a symmetry between the age distribution and the distribution of remaining lifetimes in stationary populations. Demographers often think of time as a continuous process, which permits the use of differential calculus on demographic functions, and facilitates the proof of demographic relationships among these functions. However, in daily work, researchers usually deal with data collected at discrete-time points, such as surveys, censuses, or capture-recapture data. Therefore, one of the main contributions of this paper is to offer, under certain common assumptions, a clear and formal proof of the Brouard-Carey equality for empirical populations in a discrete-time framework. The discretization of time provides a more realistic reference, but makes the demonstration more complex than in the continuous case.

Following Rao and Carey's idea, our inspiration also comes from an empirical set-up in which capture ages are unknown, but individuals are followed until death, which is a common scenario in many studies on wild animals (Müller et al. 2004, 2007). However, it also applies to human populations when the ages of individuals are unknown, as may be the case in historical data or other kinds of incomplete demographic data. For instance, in the Barcelona Historical Marriage Database – which collects information about marriage

licenses of Barcelona (Spain) from the mid-15<sup>th</sup> century until the early 20<sup>th</sup> century – individuals are first identified in their marriage record and then followed up, but no information is available about their birth date or their age at marriage (Villavicencio, Jordà, and Pujadas-Mora 2015).

Perfect stationarity is rarely observed, but it is often a useful assumption if the error implied is not overly distortive. The Brouard-Carey equality serves as a useful rule of thumb when the assumption of stationarity is acceptable for the population universe, and when the sampling of the population at large is unbiased with respect to age and lifespan.

## **6. Acknowledgements**

We are both grateful to Jim Vaupel for arousing and encouraging our interest in formal demography, and for his support in this project. We would also like to thank Fernando Colchero, Josh Goldstein, and two anonymous reviewers for critical and helpful comments. Part of this work was carried out while TR was at the Department of Demography, University of California, Berkeley. This research was partially supported by a grant from the SDU eScience Center and by grants No. R01-AG011552 and R01-AG040245 of the U.S. National Institute on Aging of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

## References

- Brouard, N. (1986). Structure et dynamique des populations: La pyramide des années à vivre, aspects nationaux et exemples régionaux. *Espace, populations, sociétés* 4(2): 157–168. doi:10.3406/espos.1986.1120.
- Brouard, N. (1989). *Mouvements et modèles de population*. Yaoundé: Institut de Formation et de Recherche Démographiques (IFORD).
- Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation*. Sunderland, MA: Sinauer.
- Cha, J.H. and Finkelstein, M. (2016). On stochastic comparisons for population age and remaining lifetime. *Statistical Papers*: 1–15. doi:10.1007/s00362-016-0759-6.
- Chiang, C.L. (1984). *The life table and its applications*. Malabar, FL: Robert E. Krieger.
- Finkelstein, M. and Vaupel, J.W. (2015). On random age and remaining lifetime for populations of items. *Applied Stochastic Models in Business and Industry* 31(5): 681–689. doi:10.1002/asmb.2073.
- Goldstein, J.R. (2009). Life lived equals life left in stationary populations. *Demographic Research* 20(2): 3–6. doi:10.4054/DemRes.2009.20.2.
- Goldstein, J.R. (2012). Historical addendum to “Life lived equals life left in stationary populations”. *Demographic Research* 26(7): 167–172. doi:10.4054/DemRes.2012.26.7.
- Kim, Y.J. and Aron, J.L. (1989). On the equality of average age and average expectation of remaining life in a stationary population. *SIAM Review* 31(1): 110–113. doi:10.1137/1031005.
- Leslie, P.H. (1945). On the use of matrices in certain population mathematics. *Biometrika* 33(3): 183–212. doi:10.1093/biomet/33.3.183.
- Müller, H.G., Wang, J.L., Carey, J.R., Caswell-Chen, E.P., Chen, C., Papadopoulos, N., and Yao, F. (2004). Demographic window to aging in the wild: Constructing life tables and estimating survival functions from marked individuals of unknown age. *Aging Cell* 3(3): 125–131. doi:10.1111/j.1474-9728.2004.00096.x.
- Müller, H.G., Wang, J.L., Yu, W., Delaigle, A., and Carey, J.R. (2007). Survival and aging in the wild via residual demography. *Theoretical Population Biology* 72(4): 513–522. doi:10.1016/j.tpb.2007.07.003.
- Preston, S.H., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Oxford: Blackwell.

- Rao, A.S.R.S. and Carey, J.R. (2015). Generalization of Carey's equality and a theorem on stationary population. *Journal of Mathematical Biology* 71(3): 583–594. doi:10.1007/s00285-014-0831-6.
- Riffe, T. (2015). The force of mortality by years lived is the force of increment by years left in stationary populations. *Demographic Research* 32(29): 827–834. doi:10.4054/DemRes.2015.32.29.
- Vaupel, J.W. (2009). Life lived and left: Carey's equality. *Demographic Research* 20(3): 7–10. doi:10.4054/DemRes.2009.20.3.
- Villavicencio, F., Jordà, J.P., and Pujadas-Mora, J.M. (2015). Reconstructing lifespans through historical marriage records of Barcelona from the sixteenth and seventeenth centuries. In: Bloothoof, G., Christen, P., Mandemakers, K., and Schraagen, M. (eds.). *Population reconstruction*. Cham: Springer: 199–216. doi:10.1007/978-3-319-19884-2\_10.
- Wilmoth, J.R., Andreev, K., Jdanov, D., Gleijer, D.A., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., and Vachon, P. (2007). *Methods protocol for the Human Mortality Database*. Tech. rep., Berkeley: University of California, and Rostock: Max Planck Institute for Demographic Research.

