EUROPEAN COMMISSION
EUROSTAT

Directorate F: Social statistics
Unit F-2: Population and migration

Luxembourg, 14 April 2020
ESTAT/F-2/GL*

TECHNICAL NOTE

Subject: an alternative life table based on probabilities of dying within the calendar year (annual period-cohort life table)

1. INTRODUCTION

Life tables and especially its prominent indicator – the life expectancy at birth – are among the most important indicators computed by statistical offices. The traditional computation of a life table starts from a set of period age-specific mortality rates, which are transformed in age-specific probabilities of dying before the next birthday. This is the standard approach taken by Eurostat as well in the currently published life tables[1].

This choice was made at a time in which life table computations were based on data provided on voluntary basis by the countries, for which it was not always possible to comply with the data structure suggested by Eurostat. Such data heterogeneity led to choose a life table methodology simple but robust enough to be applicable across a variety of mortality profiles, even with rather limited data. The entry into force in 2013 of the EU regulation on demographic statistics[2] has opened new perspectives, for it has ensured the regular provision to Eurostat of basic demographic data with agreed structure.

After a few years of implementation of the above-mentioned EU regulation it is now available a dataset, complying with its requirements, which is still limited in size but large enough for testing purposes. Taking advantage of these first data, this note explores an option for improving the calculation of the life table. Given the current structure of the data, the study refers only to 'complete' life tables (i.e., by single age) to be produced on

* Giampaolo Lanzieri (giampaolo.lanzieri@ec.europa.eu). This note has not been officially edited and it is released to inform interested parties about ongoing work. The information and views set out in this note are those of the author and do not necessarily reflect the official opinion of the European Commission / Eurostat.

[1] See the "Description of the Eurostat method for the calculation of the life expectancies at all ages". https://ec.europa.eu/eurostat/cache/metadata/Annexes/demo_mor_esms_an1.pdf

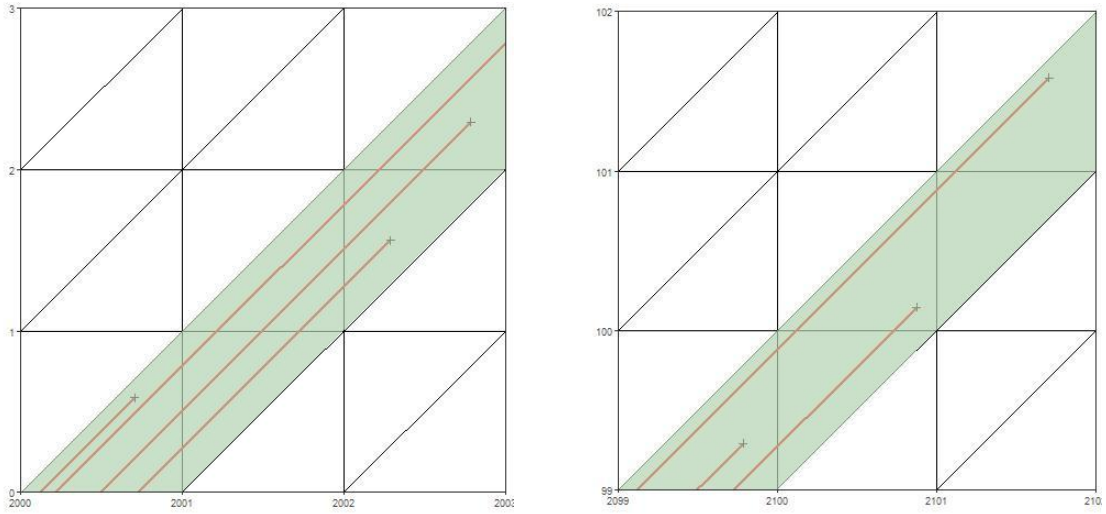
[2] Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European demographic statistics (text with EEA relevance) and its implementing measures.

annual basis from data referring to calendar years. Further, this note is merely an introductory analysis to more extensive investigations on alternative methods of computation of the life table.

2. BASIC FEATURES IN LIFE TABLE

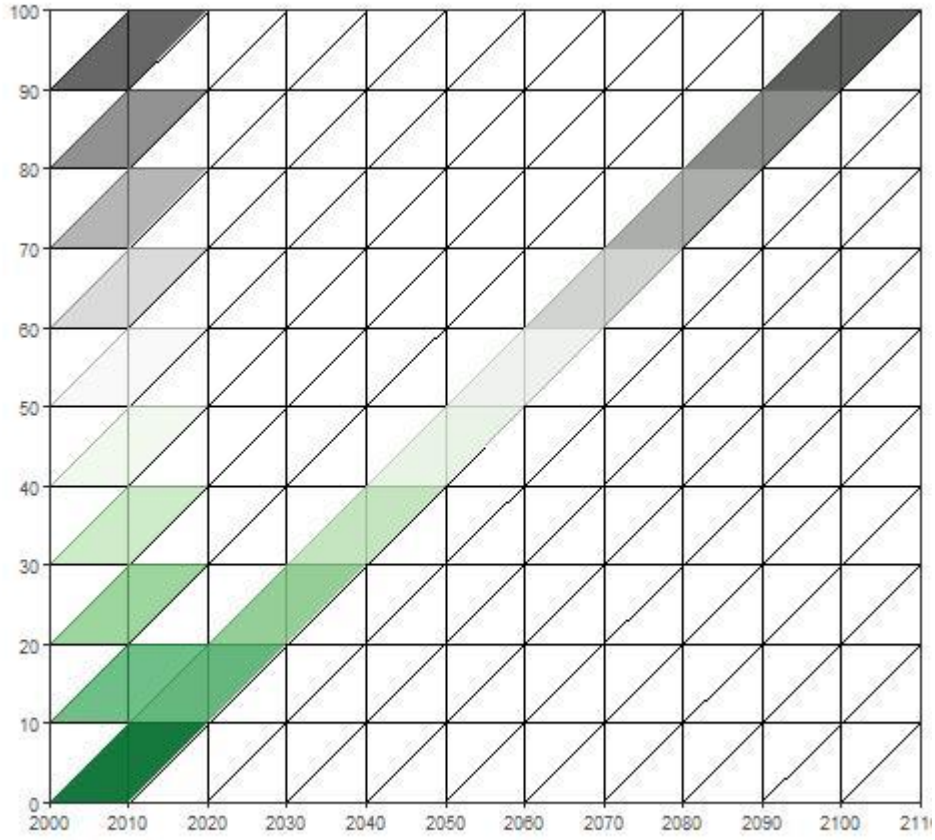
The basic idea of a life table is to follow the mortality of a cohort until its extinction. For instance, the cohort born in the year 2000 would experience deaths in the same year, in 2001, 2002 (see the left panel of the Figure 1), and so forth, until its complete extinction (see the right panel of the Figure 1).

Figure 1: theoretical cohort with examples of individual lifelines (in red)



The mortality experience of such cohort can be broken down in age-specific horizontal sides parallelograms, so-called 'age-cohort' parallelograms because uniquely identified by the 'age' (i.e., more precisely, by the age span between exact age x and the next exact age $x + 1$ excluded) and by the 'cohort' (i.e., the year of birth of the selected generation), so that the deaths at age x occurring into a selected parallelogram can be related to its population at exact age x to measure the probability of dying before reaching exact age $x + 1$. This is represented by the coloured diagonal corridor shown in the Figure 2 (to easier the visualization, ages are there represented as 10-year age groups).

Figure 2: shift of age-cohort parallelograms

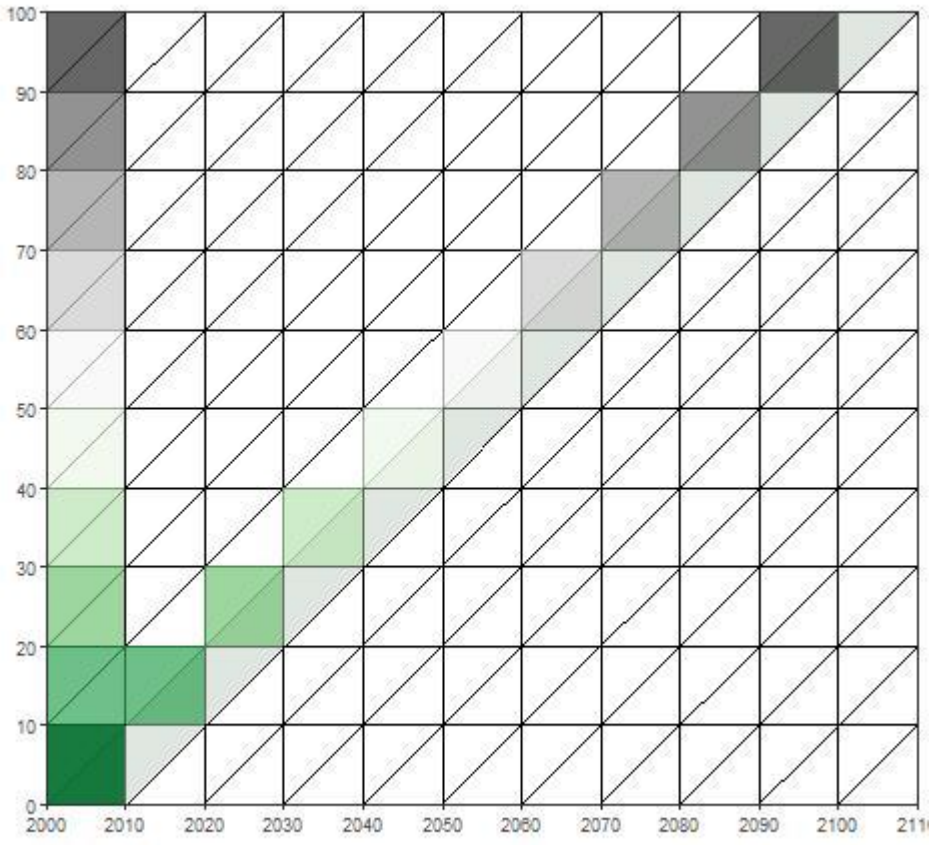


Because computing a cohort life table requires an observation period spanning about a century, the usual stratagem is to refer to a synthetic cohort who would experience the same mortality as observed in the selected year. This means shifting rightwards the period-cohort parallelograms, for as many years as the difference between the period and the cohort. For instance, the parallelogram containing the deaths at age 1 in 2000-2001 will be representing the deaths occurred in 2001-2002 to the generation born in the year 2000, thus with a shift of 1 year; the parallelogram containing the deaths at age 2 in 2000-2001 will be representing the deaths occurred in 2002-2003 to the generation born in the year 2000, thus with a shift of 2 years; and so forth. In the Figure 2 this is shown for 10-year age groups: the observed period mortality in the parallelograms of the coloured vertical corridor is assumed to represent the mortality in the coloured diagonal corridor. Whilst at younger ages such assumption sounds reasonable, at older ages is less robust, for it implies that the mortality observed today would apply in several decades (up to over a century) from now. Nonetheless, this is the ploy regularly adopted by demographers to obtain a timely estimate of the life expectancy (particularly at birth) and this shifting feature is commonly overlooked.

There is a further aspect to be taken into account. The age-cohort parallelograms require an observation extended over 2 years (at least). In a time of annually available data and high demand for timeliness, this is considered an undesirable feature. A common alternative is to group the death events in the Lexis squares identified by the age and the period (i.e., the year of the observations). By doing so, only the mortality data in the selected period (i.e., the calendar year) are necessary to estimate the life expectancies of

the generation born in the same period. However, here the shift rightwards of the age-period squares would not fully overlap with the cohort mortality of that same generation. As shown in the Figure 3, the squares would actually be partially representative of the mortality experience of two cohorts once shifted rightwards.

Figure 3: shift of age-period squares



The approach based on age-period is widely used[3] and perhaps the most used worldwide. This might be a heritage of a time in which there was no possibility of getting detailed information on deaths, such to distribute them according to Lexis triangles, i.e. by age, period and cohort. Further, the easiest way to collect deaths data is by age 'completed', which would justify the predominant approach based on age-period mortality rates.

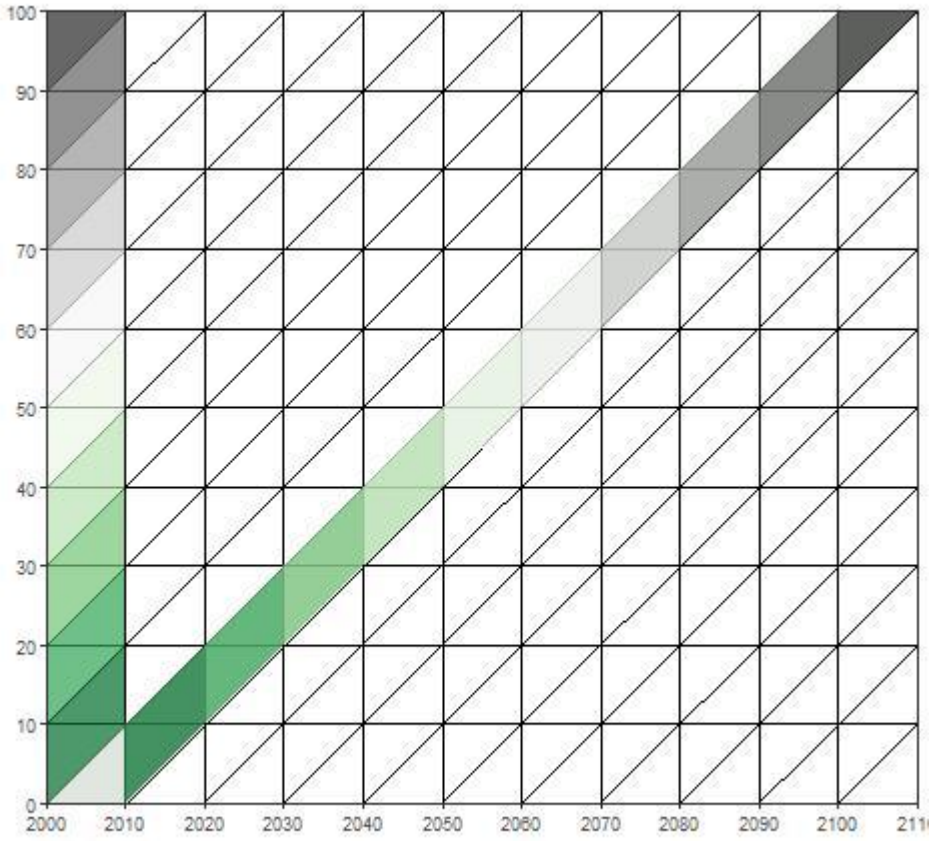
In fact, as for the way age is measured, in demography there are two main concepts: 'age reached during the year' (also known as 'age at the end of the year') and 'age completed' (also known as 'age at last birthday'). The former can be computed by the simple difference between the year of reference and the year of birth, while the latter is the number of completed calendar years between the date of reference and the birthdate. According to a Eurostat survey carried out in 2013[4], all (44) European countries compute death rates by age completed, while only a subset of countries compute them also by age

[3] See Eurostat (2015): "Demographic statistics: a review of definitions and methods of collection in 44 European countries".

[4] Ibid.

reached. Yet, the period-cohort observational plan is more 'natural' to the demographic dynamic, as it avoids the concurrent presence of demographic events from two different cohorts in the same observation period (as shown as well in the Figure 3) and put in a clear fashion the ageing over time. This would lead to a breakdown of the mortality experience of a given cohort in parallelograms with vertical sides, i.e. period-cohort parallelograms – but the one at age 0, which would be a triangle. Their shift, as shown by the diagonal corridor in the Figure 4, would perfectly overlap with the mortality of the cohort born in the same calendar year.

Figure 4: shift of period-cohort parallelograms



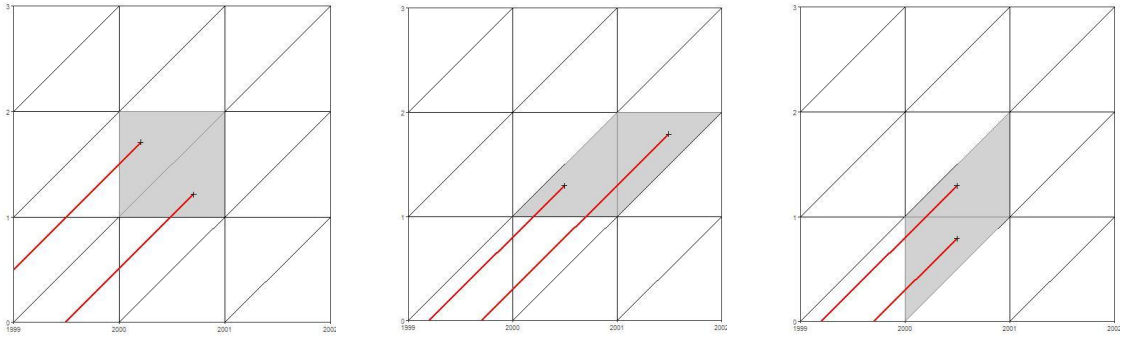
In the synthetic cohort, whose grouping of the deaths is either based on age-cohort parallelograms or on age-period squares, the functions of the life table hold the same meaning as in the cohort life table: m_x is the mortality rate[5] between exact ages x and $x+1$ (more precisely, in the interval $[x, x+1)$), q_x is the probability of dying within the same age range, p_x is the complementary probability of surviving from exact age x to exact age $x+1$, l_x are the number left alive at exact age x (the 'survivors'), a_x are the average person-years lived in the interval by those dying, L_x are the person-years lived within the age interval $[x, x+1)$, T_x are the person-years lived above exact age x and e_x

[5] For the sake of precision, M_x is the set of observed period age-specific death rates and m_x are the corresponding rates in the hypothetical cohort. The former are derived from the events in the squares of the coloured vertical corridor in the Figure 3, the latter are those corresponding to the squares in the coloured diagonal corridor in the same figure. It is usually assumed that $m_x = M_x$.

are the expectation of life at exact age x. The functions l_x, T_x and e_x refer to an exact age, the others to age intervals.

When the input data are observed period age-specific mortality rates M_x, the most cumbersome step in the computation of the so-called period life tables is their conversion to probabilities of dying[6]. Mortality rates M_x are however not the only option for data input to the life table. When deaths data are grouped by age-cohort parallelograms in principle it is possible to compute directly the probabilities of dying in the selected age interval. The Figure 5 shows an example for the age 1: in both the age-period square (in grey in the left panel) and the age-cohort parallelogram (in grey in the central panel), all the deaths there included the line of age 1 had already been crossed by the individual lifeline; with input data grouped by period-cohort parallelograms (in grey in the right panel), this is not necessarily the case, as it may cover deaths occurred at ages lower than 1, but those individuals would have reached age 1 by the end of the selected year. These aspects are further exploited in the next section.

Figure 5: comparison of input data for life table



3. AN ALTERNATIVE APPROACH

Taking advantage of the higher granularity now regularly available in deaths data, this alternative method bypass the issues related to the conversion from mortality rates to probabilities of dying by estimating directly these latter. For sake of simplicity, the method is here presented in its simplest version[7].

3.1. Estimation of the probabilities of dying

Assuming a period-cohort observational plan, the probability of dying q in the year t (the period) at age (reached) x for the cohort c is:

$$q_{c,x}(t-1,t] = \frac{D_{c,x}(t-1,t]}{C[t-1]} \qquad [1]$$

[6] There are also other problematic steps, such as the estimation of the person-years of exposure to the risk of dying to be used in the rate.

[7] This means for a closed population, without considerations for smoothing, or estimates of missing values, or graduation of the average person-years lived in the interval by those dying in the year, etc.

where $D_c(t-1,t]$ are the deaths occurring in the selected cohort from the beginning to the end of the year t and $C[t-1]$ is the cohort size at the end of the previous year. Dropping the information about the cohort, the [1] can be reformulated in terms of age and period alone:

$$q_x(t-1,t] = \frac{D_x(t-1,t]}{P_{x-1}[t-1]} \qquad [2]$$

where $D_x(t-1,t]$ are the deaths occurring during the year t to persons who would have been of age x at the end of the same year, and $P_{x-1}[t-1]$ is the population size of persons aged $x-1$ at the end of previous year who thus could reach age x at the end of the reference calendar year t.

Particular attention must be paid to the tails of the age distribution. At the upper end, the probability of dying in the open-ended age class ω is:

$$q_\omega(t-1,t] = \frac{D_\omega(t-1,t]}{P_{\omega-1}[t-1] + P_\omega[t-1]} \qquad [3]$$

i.e. the events in the open-ended age class are attributable to persons aged $\omega-1$ and above.

At the bottom end of the age distribution, the population that will reach age 0 at the end of the year is actually the cohort of new-born children. The [2] thus becomes:

$$q_0(t-1,t] = \frac{D_0(t-1,t]}{B(t-1,t]} \qquad [4]$$

where $B(t-1,t]$ are the live births occurring during the year t and $D_0(t-1,t]$ are the deaths occurring in the lower Lexis triangle at the bottom of the age distribution.

Putting all together and setting $\omega = 100+$, the estimated probabilities of dying are:

$$\begin{cases} q_0(t-1,t] = \dfrac{D_0(t-1,t]}{B(t-1,t]} & x = 0 \\ q_x(t-1,t] = \dfrac{D_x(t-1,t]}{P_{x-1}[t-1]} & x \in I[1,99] \\ q_{100+}(t-1,t] = \dfrac{D_{100+}(t-1,t]}{P_{99}[t-1] + P_{100+}[t-1]} & x = 100 \end{cases} \qquad [5]$$

Values of q_x are by construction bounded within the interval [0, 1].

There is a distinction between probabilities of dying q_x in the two approaches to life table computation: in the traditional one, q_x is the probability of dying *during one year of ageing*, in the alternative approach here presented is the probability of dying *within the calendar year*.

3.2. Notation about the time

Using time indicators with age reached requires some carefulness in the notation.

The end of the year is considered the midnight of 31 December. At this moment in time, age completed and age reached are perfectly equal, that is why for the population at the end of the year this is simply referred as 'age'. It is often the case that the population at the end of the year is disseminated as 'population on 1 January'; however, as soon as the midnight of 31 December has passed, all the persons have an 'age reached' increased by one year, even though this is not reflected in the disseminated population data. Therefore, here the population on 1 January is equal by definition to the population at the end of the previous year ($P_{31.12.t-1} \equiv P_{01.01.t}$).

Likewise, the time interval $(t-1, t]$ is in fact $(31.12.t-1, 31.12.t]$, that, noting the open interval on the left, can be reformulated as $[01.01.t, 31.12.t]$, which can be indicated in short as year *t*.

Bearing in mind that some components refer to a point in time, while others to time intervals, the notation can be consequently eased. For instance, for the calendar year *t* the [5] becomes:

$$\begin{cases} q_{t,0} = \dfrac{D_{t,0}}{B_t} & x = 0 \\ q_{t,x} = \dfrac{D_{t,x}}{P_{t,x-1}} & x \in I[1,99] \\ q_{t,100+} = \dfrac{D_{t,100+}}{P_{t,99} + P_{t,100+}} & x = 100 \end{cases} \quad [6]$$

where the subscript *t* means either the entire calendar year or the 1 January of the same year (first moment after the end of previous year), as appropriate, and it could also be dropped.

3.3. Computation of the alternative life table

Starting from the q_x as defined in [5], it is immediate to derive the probability of surviving to the end of the selected year as input to the life table. The usual assumption here is that:

$$q_{t,x} = q_{t+x,x} \quad [7]$$

i.e., the probability of dying at age (reached) x in the year t is equal to the probability of dying at the same age in the year $t + x$ (the shift rightwards in the Lexis diagram). Unlike the conventional life table, here $q_{t,\omega}$ is not set by definition equal to one because not everybody in the open-ended age group must die in the given year.

Because $q_x \in [0,1]$, its complement to 1 is also bounded within the same interval:

$$p_{t,x} = 1 - q_{t,x} \quad [8]$$

The number of survivors $l_{t,x}$ of a cohort of arbitrary size (the so-called *radix*) at the end of the selected year *t* is then easily derived by:

$$l_{t,x} = p_{t,x} * l_{t-1,x-1} \quad [9]$$

Likewise q_x, l_x has here a different meaning as well: it is not anymore the number of persons reaching age *x*, but instead they are the number of persons aged *x* who have reached the end of the year. Therefore, they are not the *survivors at age x* (an horizontal segment in the Lexis diagram), but the *survivors aged x at the end of the year* (a vertical segment in the Lexis diagram). In a cohort perspective, it would be:

$$\begin{cases} l_{t+1,0} = p_{t,0} * radix \\ l_{t+2,1} = p_{t+1,1} * l_{t+1,0} \\ l_{t+3,2} = p_{t+2,2} * l_{t+2} \\ \vdots \\ l_{t+x+1,x} = p_{t+x,x} * l_{t+x,x-1} \\ \vdots \end{cases} \quad [10]$$

or, taking as example the cohort born in the year 2000:

$$\begin{cases} l_{1.1.2001,0} = p_{2000,0} * radix \\ l_{1.1.2002,1} = p_{2001,1} * l_{1.1.2001,0} \\ l_{1.1.2003,2} = p_{2002,2} * l_{1.1.2002,1} \\ \vdots \\ l_{t+x+1,x} = p_{t+x,x} * l_{t+x,x-1} \\ \vdots \end{cases} \quad [11]$$

but because of the assumption of the invariance over time $l_{t,x} = l_{t+x,x}$ mentioned above, the [9] becomes:

$$l_{t,x} = p_{t,x} * l_{t,x-1} \quad [12]$$

The fact that the l_x can be seen as one or the other of the vertical lines in a Lexis diagram along the corridor over the selected year, i.e. either as $l_{t-1,x}$ or as $l_{t,x}$, may be source of confusion as for the time index. In fact, as shown in the [10] and [11], the survivors at the end of the year (vertical line on the right) become the new base for the next-age probability of surviving (vertical line on the left)[8]. In a life table it is therefore possible to list the $l_{t,x}$ starting from the radix (thus as vertical line on the left) or as in the cohort perspective (thus as vertical line on the right). Following the practice of having the radix listed in the life table column of the l_x, it is here proposed to report the $l_{t-1,x}$. This approach would allow an age alignment for the sequence of computations in the life table as well as to have the "correct" l_x at the closing of the life table. The disadvantage is that they are not exactly the survivors aged *x* at the end of the year, but rather a synthetic cohort on 1 January. Once clarified this point (which will also be illustrated in an example below), the notation can be further simplified and the notation about the year of reference *t* can be dropped. Hereinafter, l_0 is the radix and the following l_x are the synthetic base population of age *x* for the probability of surviving, which leads to replace the [12] with:

$$l_{x+1} = p_x * l_x \quad [13]$$

The number of life table deaths is then:

[8] This can be imagined as a ball rebounding off two opposite walls, and climbing of one level each two rebounds.

$$d_x = l_x - l_{x+1} \quad [14]$$

As for the open-ended age class, by definition it is $d_\omega \equiv l_\omega$ because everybody eventually dies. The forcing here is that all those deaths are virtually attributed to a single year, while they could be spanned over several. This works well enough when the life table is extended to the oldest ages observed in the population, even if closed before of the complete extinction of the synthetic cohort[9].

It is now possible to compute the person-years lived in the year. They are the sum of the person-years lived by the survivors ($l_{x+1} \cdot 1$) and of the person-years lived by those dying in the selected year. This latter quantity is a function of the average time *a* lived by those persons in the same calendar year of their death. For all ages higher than 0, a common simplifying assumption[10] is to set $a_x = 0.5$, i.e. those persons have lived half a year on average, meaning an uniform distribution of deaths. Therefore, for $x \geq 1$:

$$L_x = l_{x+1} + 0.5 \cdot d_x \quad [15]$$

As for the age 0, the average time spent in life is usually shorter because infant mortality is concentrated in the first period of life. For the sake of simplicity and to easier the comparison with the conventional approach, it is here adopted the same value as in the current Eurostat life table, that is $a_0 = 0.2$.

It should be noted that at the open-ended age class there are by definition no survivors at the end of the year and therefore the first quantity of the right side of the [15] is also zero by definition.

$$L_\omega = a_\omega \cdot d_\omega \quad [16]$$

A first option would seem to set the value of the parameter a_ω again at 0.5, but that would mean that all those dying after reaching the oldest age (i.e., everybody) will live further 6 months only, which seems a too conservative option especially when the life table is not closed at the highest observed age. Otherwise, the average time lived in the open-ended age group by those surviving at the age ω, can be approximated by dividing that number of survivors by the probability of dying in the calendar year for that specific age group, that is:

$$L_\omega = \frac{l_\omega}{q_\omega} \quad [17]$$

This solution mimics the one taken in conventional life tables, where:

$$L_\omega = \frac{l_\omega}{m_\omega} \quad [18]$$

[9] Randomness and significance of small numbers of events as well as frequent data quality issues at older ages are among the reasons to opt for an earlier-than-extinction closing of the life table, grouping all residual cases in an open-ended age group.

[10] The estimates of a_x can be improved, but this aspect will be the dealt in a separate report that will present further extensions of the alternative life table. In this note, the a_x are taken at their simplest value.

and it works acceptably in most real cases. However, in the extreme case in which all persons at the highest age group ($P_{\omega-1} + P_{\omega}$) have died in the same calendar year, i.e. $q_{\omega} = 1$, the [17] would return a value equal to the size of the last group of survivors l_{ω}. This causes a subtle interpretation twist, because that value does not mean that every survivor would have lived exactly one year, but rather that on average their life would have been about one year, some dying earlier in the same calendar year, other in the following calendar year(s). Likewise, if there are not observed deaths in the open-ended age group, $q_{\omega} = 0$ and the [17] is equal to infinity[11,12].

A different way of arriving to the [17] is to identify an estimate for a_{ω}. Such an approximate estimate of a_{ω} can be obtained by the number of persons exposed to the risk at the beginning of the year divided by the number of observed deaths in the open-ended age class:

$$a_{\omega} = \frac{P_{\omega-1} + P_{\omega}}{D_{\omega}} = \frac{1}{q_{\omega}} \qquad [19]$$

Replacing a_{ω} with this estimate in [16] gives:

$$L_{\omega} = \frac{1}{q_{\omega}} \cdot d_{\omega} = \frac{l_{\omega}}{q_{\omega}} \qquad [20]$$

because of the identity between number of deaths and number of survivors in the open-ended age group.

Once estimated the L_x, the computation of the life table follows the standard method. The total person-years lived above age (reached) *x* is:

$$T_x = \sum_{a=x}^{\omega} L_a \qquad [21]$$

Consequently, the expectation of life at age (reached) *x* is:

$$e_x = \frac{T_x}{l_x} \qquad [22]$$

3.4. An example

The following Table 1 reports the values of the life table functions computed for women in Austria in 2018.

[11] The indefinite form $0/0$ is in principle not possible in L_{ω}, because – in general – if there are neither deaths nor persons at the beginning of the period, then the probability is already indefinite ($q_x = 0/0$) and no further computations are possible without intervention (such as closing earlier the life table or estimating the q_x).

[12] Similar caveats apply to the conventional method as well. If nobody dies during the selected year in the observed population at the open-ended age, then $m_{\omega} = 0$ and the person-years lived by the persons in that age group become infinity. If everybody dies, then the length of life lived by those persons depends upon the way the population exposed to the risk is estimated.

Table 1: annual period-cohort life table, Austria, 2018, women

Age	Pop on 1 January / end of previous year	Live births (female total)	Deaths by age reached during the year	Base population ($\hat{P}_x$)	Probability of dying before end of year (q_x)	Probability of surviving until end of the year (p_x)	Synthetic cohort on 1 January (l_x)	Survivors at the end of the year ("true" l_x)	Synthetic deaths (d_x)	Average time lived by those dying in the year (a_x)	Person-years (L_x)	Cumulative remaining person-years (T_x)	Life expectancy (e_x)
0	41 851	41 739	91	41 739	0.00218	0.99782	100 000	99 782	218	0.20000	99 826	8 452 420	84.5
1	43 340		13	41 851	0.00031	0.99969	99 782	99 751	31	0.50000	99 766	8 352 595	83.7
2	41 794		5	43 340	0.00012	0.99988	99 751	99 739	12	0.50000	99 745	8 252 828	82.7
3	41 553		3	41 794	0.00007	0.99993	99 739	99 732	7	0.50000	99 736	8 153 083	81.7
4	40 760		2	41 553	0.00005	0.99995	99 732	99 728	5	0.50000	99 730	8 053 347	80.7
5	40 853		7	40 760	0.00017	0.99983	99 728	99 710	17	0.50000	99 719	7 953 617	79.8
6	40 470		0	40 853	0.00000	1.00000	99 710	99 710	0	0.50000	99 710	7 853 898	78.8
7	41 159		2	40 470	0.00005	0.99995	99 710	99 705	5	0.50000	99 708	7 754 188	77.8
8	40 206		4	41 159	0.00010	0.99990	99 705	99 696	10	0.50000	99 701	7 654 480	76.8
9	40 493		2	40 206	0.00005	0.99995	99 696	99 691	5	0.50000	99 693	7 554 779	75.8
10	40 348		3	40 493	0.00007	0.99993	99 691	99 683	7	0.50000	99 687	7 455 086	74.8
11	41 318		1	40 348	0.00002	0.99998	99 683	99 681	2	0.50000	99 682	7 355 399	73.8
12	41 317		3	41 318	0.00007	0.99993	99 681	99 674	7	0.50000	99 677	7 255 717	72.8
13	41 681		3	41 317	0.00007	0.99993	99 674	99 666	7	0.50000	99 670	7 156 039	71.8
14	41 027		4	41 681	0.00010	0.99990	99 666	99 657	10	0.50000	99 662	7 056 369	70.8
15	42 239		2	41 027	0.00005	0.99995	99 657	99 652	5	0.50000	99 654	6 956 708	69.8
…	…		…	…	…	…	…	…	…	…	…	…	…
96	3 291		1 168	4 503	0.25938	0.74062	10 698	7 923	2 775	0.50000	9 310	31 486	2.9
97	2 138		916	3 291	0.27833	0.72167	7 923	5 718	2 205	0.50000	6 820	22 175	2.8
98	1 182		658	2 138	0.30776	0.69224	5 718	3 958	1 760	0.50000	4 838	15 355	2.7
99	506		391	1 182	0.33080	0.66920	3 958	2 649	1 309	0.50000	3 303	10 517	2.7
100+	872		506	1 378	0.36720	0.63280	2 649		2 649	2.72332	7 213	7 213	2.7

The three columns in the Table 1 after the first one are the data used as input. From those data, it is prepared the correct base population by increasing the ages by one year, inserting the live births at age 0 and putting together in the new age group 100+ the number of persons previously classified in the age groups 99 and 100+ years old. These two latter cells are highlighted in the Table 1 to draw the attention to their specificity within the same column. The column of the q_x is derived by applying the [5] or, more simply, from:

$$q_x = \frac{D_x}{\hat{P}_x} \qquad [23]$$

where $\hat{P}_x$ is the base population. It can also be noted that $q_\omega < 1$.

The computation of the p_x is straightforward and it is followed by the l_x. For illustrative purposes, the l_x are reported in the Table 1 in both the versions mentioned above: the first column refers to the l_x as synthetic cohort at the beginning of the calendar year (i.e., the vertical line in the Lexis diagram on the left of the calendar year); the next column reports the l_x as survivors aged x at the end of the year. The latter has an easier interpretation, but the former allows the inclusion of the radix, even though its age is shifted one year up in the table (but consistently with [13]).

The life table deaths are easily computed either sequentially from the column of the synthetic cohort on 1 January, or as difference between the two columns of l_x.

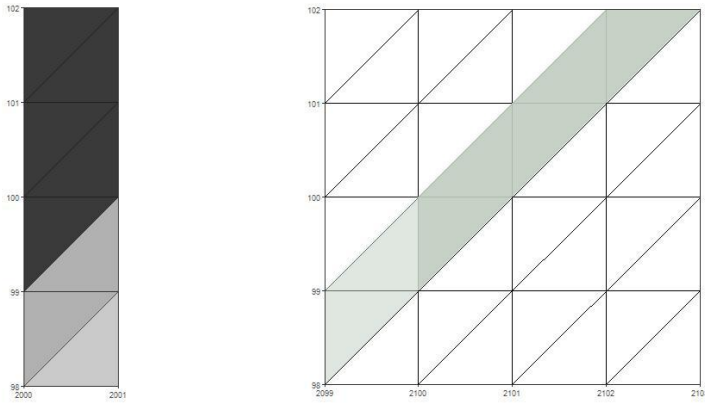
Before computing the person-years lived, it is necessary to introduce the estimates of the average time lived in the calendar year, here taken equal to the assumptions formulated in the traditional approach to easier the comparison but for a_{100+}, whose value is derived by applying the [19]. After that, the computation of the remaining functions (L_x, T_x and e_x) does not present further difficulty.

The values of the life table functions at the bottom of the Table 1 might be better explained by looking the Figure 6, that shows the synthetic cohort from the period (left panel) and cohort perspective (right panel). At age 99 is the last evaluation of the number of survivors, who are actually all persons whom are aged 99 and above. These 2,649 survivors are displayed as:

- the segment at age 99 on the right side of the Lexis diagram in the left panel;
- the vertical half-line starting at age 99 on the left side of the Lexis diagram in the left panel; and as
- the segment at age 99 on 1.1.2100 in the right panel of the Figure 6.

Eventually all those survivors will die, but not necessarily within the same calendar year. They are however the population base for deaths occurring over the following years to those who have reached age 99 by the end of the related calendar year, i.e. by the end of the 2000 year as 'representative' of the year 2099. The $d_{100+} = 2649$ are in the black polygon in the left panel of the Figure 6, and in the darker green polygon in the right panel. The 2 years and almost 9 months from the a_{100+} is the average time lived by those who have reached age 99 at the end of the year 2099 (in the cohort perspective), or by those aged 99 years old and over in the period perspective.

Figure 6: Lexis diagram at open-ended age group for period-cohort parallelograms



4. APPLICATION TO EU AND EFTA COUNTRIES

The method described above has been applied to the available data from 30 European countries (the EU27 Member States and three EFTA countries) for the years 2013-2018, i.e. using data transmitted to Eurostat since the entry into force of the EU Regulation on demographic statistics[13] until the latest available data.

The data are taken from the Eurostat database, namely from the tables:

- "Population on 1 January by age and sex" (*demo_pjan*);
- "Live births by mother's age and newborn's sex" (*demo_fasec*); and
- "Deaths by year of birth (age reached) and sex" (*demo_mager*).

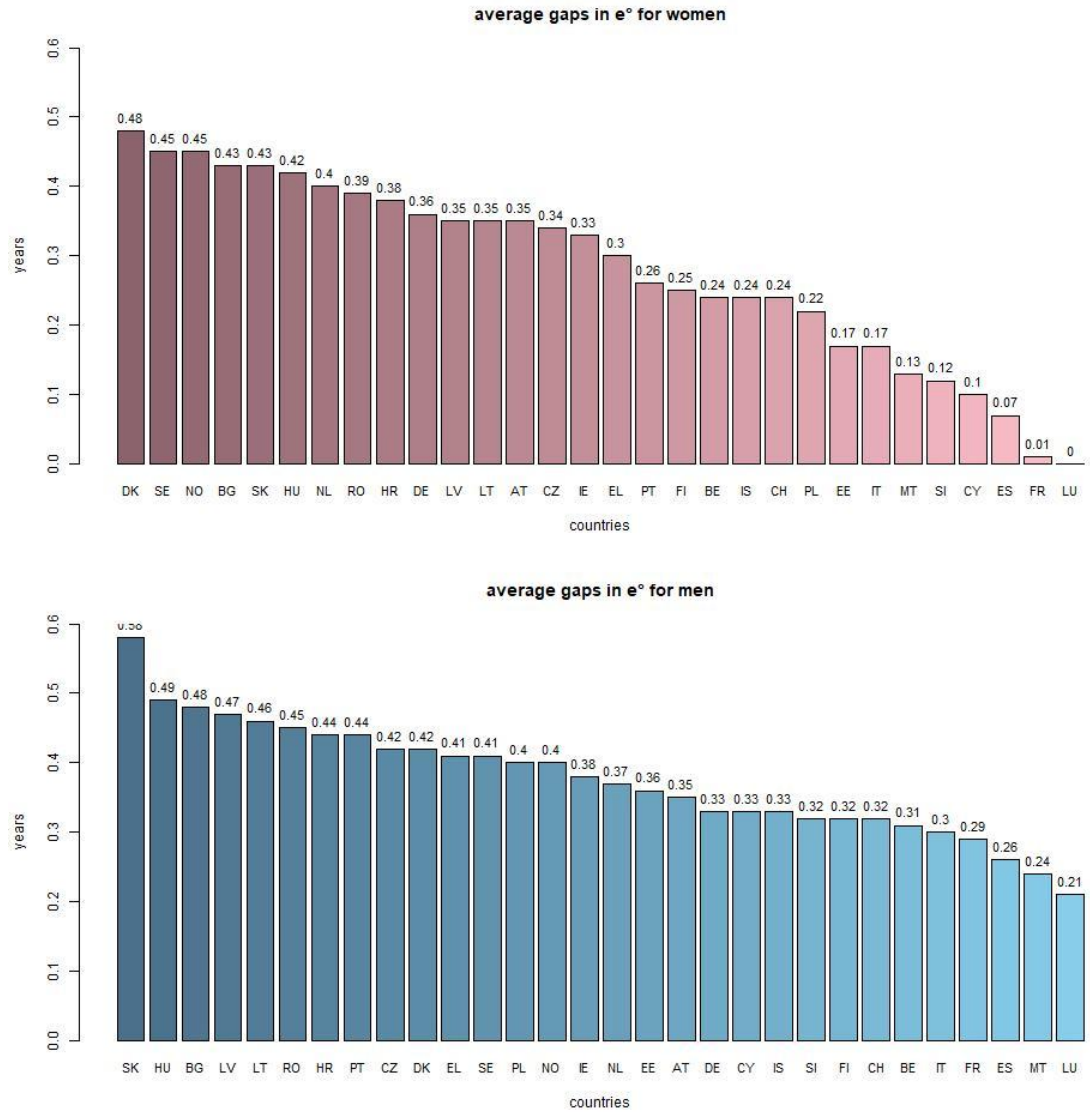
The life table is closed at age 100 and it is computed for men, women and both sexes. In order to have a better comparison with the standard method applied in Eurostat, the traditional approach has been applied to a life table closing as well at age 100, thus extending the currently applied age breakdown. Life tables for Croatia and Hungary are not computed for 2013 because their population on 1 January 2013 is not broken down up to age 100+.

The results for the life expectancy at birth obtained with the alternative method are shown in the Annex, as well as reported in the Table 2 and the Table 3. The Figure 7 shows instead the differences from the values obtained with the standard method, which are on average 0.28 years for the women, 0.38 for the men and 0.33 for both sexes. The graphs in Annex show as well the values of the life expectancy at birth obtained by extending the closing of the life table up to age 100+ in the standard method. It is there shown that the comparison with the extended life table would generally increase the differences between the estimated life expectancies. Differences from the life expectancy at birth from the standard method are computed against this latter value rounded to one decimal digit, as available from the Eurostat database, which may introduce a minimal lack of precision in these estimated differences.

[13] The only exception is the population on 1 January 2013, which was not yet subject to the EU Regulation.

All computations have been done in R[14] version 3.6.3 (2020-02-29, "Holding the Windsock"), using as well the "eurostat"[15], "LexiPlotR"[16] and "RColorBrewer"[17] packages, within the R Studio[18] version 1.2.5033 (2019-12-04, "Orange Blossom").

Figure 7: average gaps in life expectancy at birth between the alternative and the standard method by country and sex



[14] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[15] Leo Lahti, Janne Huovari, Markus Kainu, Przemyslaw Biecek (2017). "Retrieval and analysis of Eurostat open data with the eurostat package". R Journal 9(1):385-392, 2017. R package version 3.6.1.

[16] Philipp Ottolinger (2020). LexisPlotR: Plot Lexis Diagrams for Demographic Purposes. R package version 0.4.0.

[17] Erich Neuwirth (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2.

[18] RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

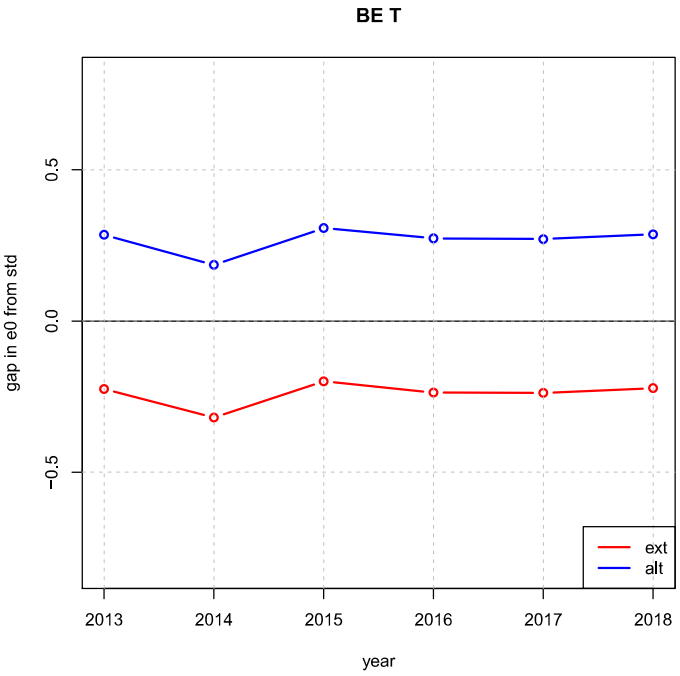
Table 2: life expectancy at birth for women as from the alternative method

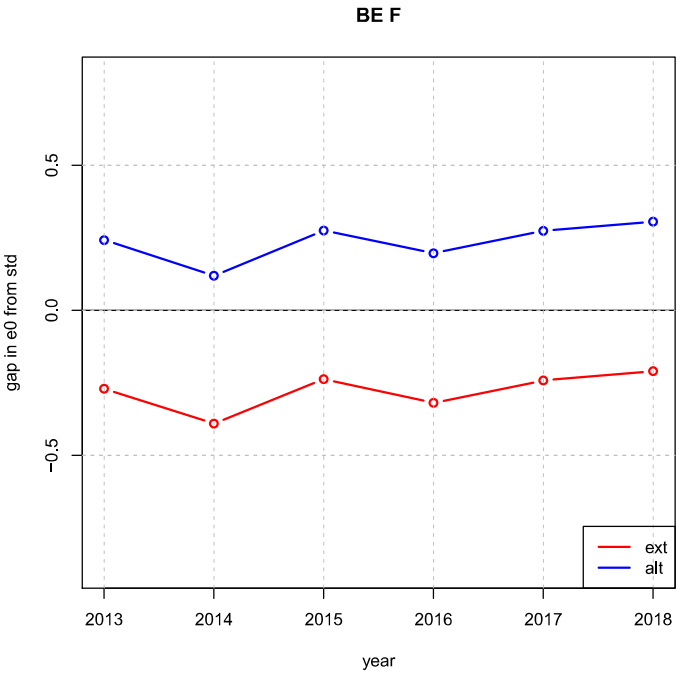
	2013	2014	2015	2016	2017	2018
BE	83.4	84.0	83.7	84.2	84.2	84.2
BG	79.0	78.5	78.6	79.0	78.8	79.0
CZ	81.7	82.2	82.0	82.4	82.4	82.4
DK	82.8	83.2	83.2	83.3	83.6	83.5
DE	83.4	83.9	83.5	83.8	83.8	83.7
EE	81.8	82.0	82.3	82.4	82.8	82.9
IE	83.5	83.8	83.8	83.9	84.3	84.4
EL	84.2	84.4	84.1	84.3	84.2	84.7
ES	86.1	86.2	85.9	86.3	86.2	86.3
FR	85.6	85.9	85.6	85.8	85.8	86.0
HR	NA	81.4	80.9	81.7	81.4	81.8
IT	85.3	85.6	85.2	85.7	85.5	85.8
CY	85.1	84.4	83.9	85.0	84.4	84.6
LV	79.3	79.7	79.8	79.9	80.0	80.1
LT	79.9	80.4	80.1	80.5	80.9	81.1
LU	84.0	85.0	84.8	85.2	84.5	84.8
HU	NA	79.8	79.4	80.1	79.8	80.0
MT	84.2	84.4	84.3	84.4	84.7	84.7
NL	83.5	83.8	83.6	83.6	83.8	83.8
AT	84.1	84.2	84.1	84.5	84.4	84.5
PL	81.4	81.9	81.8	82.2	82.0	82.0
PT	84.3	84.7	84.6	84.6	84.8	84.7
RO	79.1	79.1	79.0	79.4	79.5	79.7
SI	83.7	84.2	84.0	84.4	84.2	84.5
SK	80.5	80.9	80.6	81.1	81.2	81.2
FI	84.3	84.4	84.6	84.6	84.7	84.8
SE	84.2	84.6	84.5	84.6	84.6	84.8
IS	84.0	84.6	84.1	84.4	84.6	84.7
NO	84.2	84.6	84.7	84.7	84.8	85.0
CH	85.3	85.6	85.4	85.7	85.9	85.9

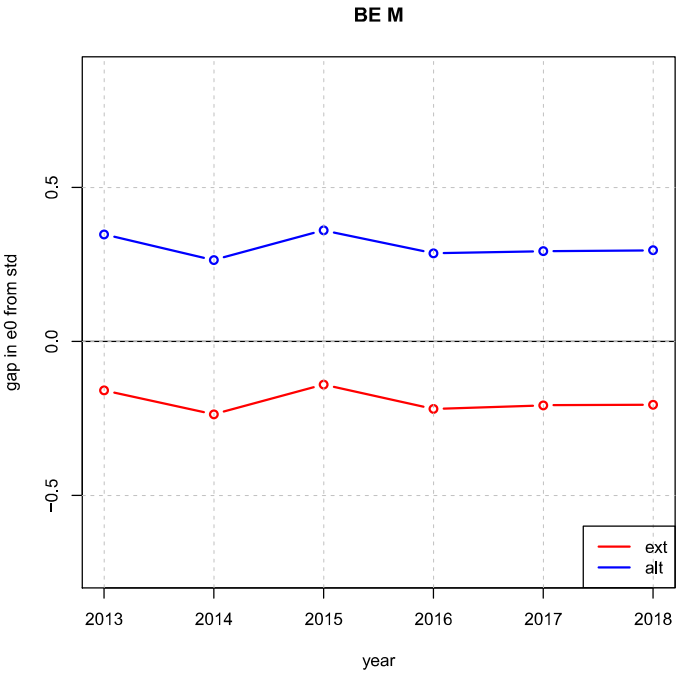
Table 3: life expectancy at birth for men as from the alternative method

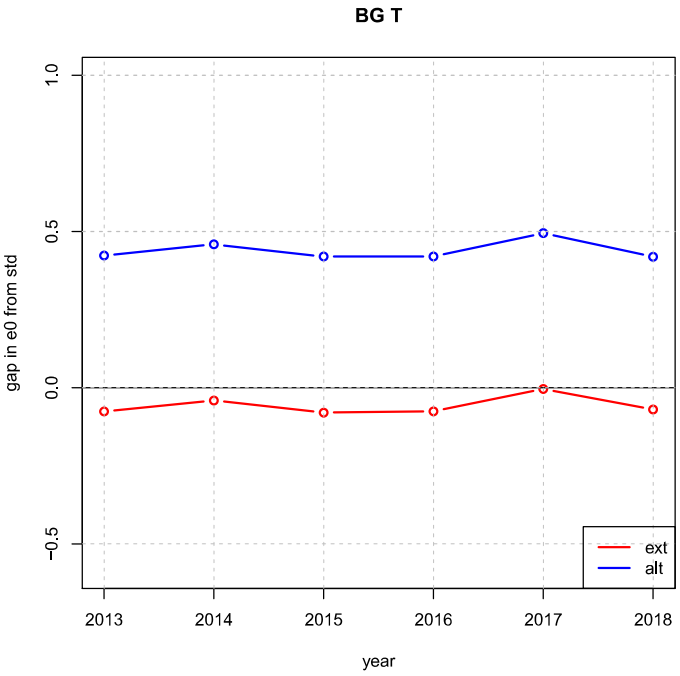
	2013	2014	2015	2016	2017	2018
BE	78.4	79.1	79.1	79.3	79.5	79.7
BG	71.8	71.6	71.7	71.8	71.9	72.0
CZ	75.6	76.2	76.1	76.5	76.5	76.6
DK	78.8	79.1	79.3	79.4	79.6	79.5
DE	78.5	78.9	78.6	78.9	79.1	79.0
EE	73.2	72.8	73.5	73.7	74.1	74.3
IE	79.4	79.7	80.0	80.1	80.7	80.9
EL	79.0	79.1	78.9	79.3	79.3	79.9
ES	80.4	80.6	80.4	80.8	80.9	80.9
FR	79.3	79.7	79.5	79.8	79.9	80.0
HR	NA	75.1	74.8	75.4	75.4	75.3
IT	80.5	81.0	80.7	81.3	81.2	81.5
CY	80.4	80.7	80.3	80.8	80.6	81.1
LV	69.8	69.6	70.2	70.3	70.3	70.5
LT	69.0	69.6	69.7	70.0	71.2	71.4
LU	80.0	79.6	80.2	80.2	80.1	80.4
HU	NA	72.8	72.8	73.1	73.0	73.2
MT	79.9	80.0	80.2	80.8	80.5	80.6
NL	79.9	80.4	80.2	80.4	80.6	80.7
AT	79.0	79.4	79.1	79.6	79.8	79.8
PL	73.4	74.1	73.9	74.3	74.3	74.2
PT	78.1	78.5	78.6	78.5	78.8	78.7
RO	72.0	71.8	71.8	72.1	72.2	72.2
SI	77.5	78.5	78.1	78.5	78.6	78.8
SK	73.4	73.8	73.6	74.4	74.4	74.6
FI	78.3	78.7	79.0	78.9	79.2	79.4
SE	80.6	80.8	80.8	81.0	81.2	81.3
IS	80.9	81.6	81.6	80.8	81.2	81.6
NO	80.2	80.5	80.9	81.1	81.4	81.5
CH	81.0	81.4	81.2	82.0	81.9	82.2

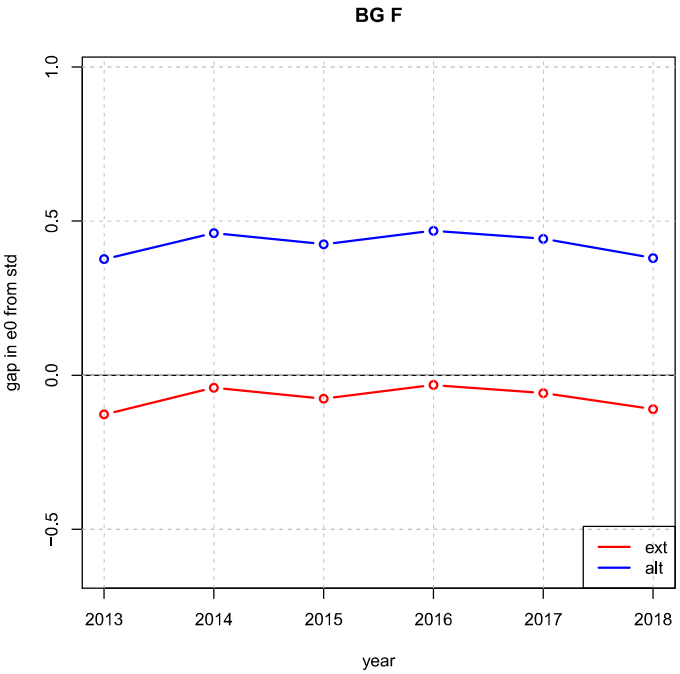
ANNEX

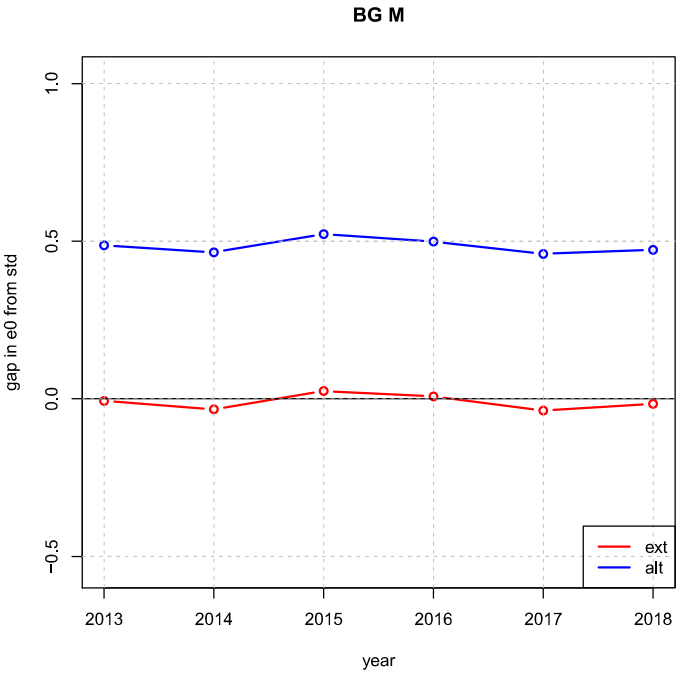

BE T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

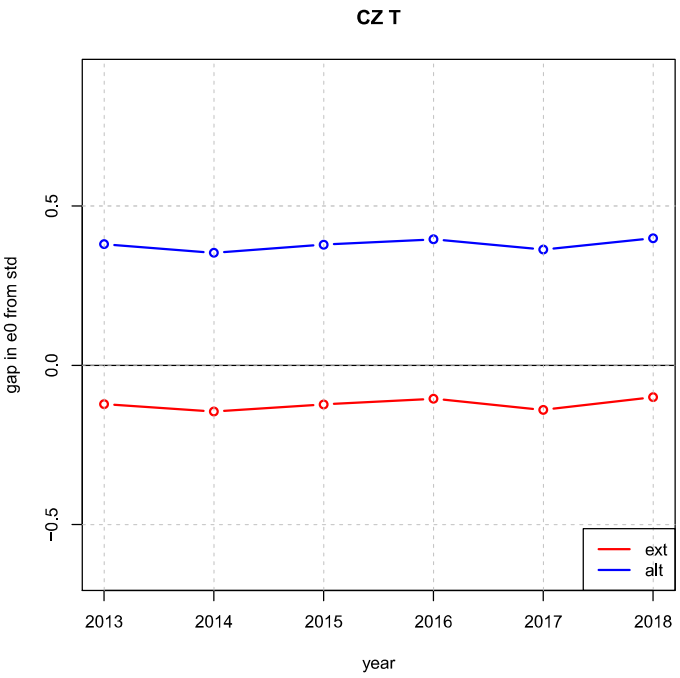

BE F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

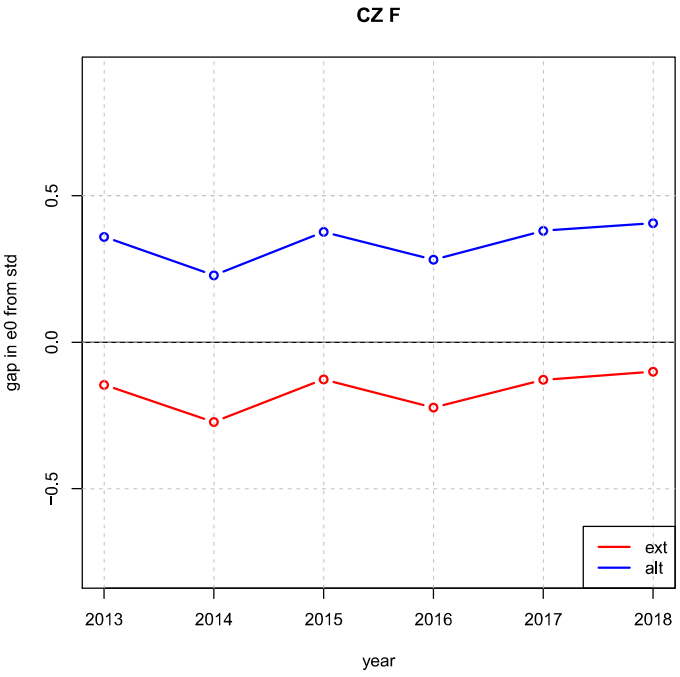

BE M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

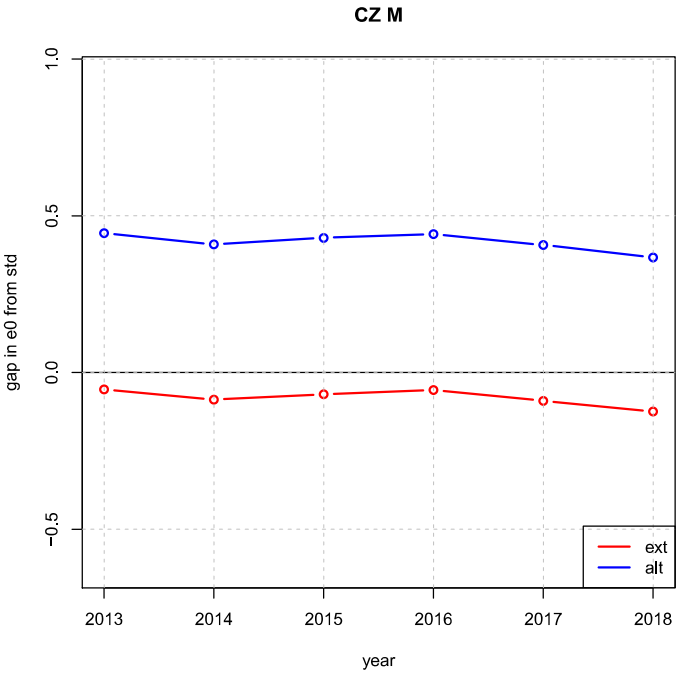

BG T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

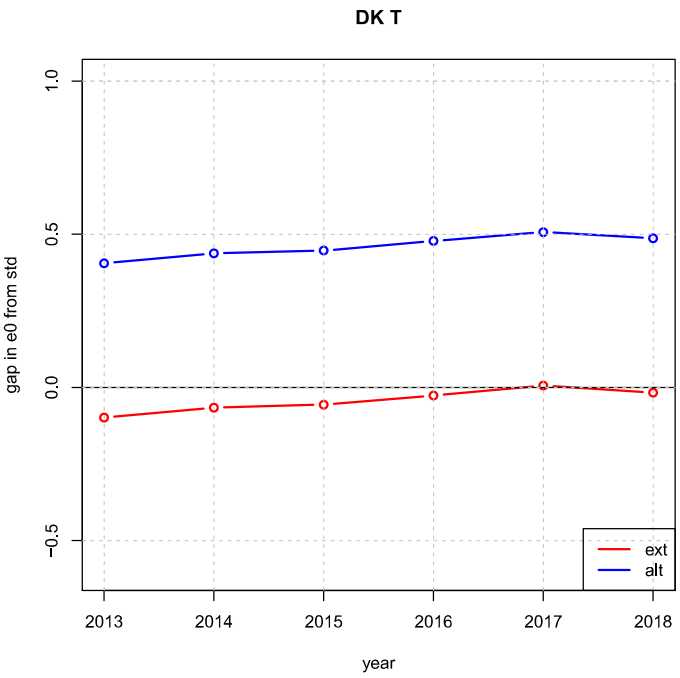

BG F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

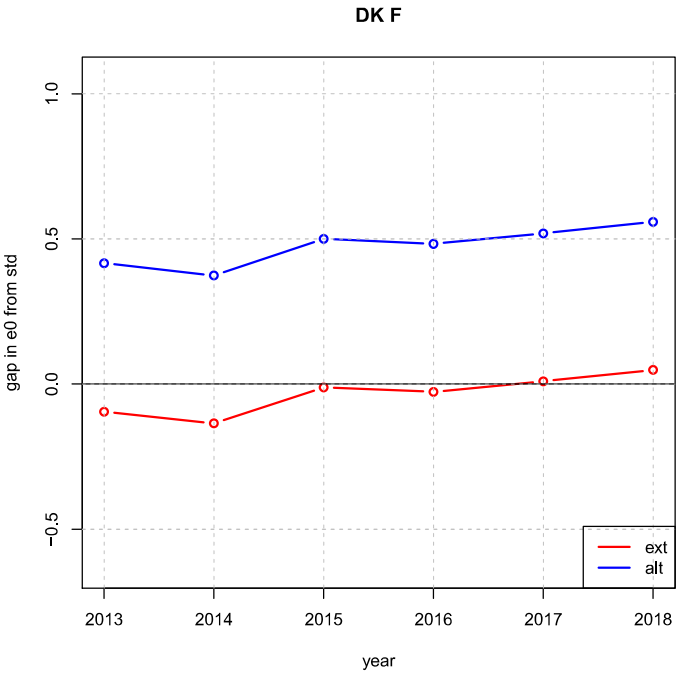

BG M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

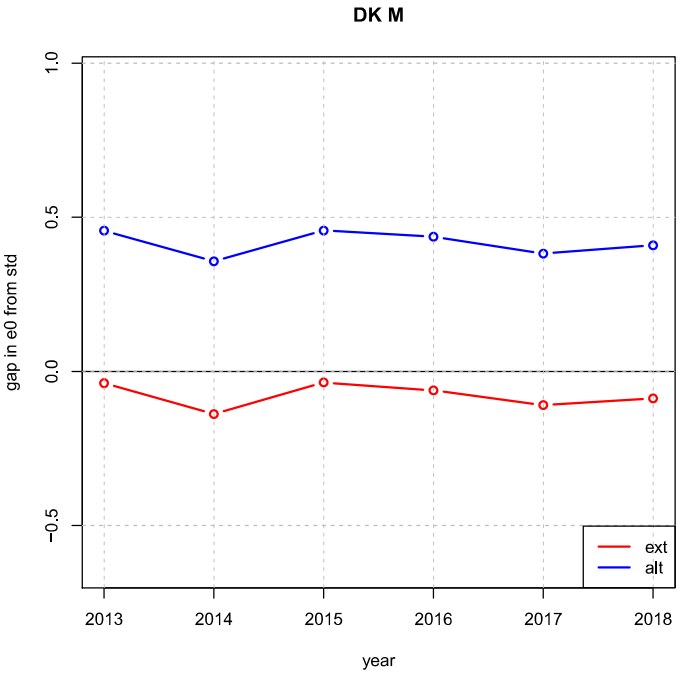

CZ T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

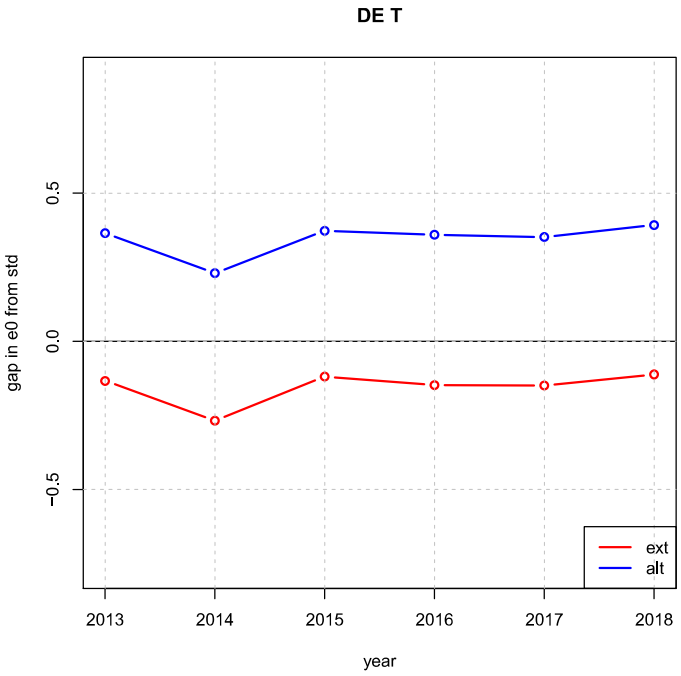

CZ F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

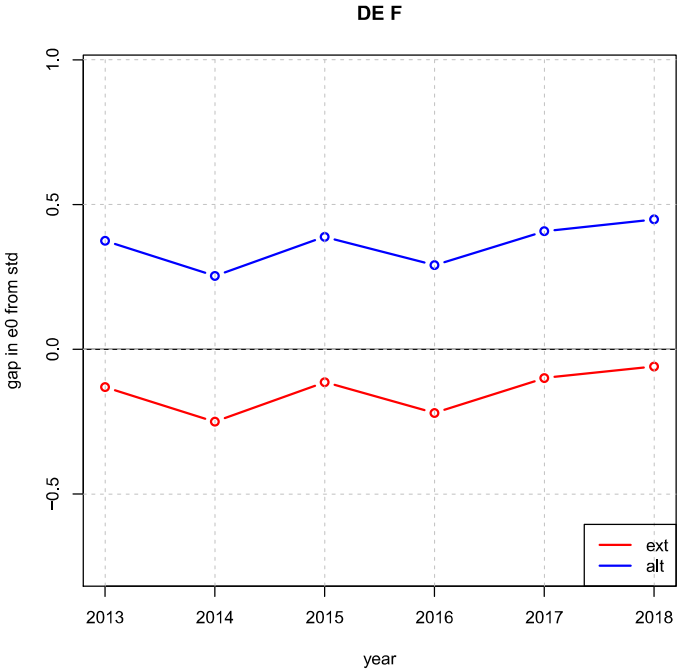

CZ M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

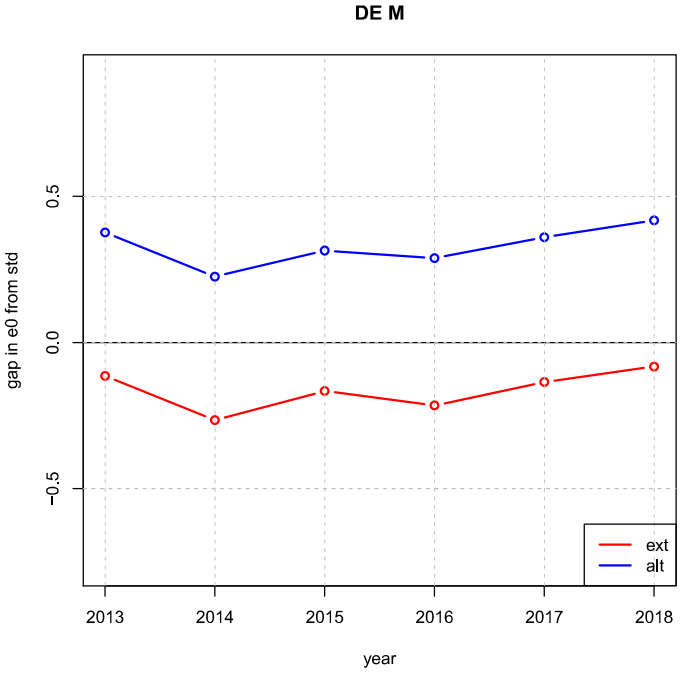

DK T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

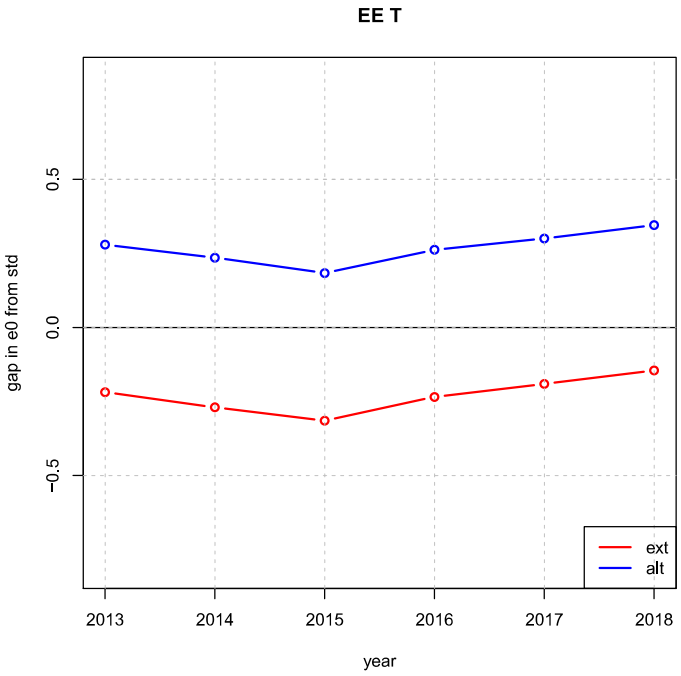

DK F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

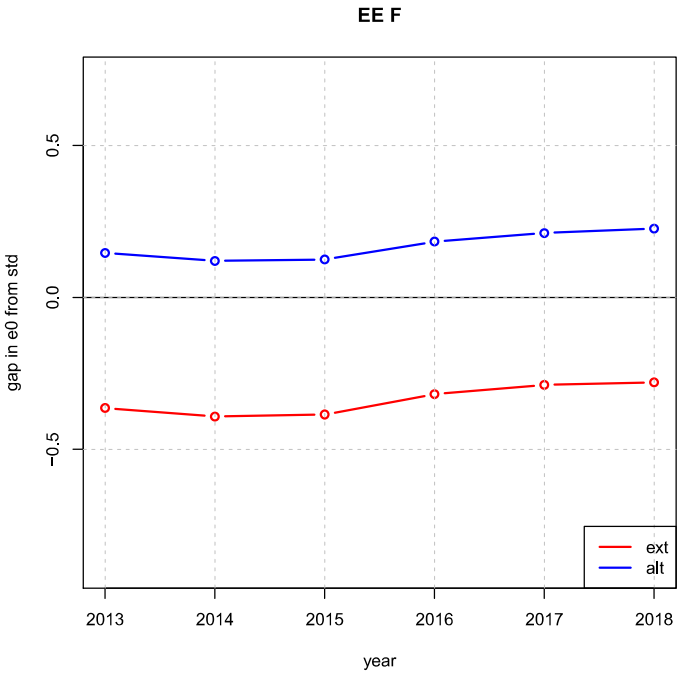

DK M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

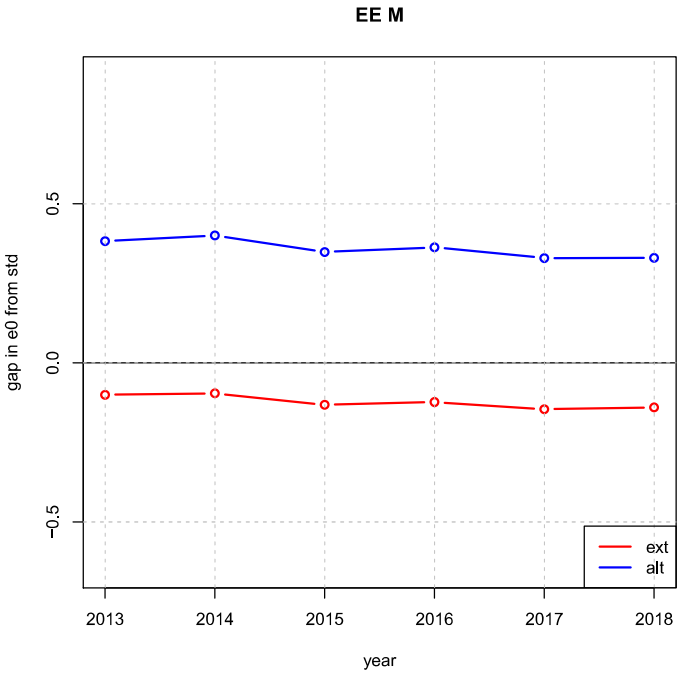

DE T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

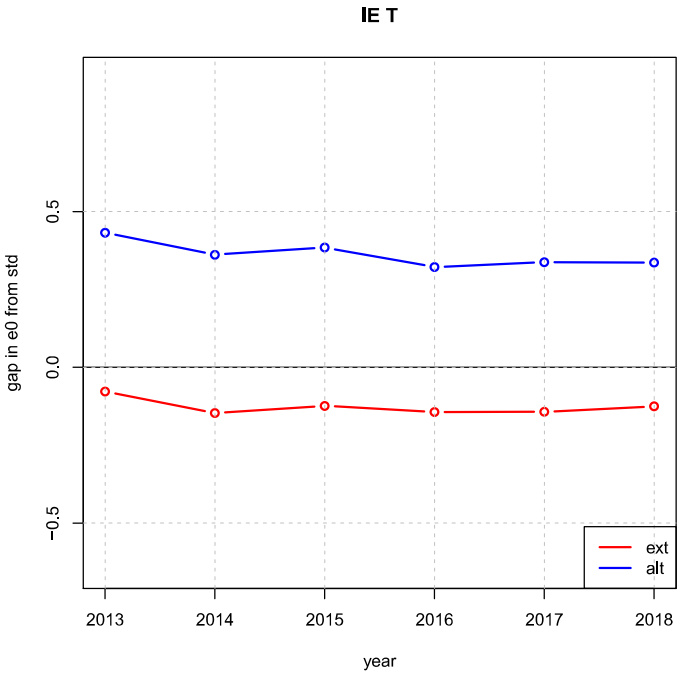

DE F
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

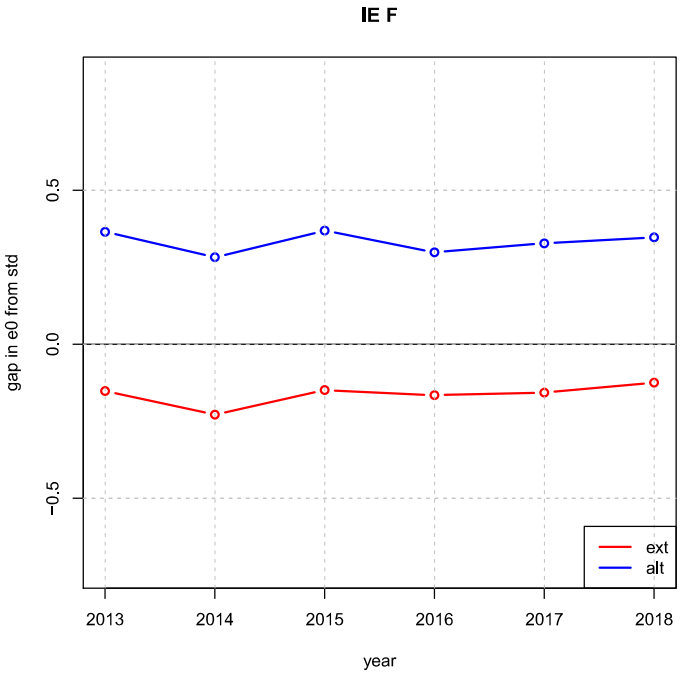

DE M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

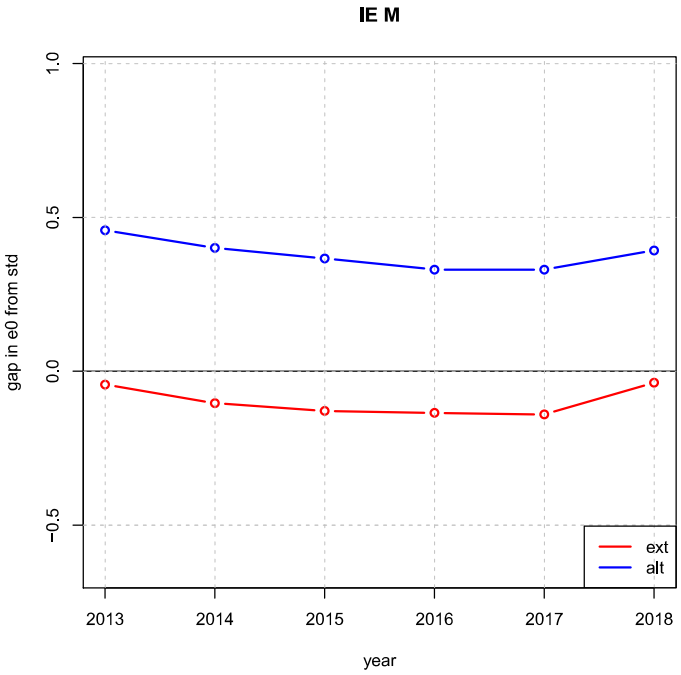

EE T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

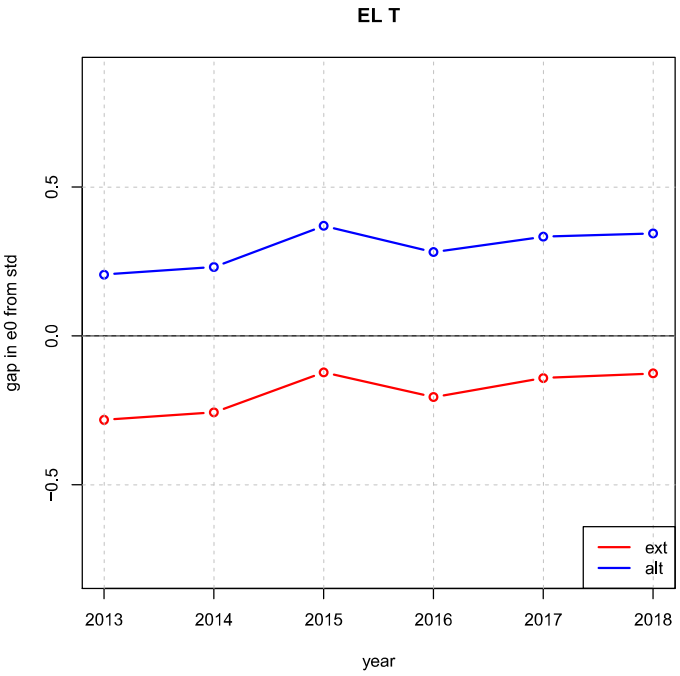

EE F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

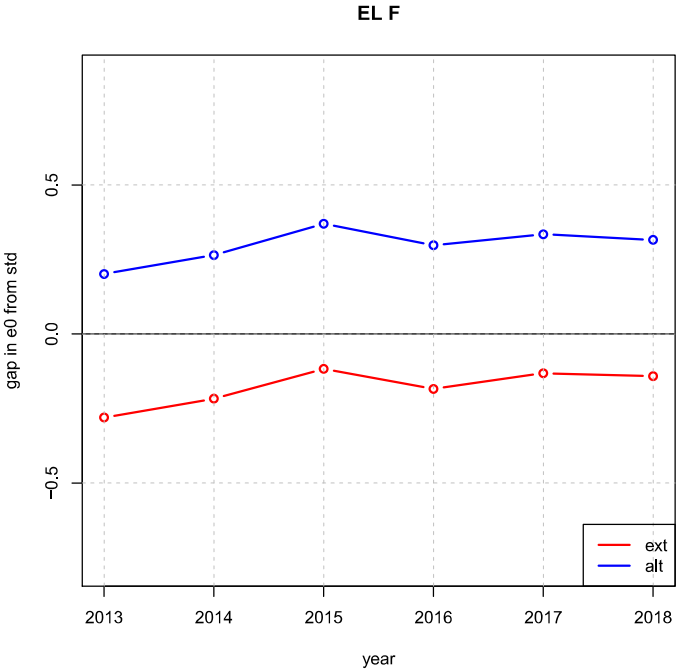

EE M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

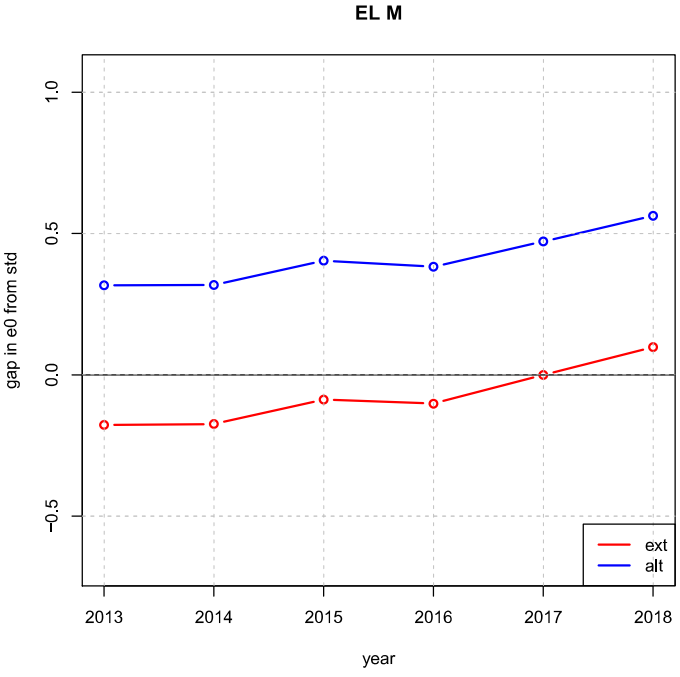

IE T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

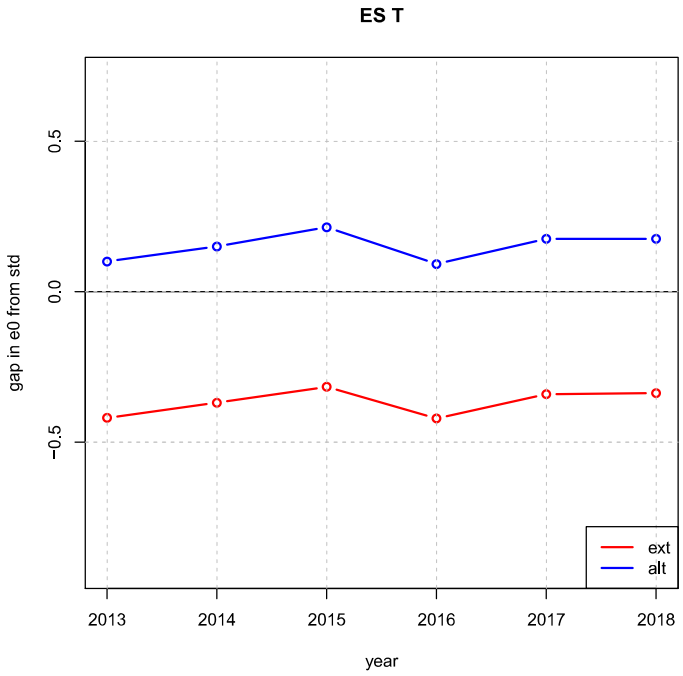

IE F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

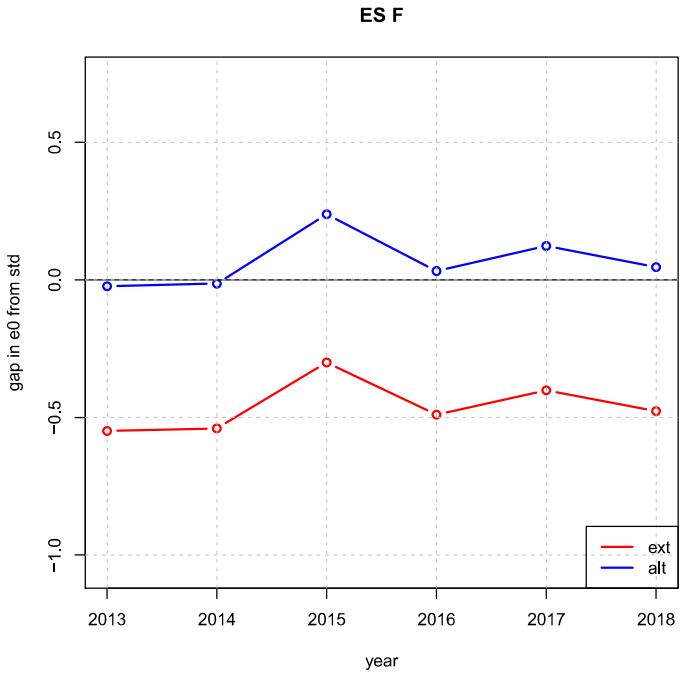

IE M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

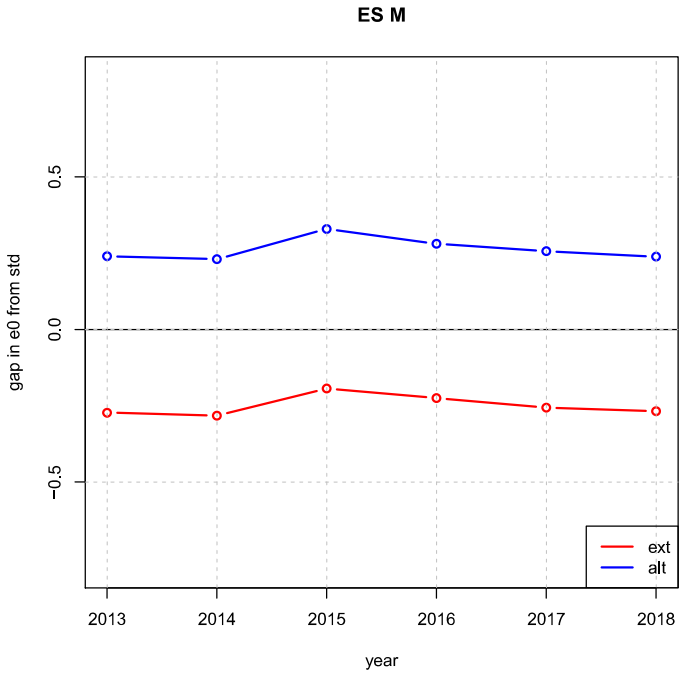

EL T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

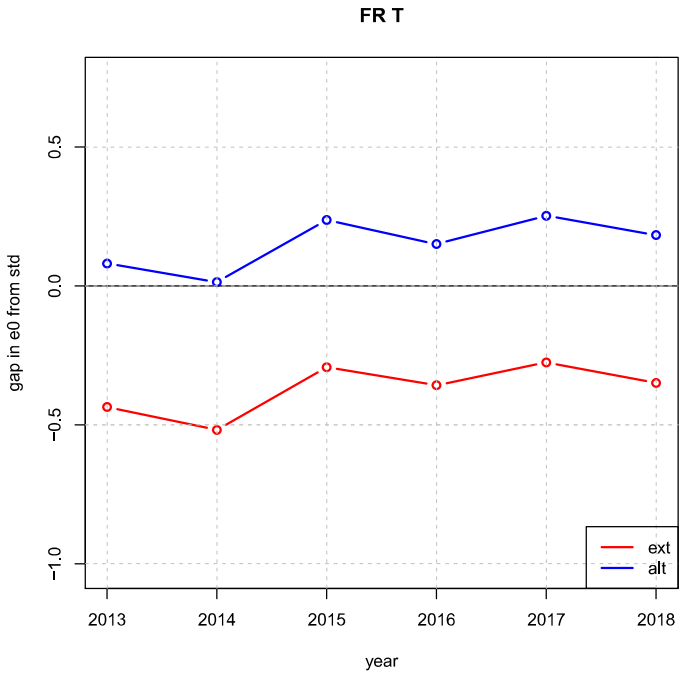

EL F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

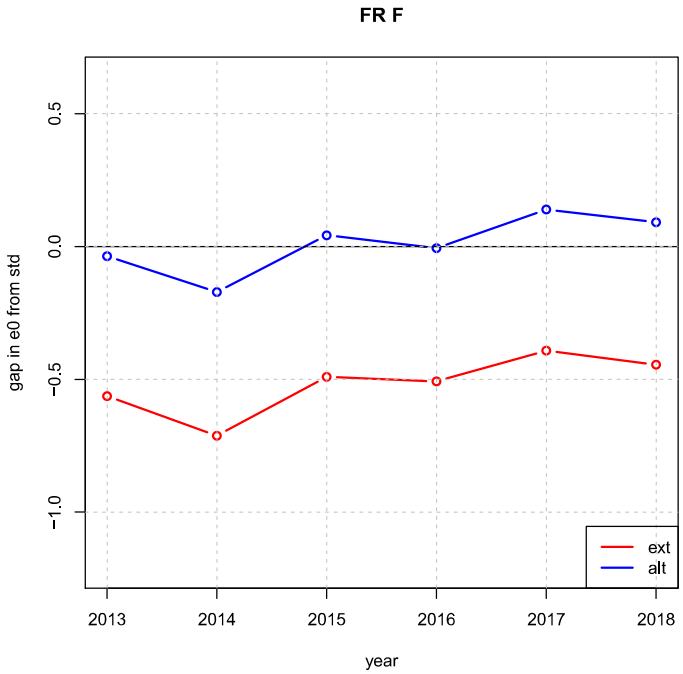

EL M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

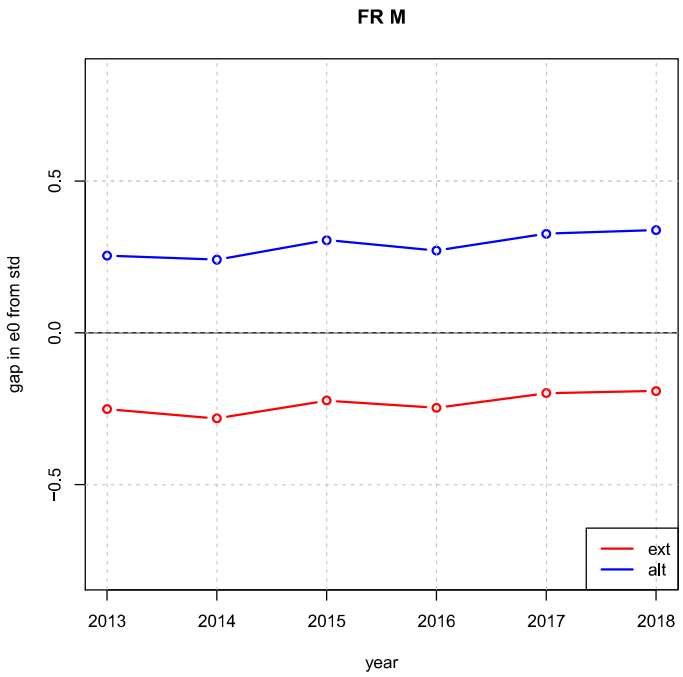

ES T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

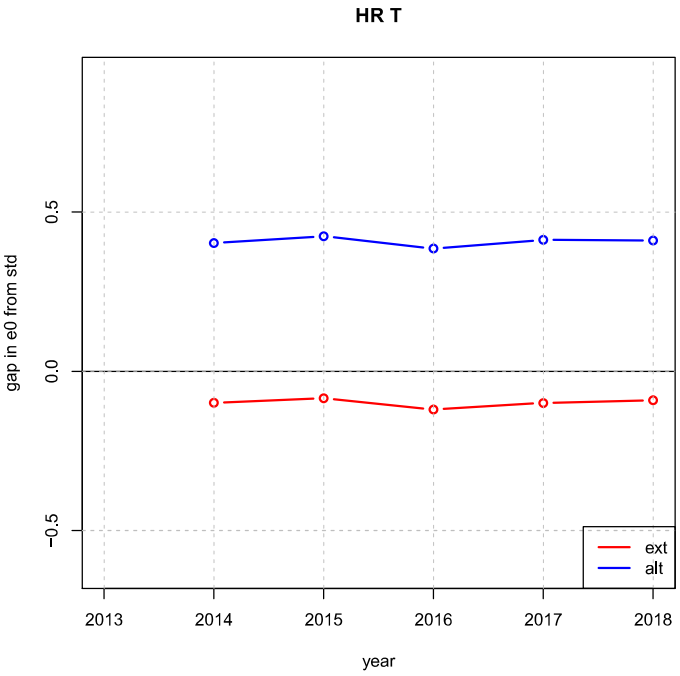

ES F
gap in e0 from std
0.5
0.0
−0.5
−1.0
ext
alt
2013
2014
2015
2016
2017
2018
year

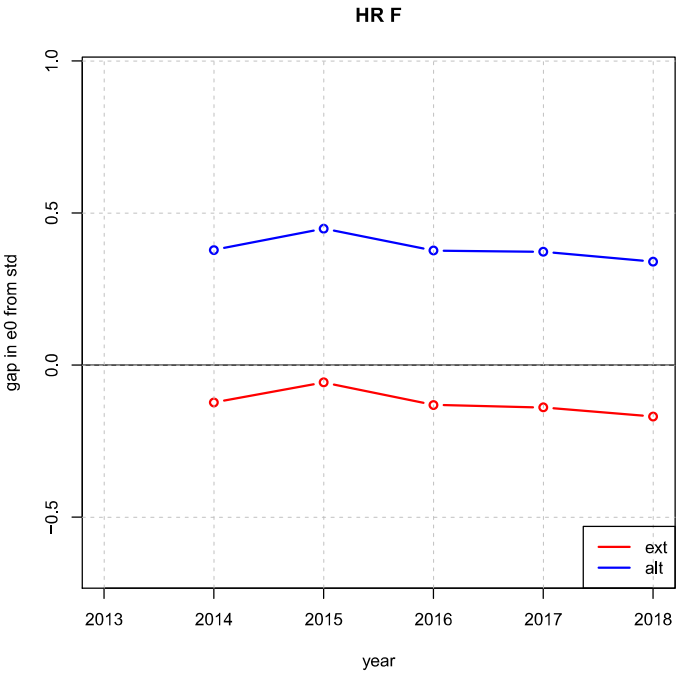

ES M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

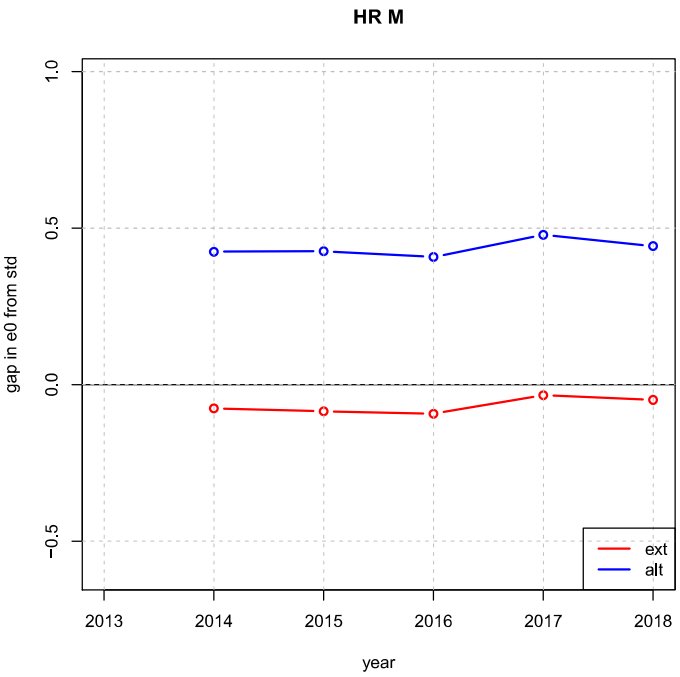

FR T
gap in e0 from std
0.5
0.0
−0.5
−1.0
2013
2014
2015
2016
2017
2018
year
ext
alt

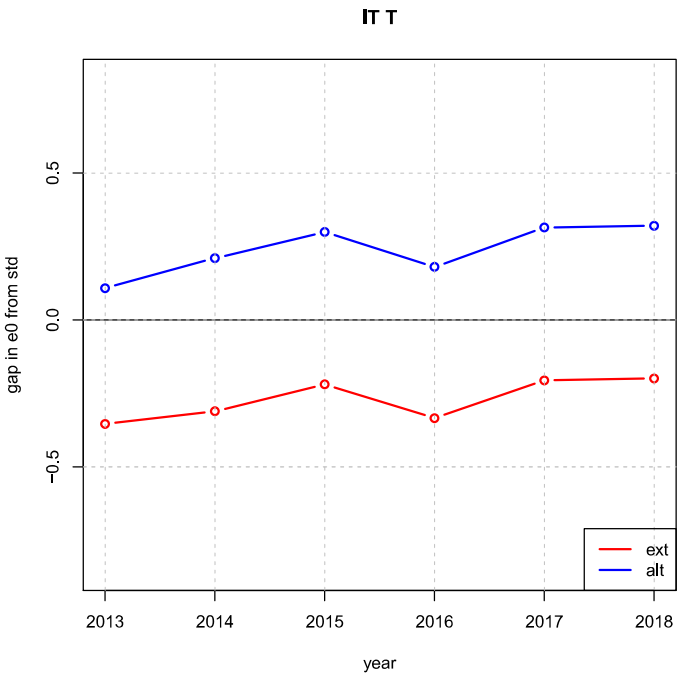

FR F
gap in e0 from std
0.5
0.0
−0.5
−1.0
2013
2014
2015
2016
2017
2018
year
ext
alt

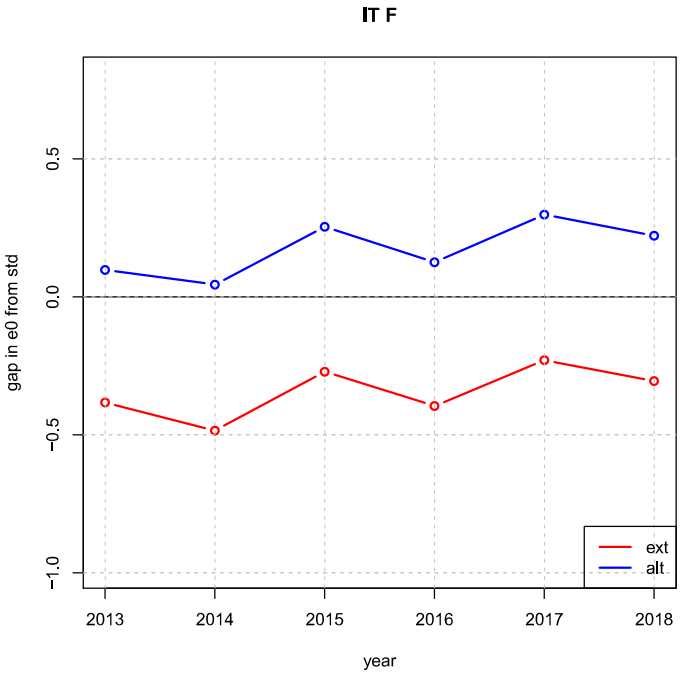

FR M
gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

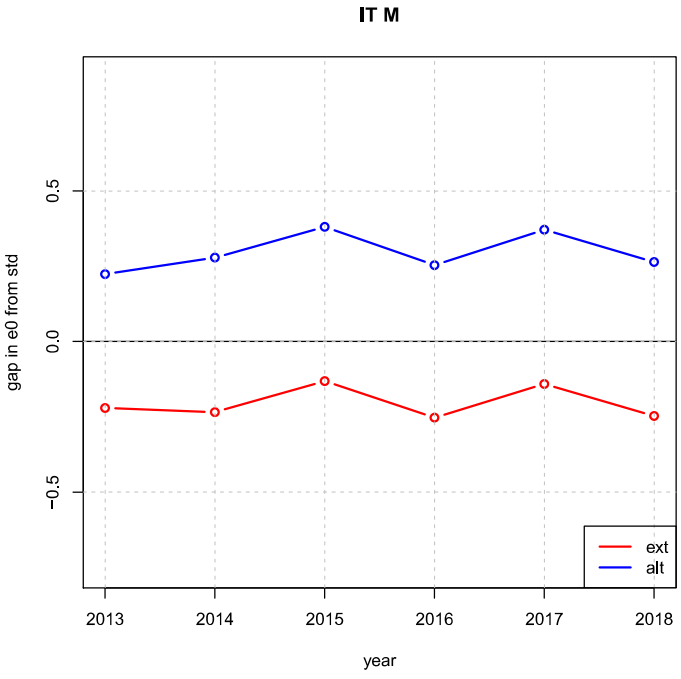

HR T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

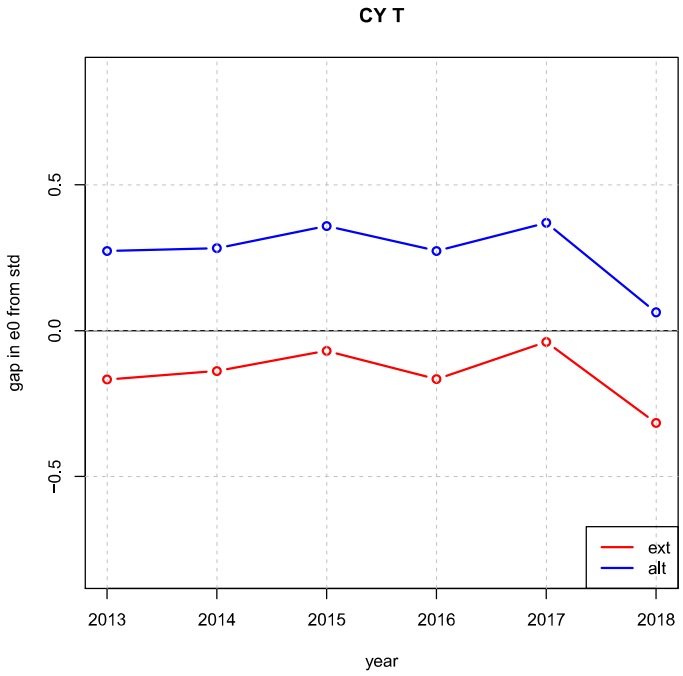

HR F
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

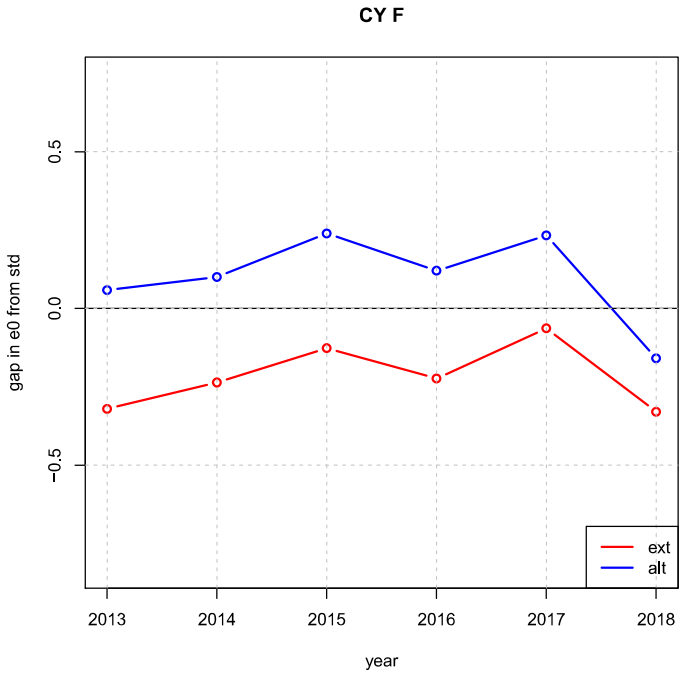

HR M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

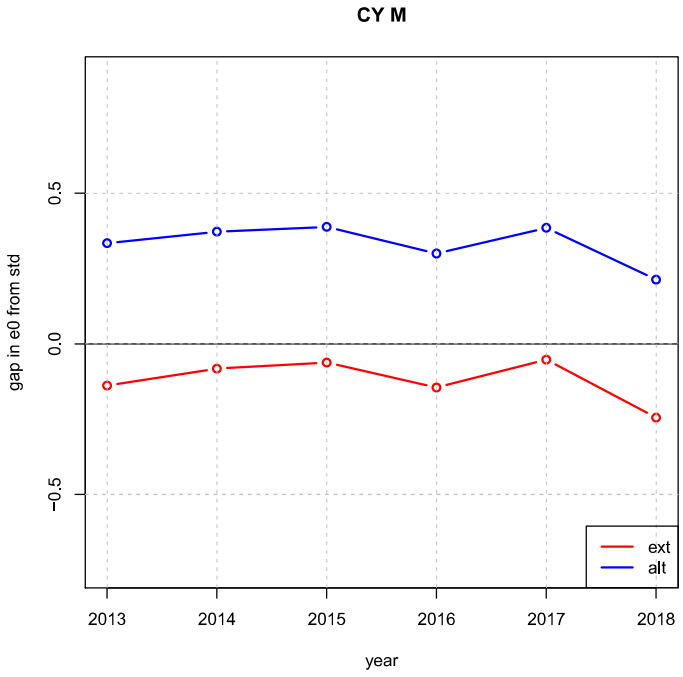

IT T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

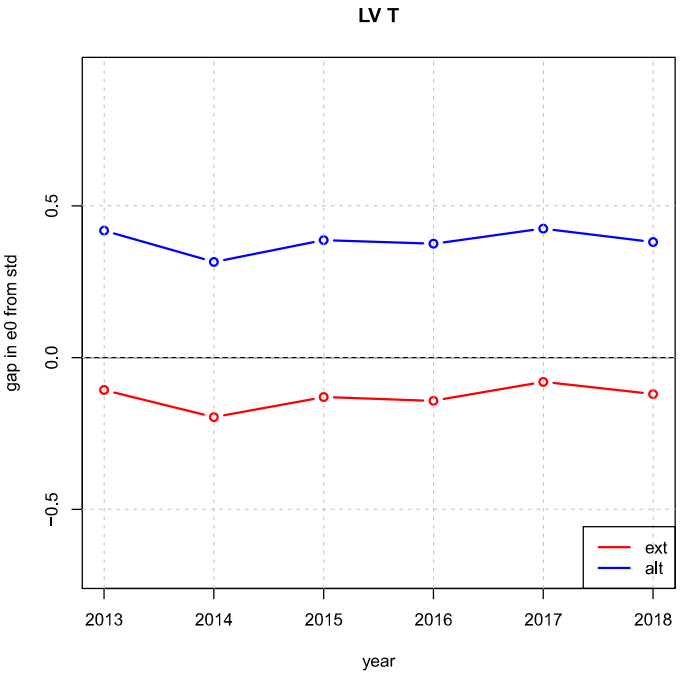

IT F
gap in e0 from std
0.5
0.0
−0.5
−1.0
ext
alt
2013
2014
2015
2016
2017
2018
year

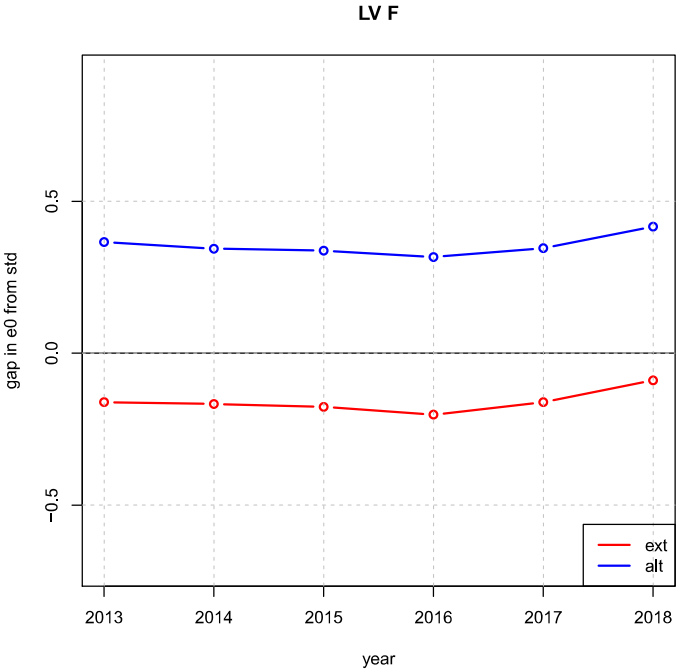

IT M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

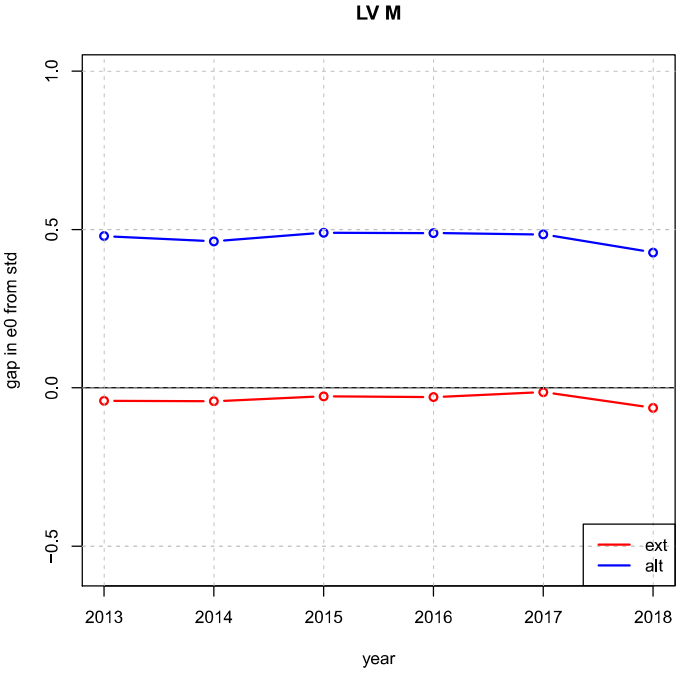

CY T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

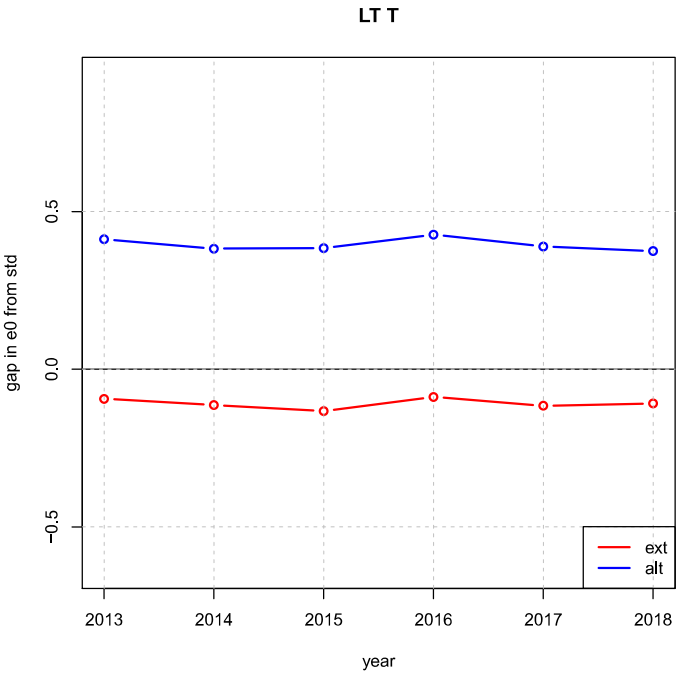

CY F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

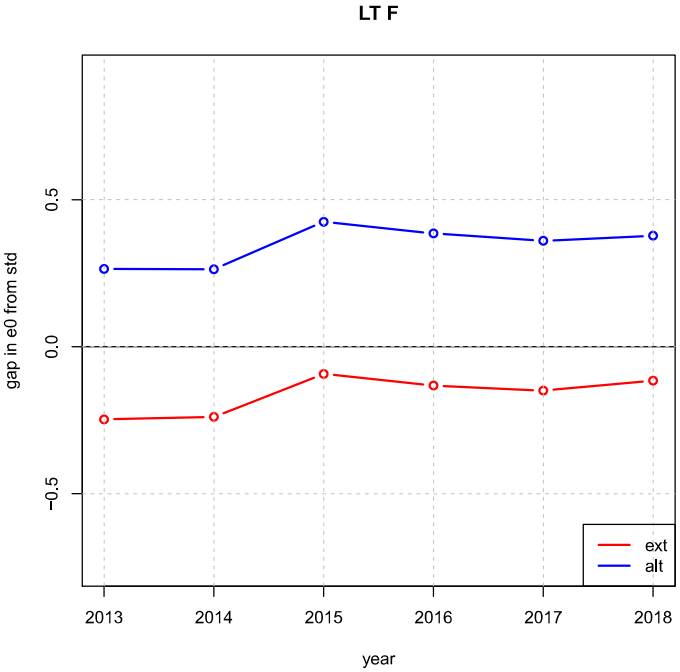

CY M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

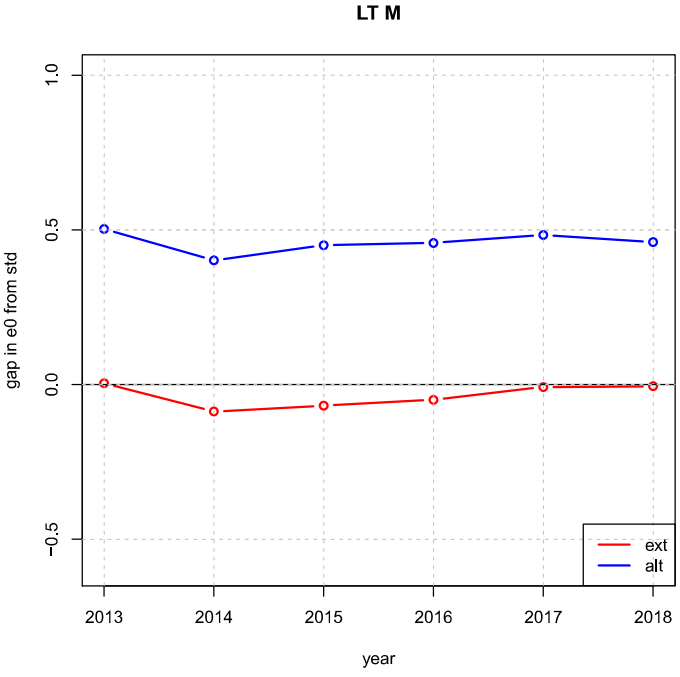

LV T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

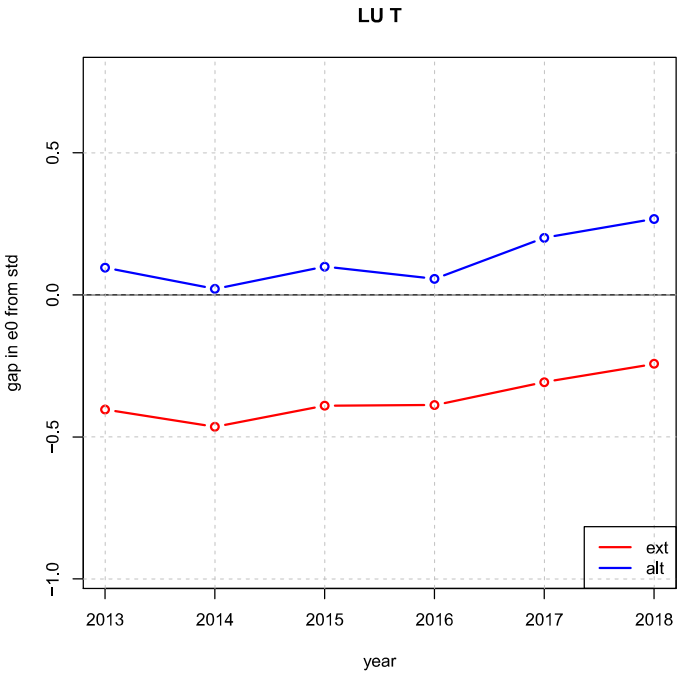

LV F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

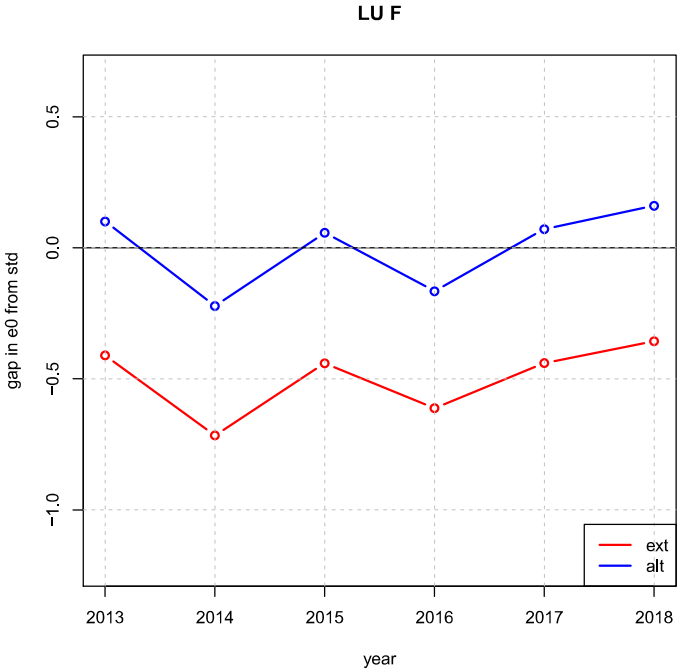

LV M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

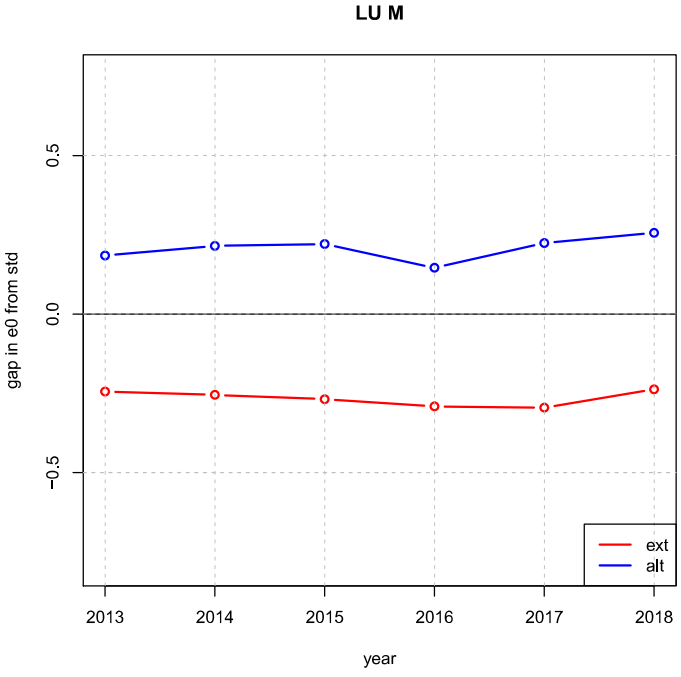

LT T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

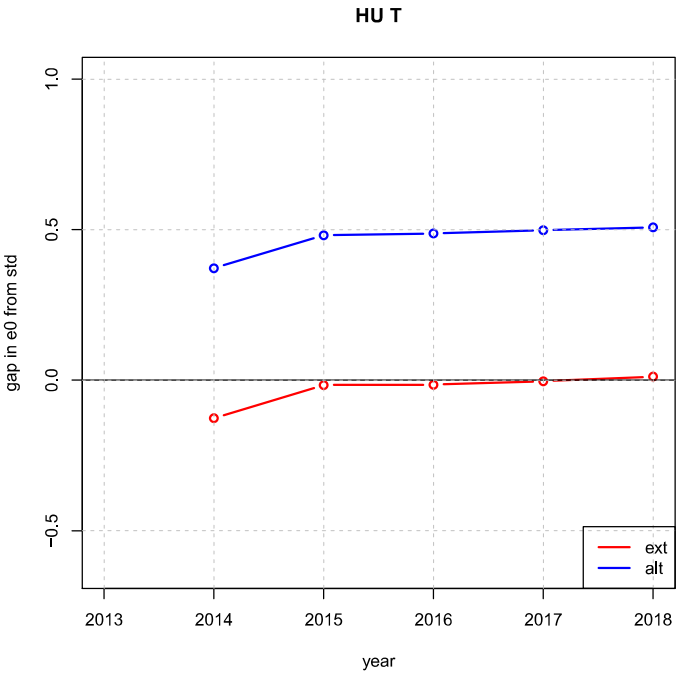

LT F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

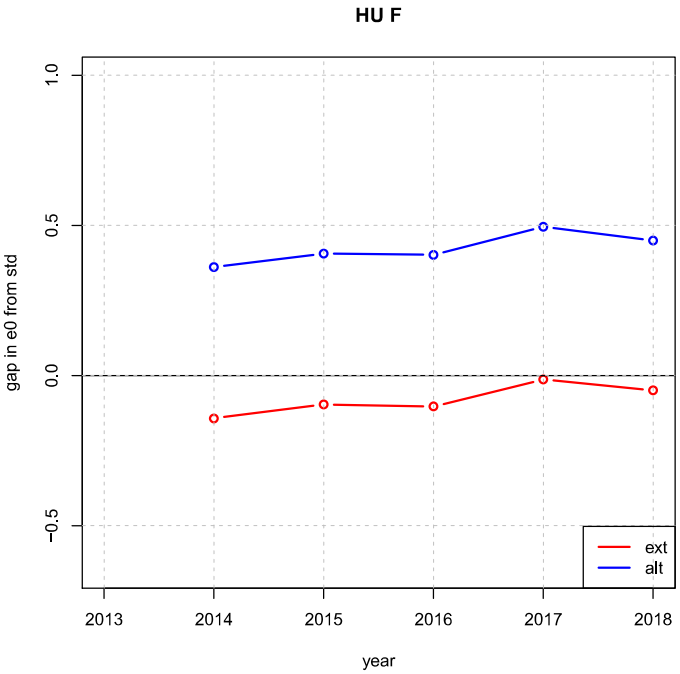

LT M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

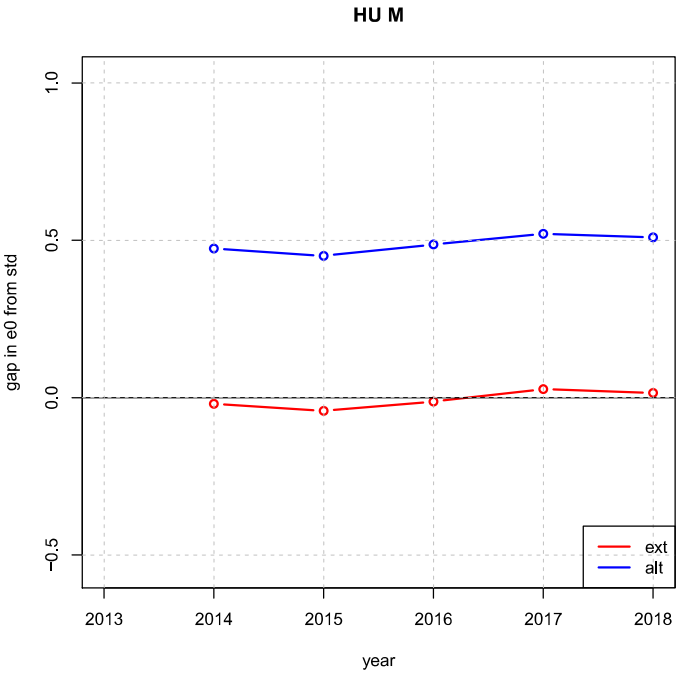

LU T
gap in e0 from std
0.5
0.0
−0.5
−1.0
2013
2014
2015
2016
2017
2018
year
ext
alt

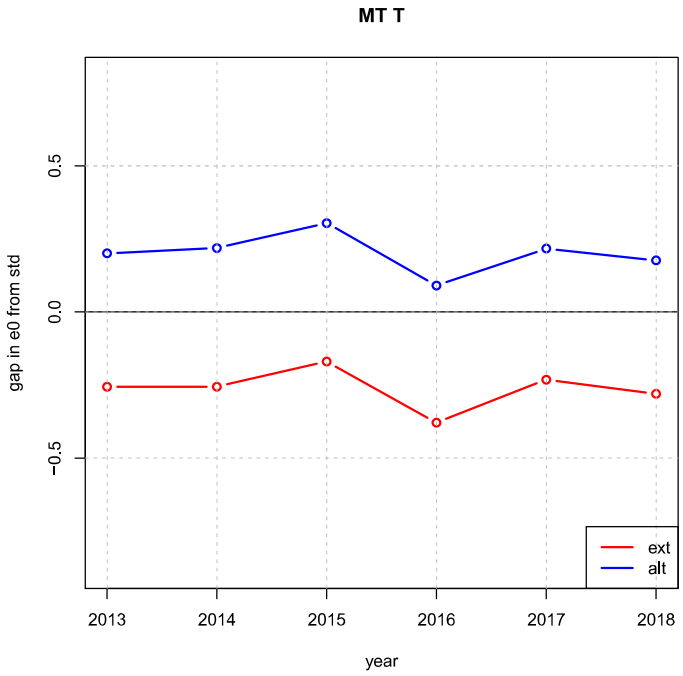

LU F
gap in e0 from std
0.5
0.0
−0.5
−1.0
2013
2014
2015
2016
2017
2018
year
ext
alt

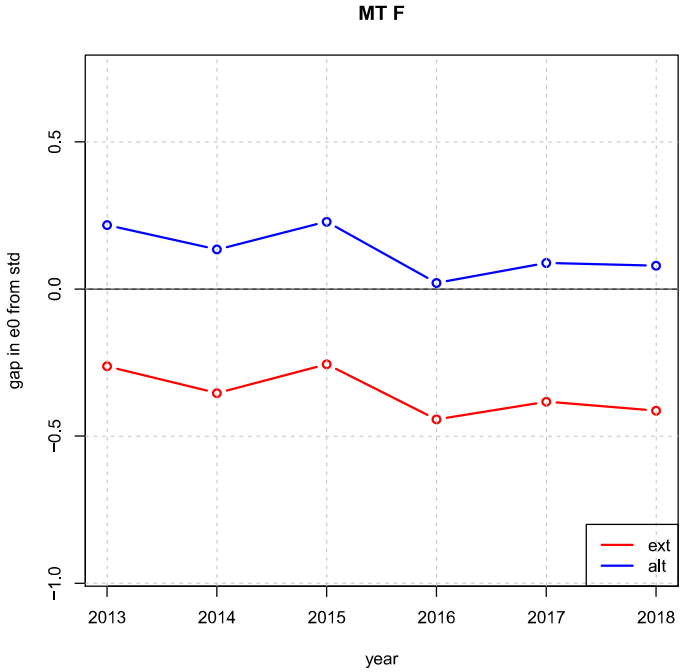

LU M
gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

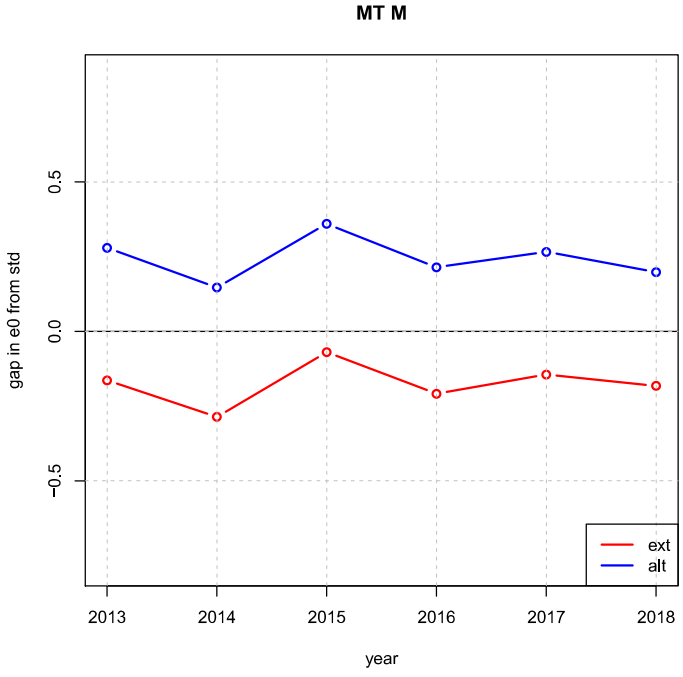

HU T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

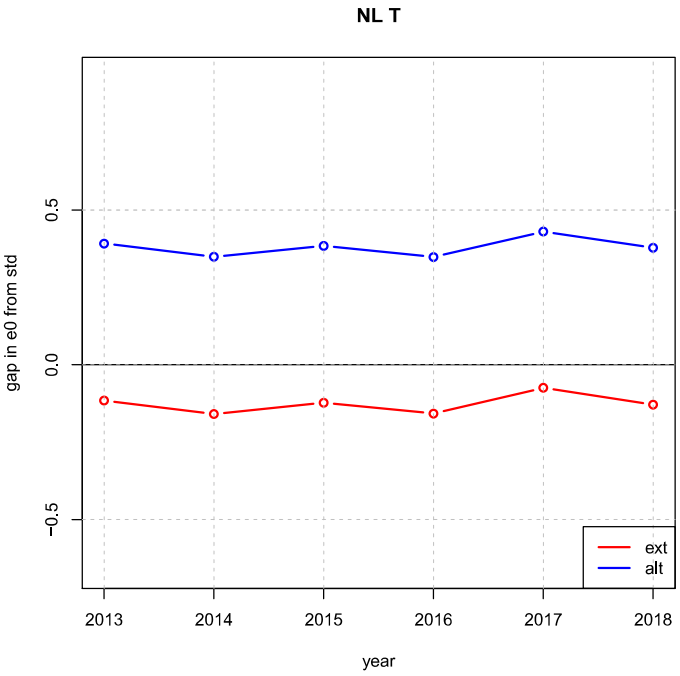

HU F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

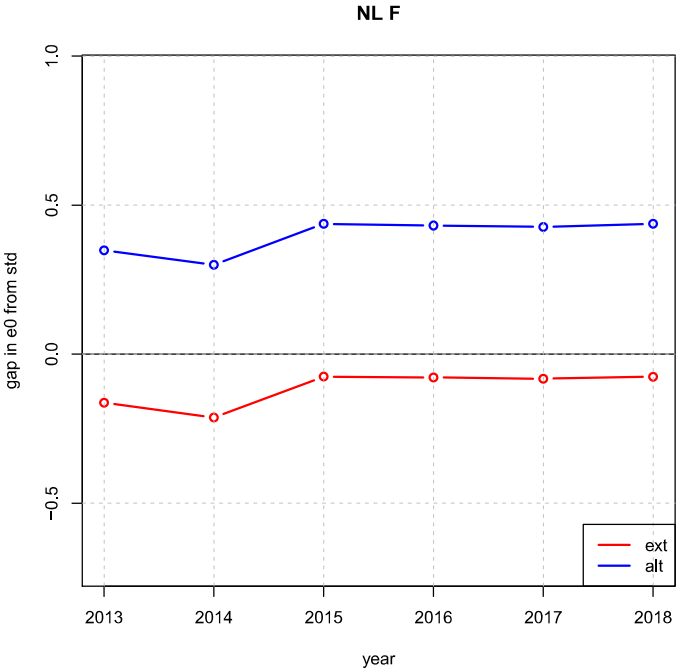

HU M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

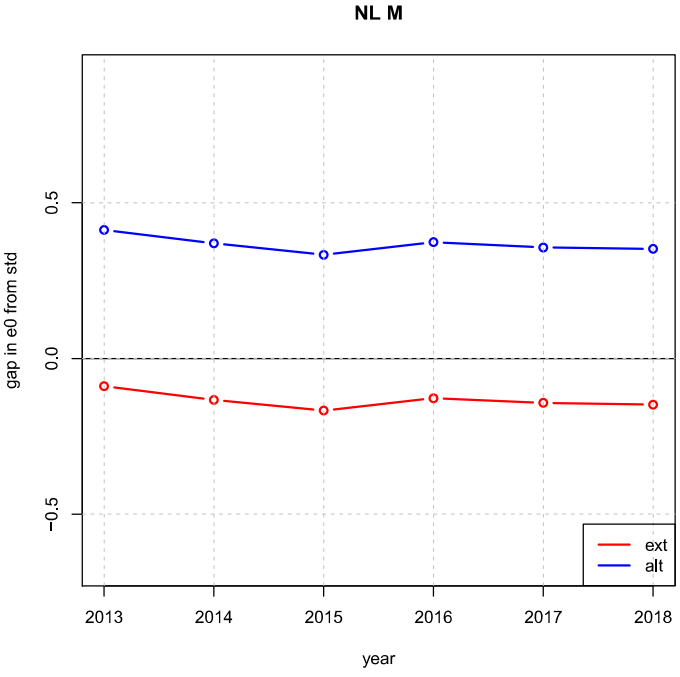

MT T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

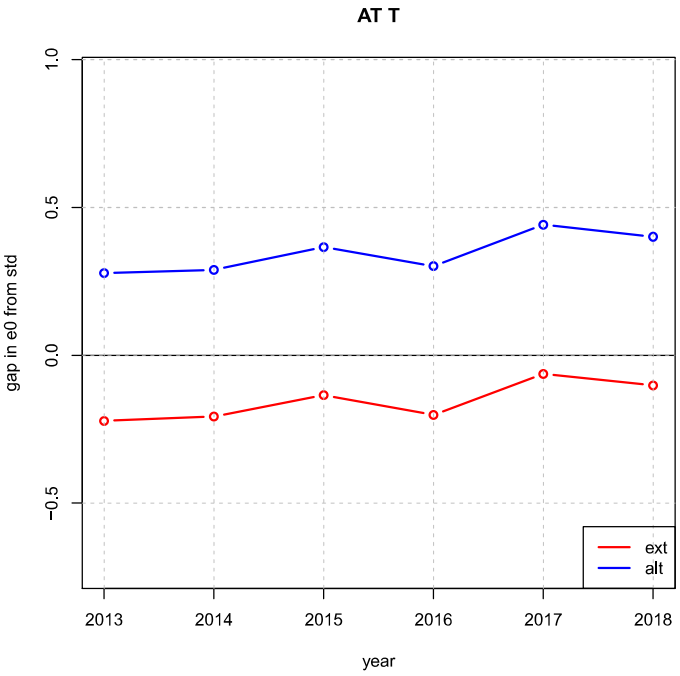

MT F
gap in e0 from std
0.5
0.0
−0.5
−1.0
ext
alt
2013
2014
2015
2016
2017
2018
year

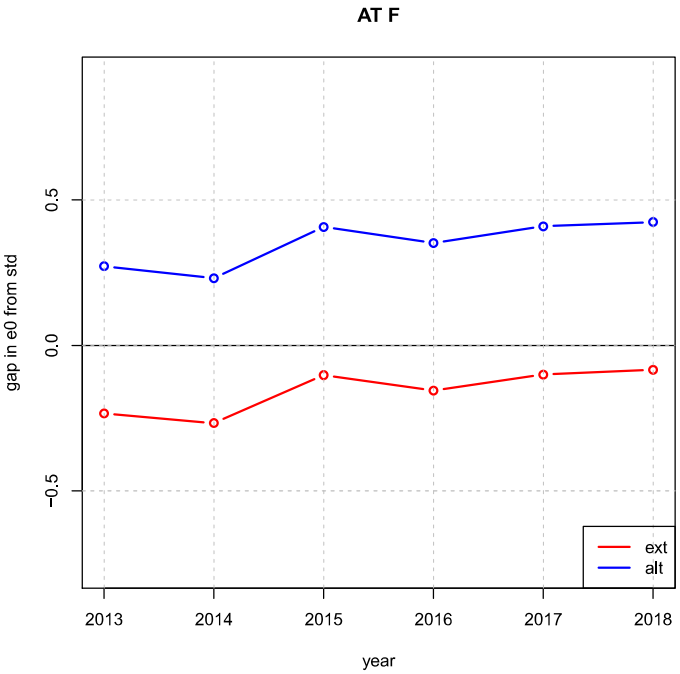

MT M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

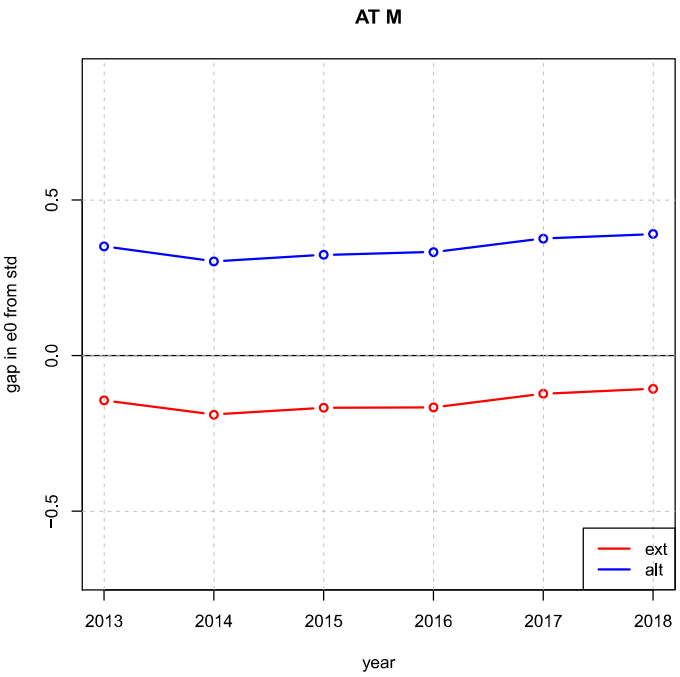

NL T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

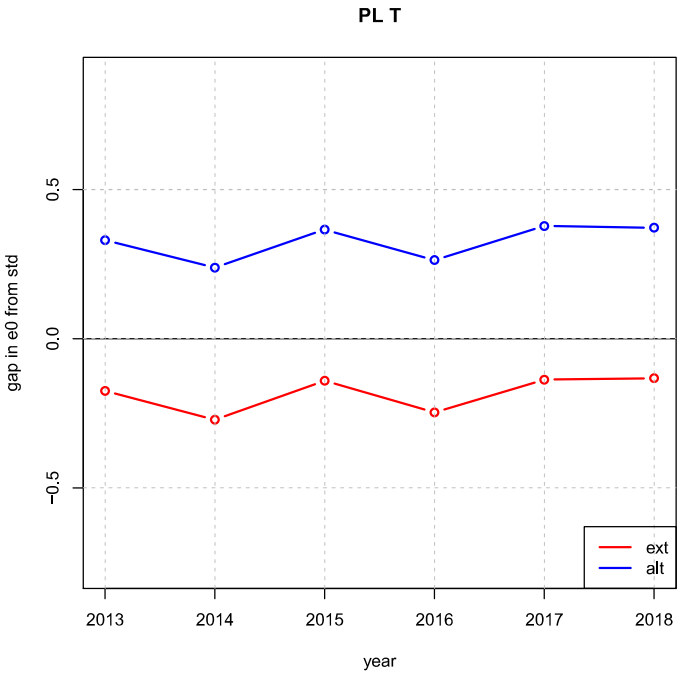

NL F
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

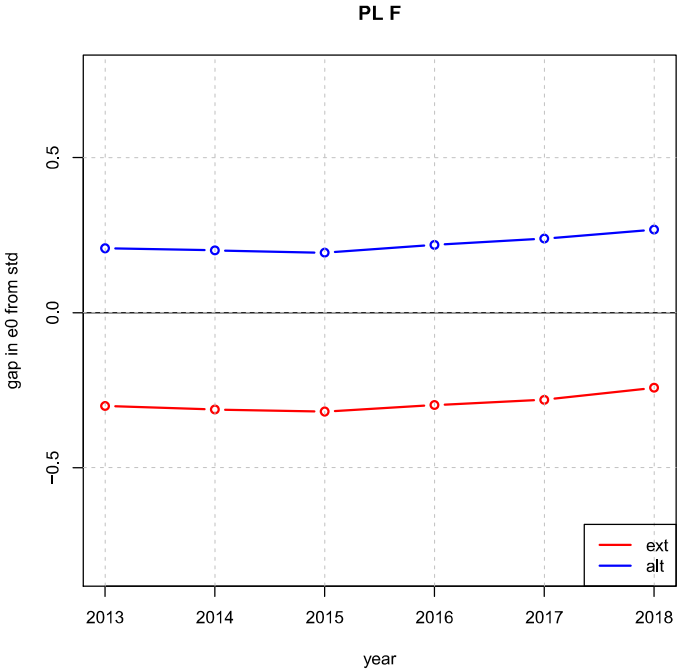

NL M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

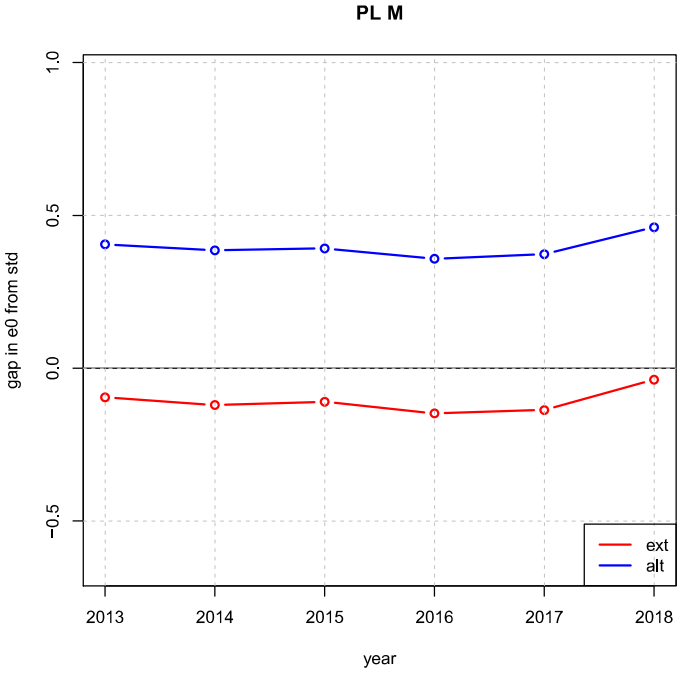

AT T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

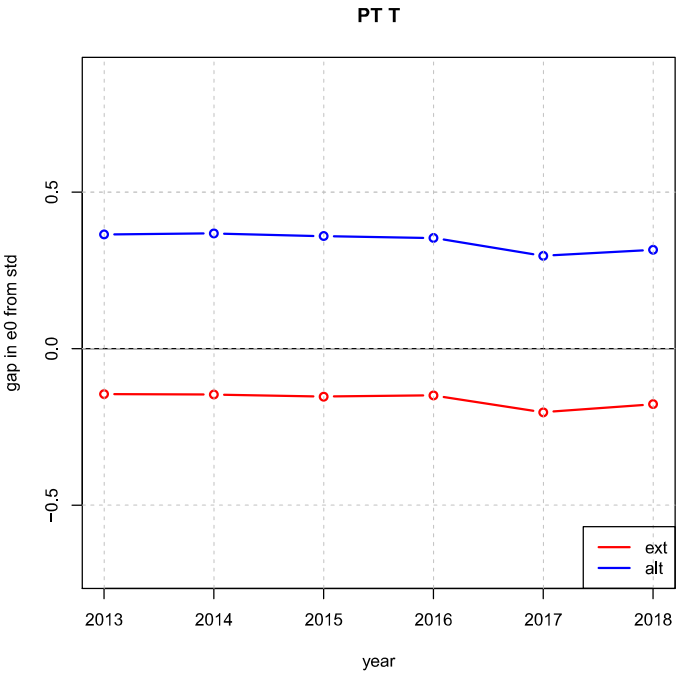

AT F
gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

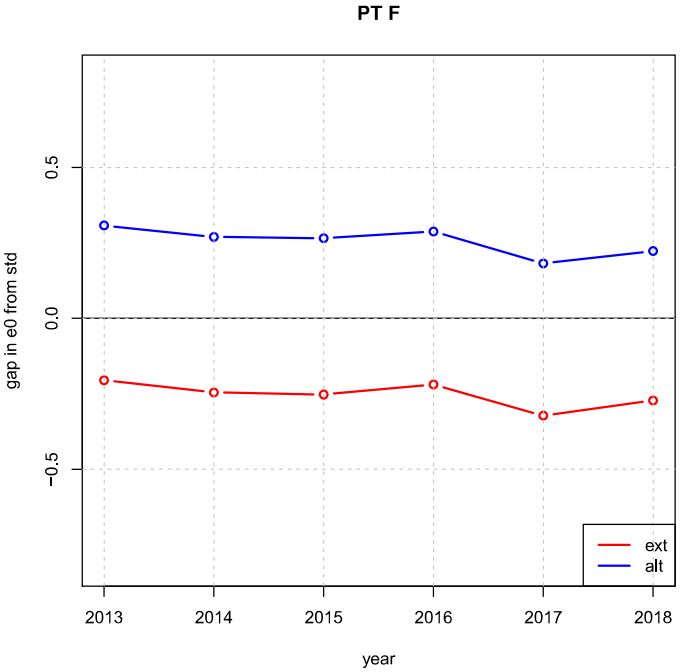

AT M
gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

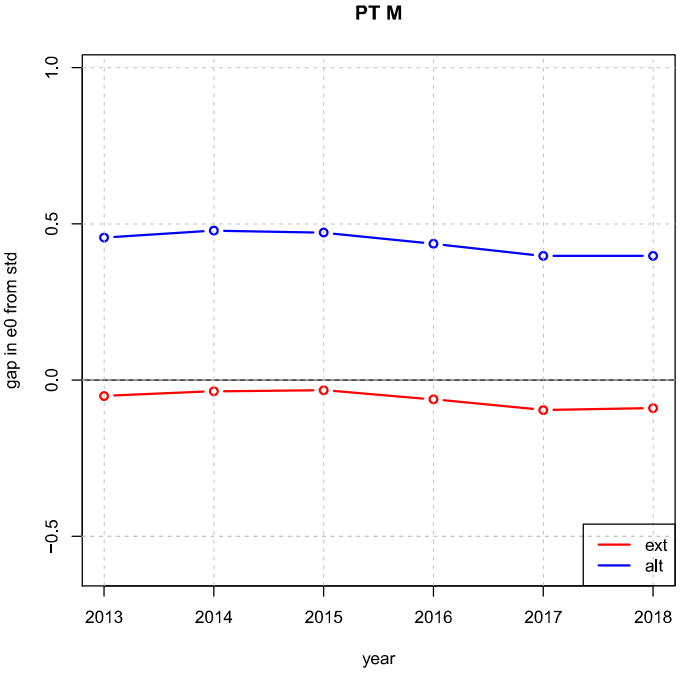

PL T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

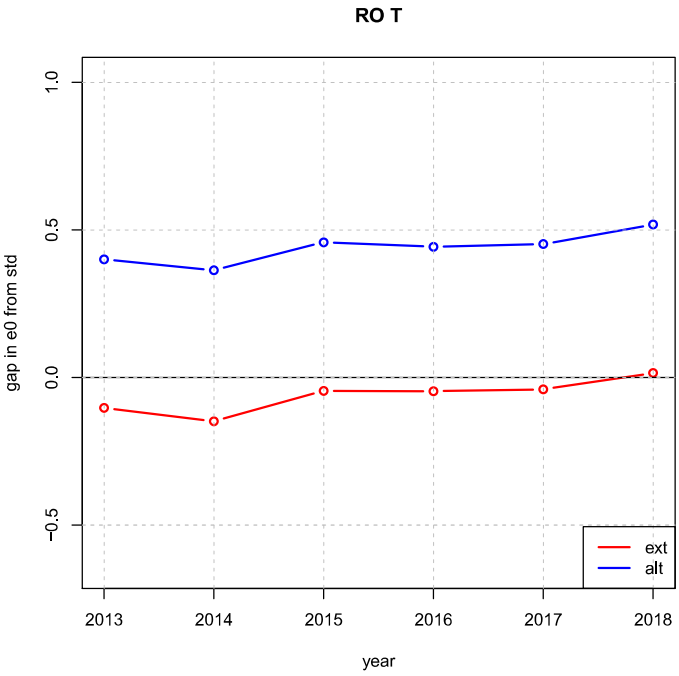

PL F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

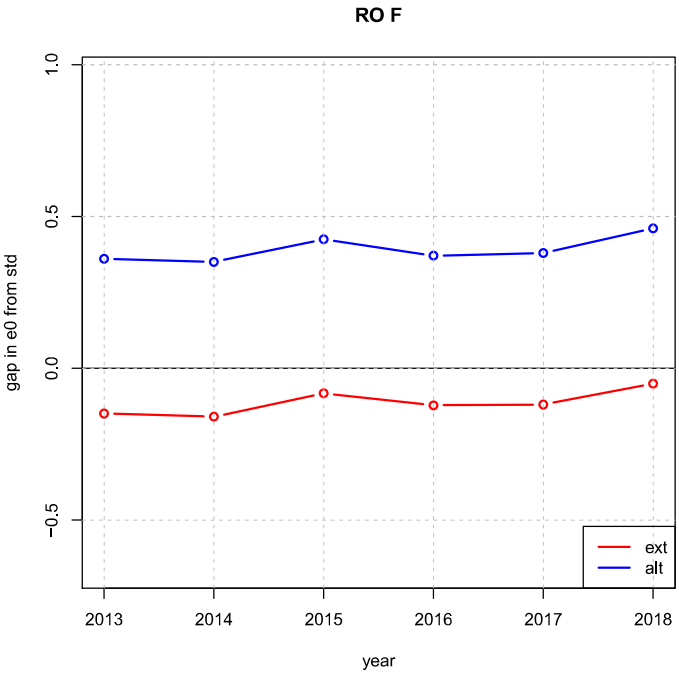

PL M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

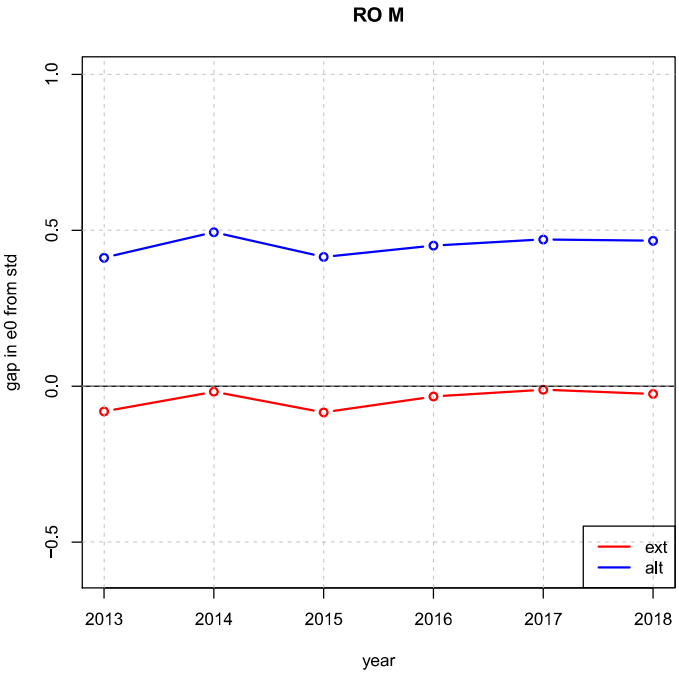

PT T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

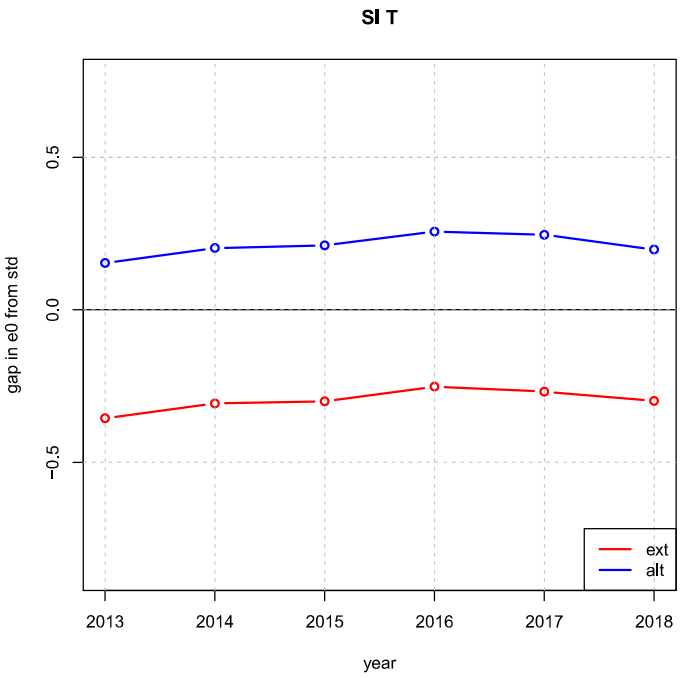

PT F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

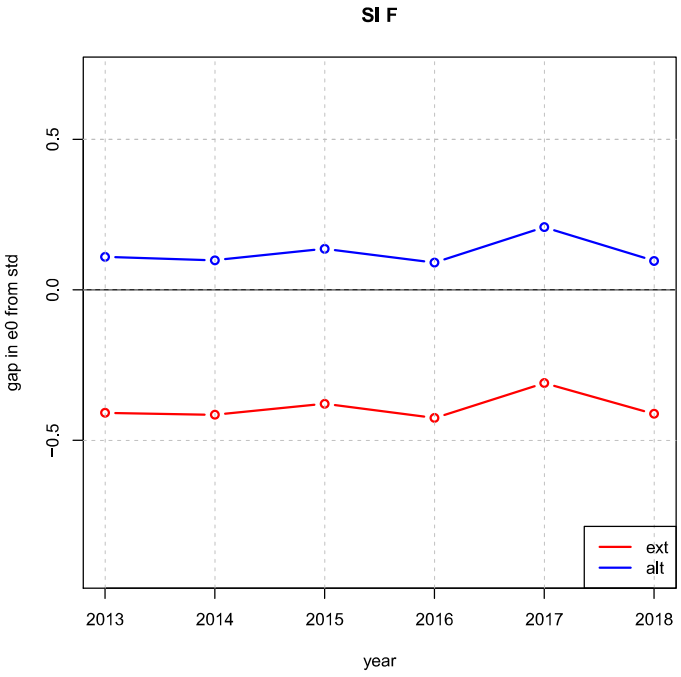

PT M
gap in e0 from std
1.0
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

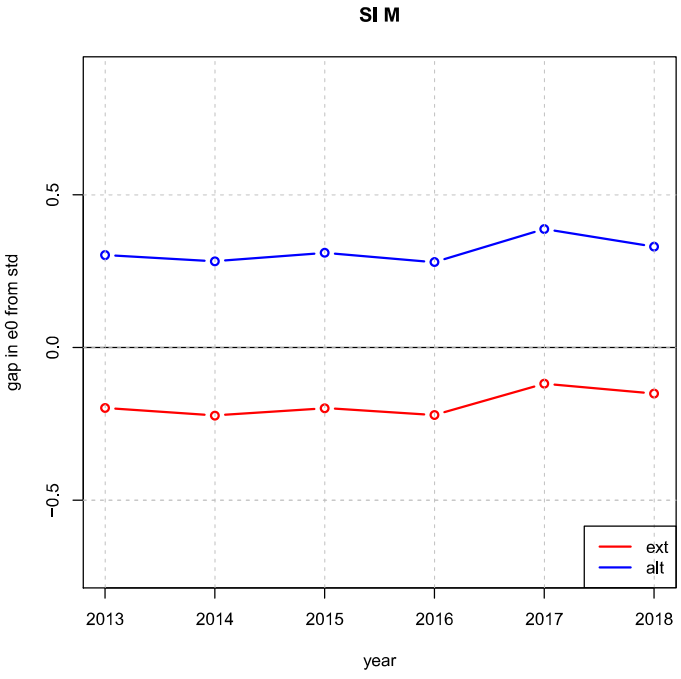

RO T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

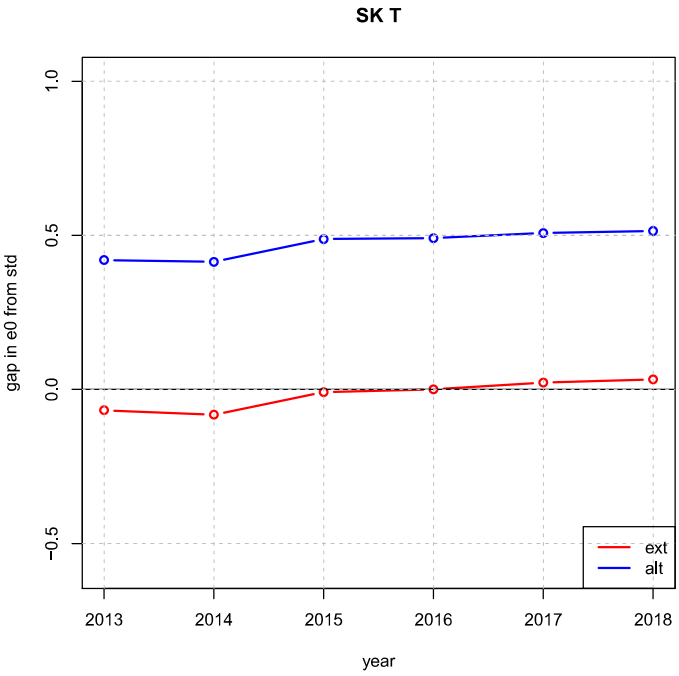

RO F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

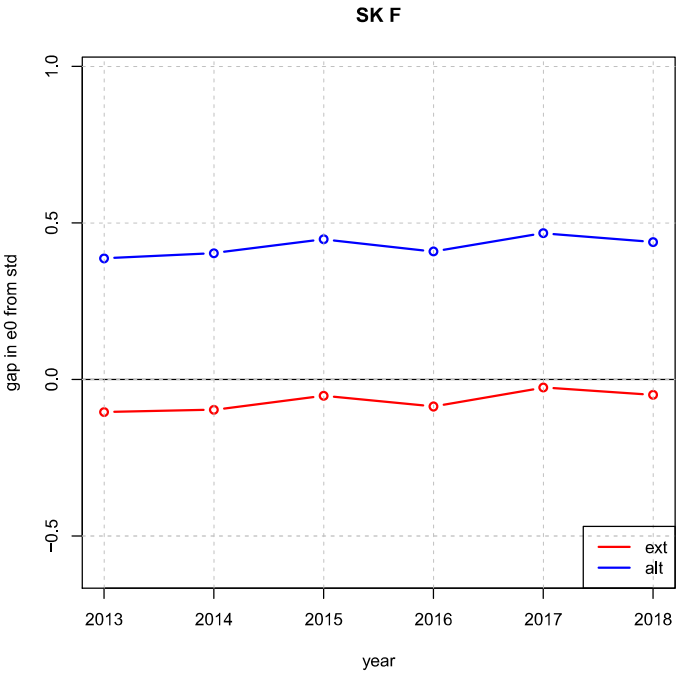

RO M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

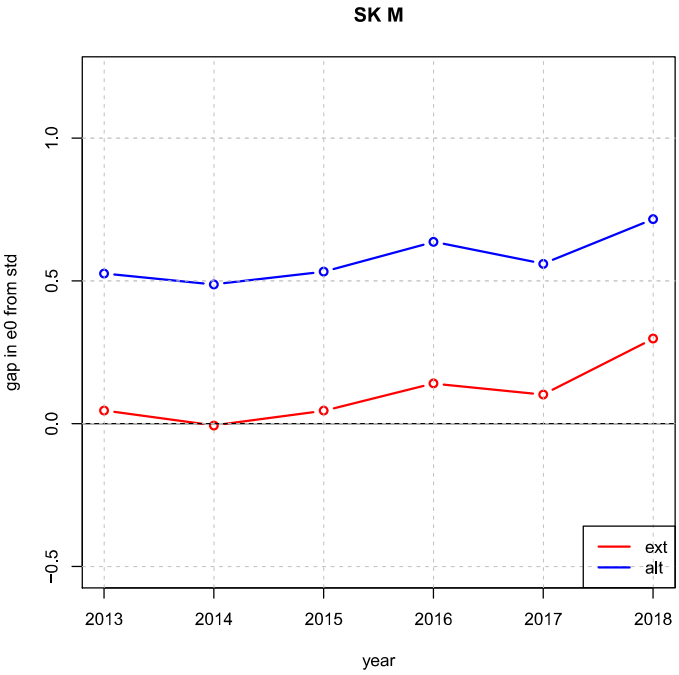

SI T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year


SI F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

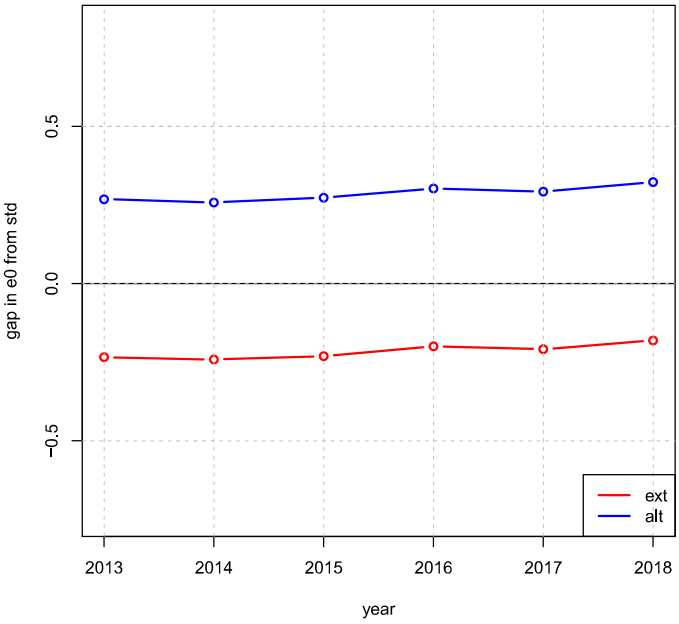

SI M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year


SK T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

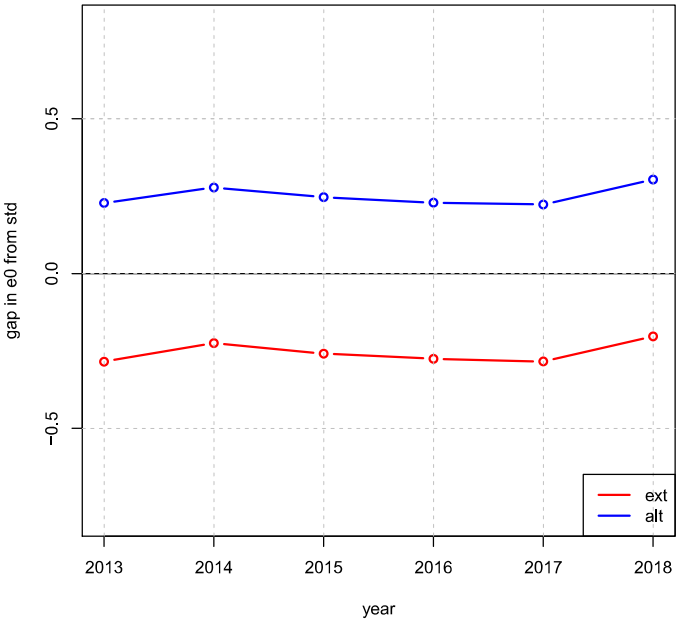

SK F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt


SK M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

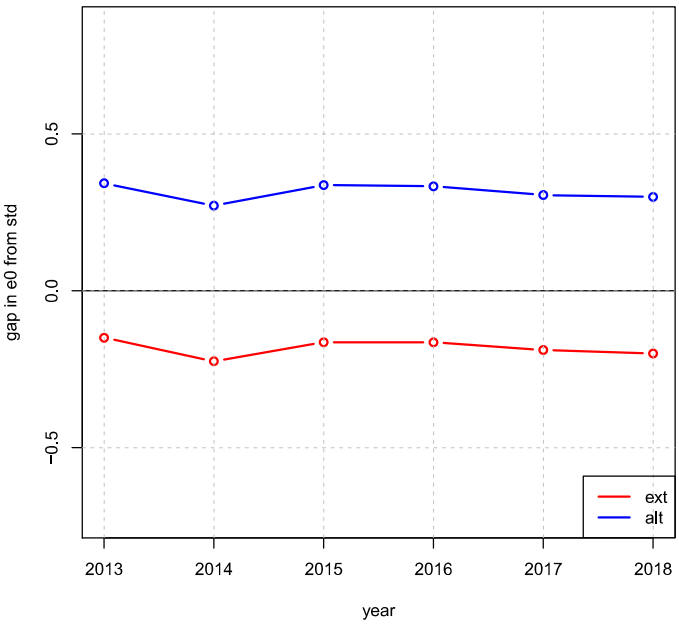
FI T

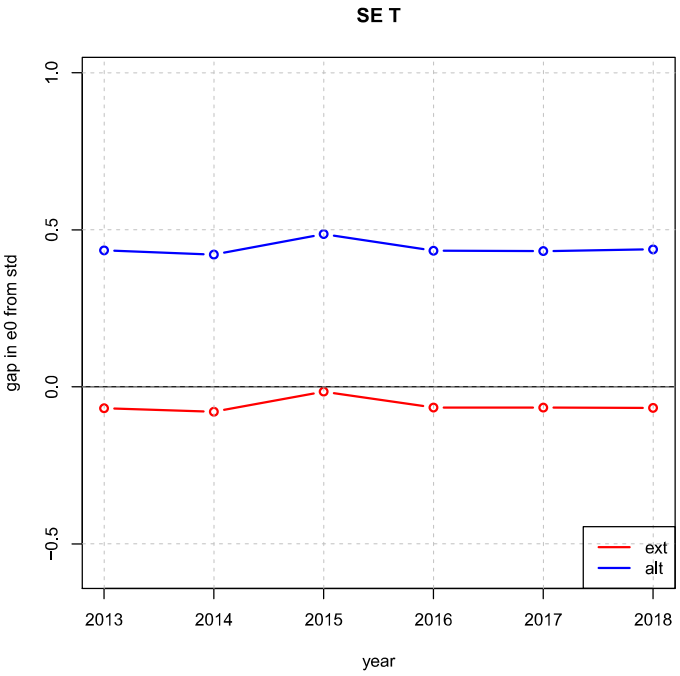

gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

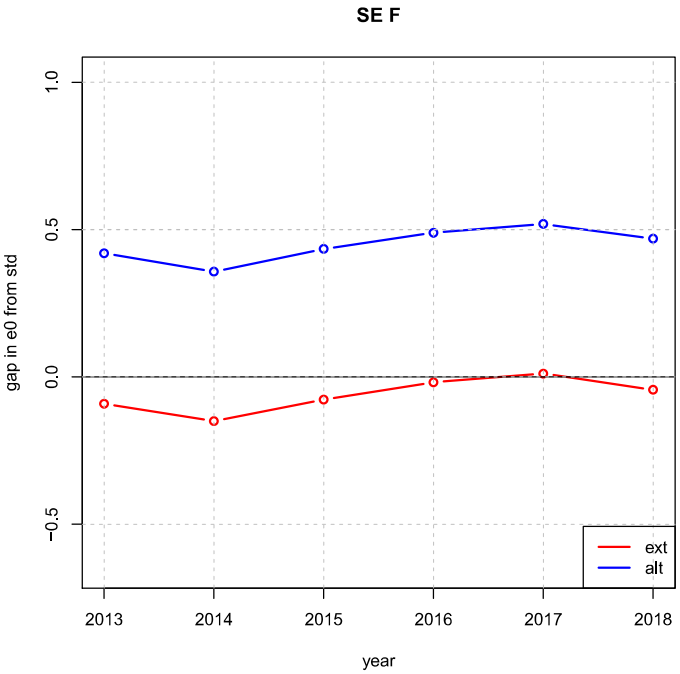
FI F

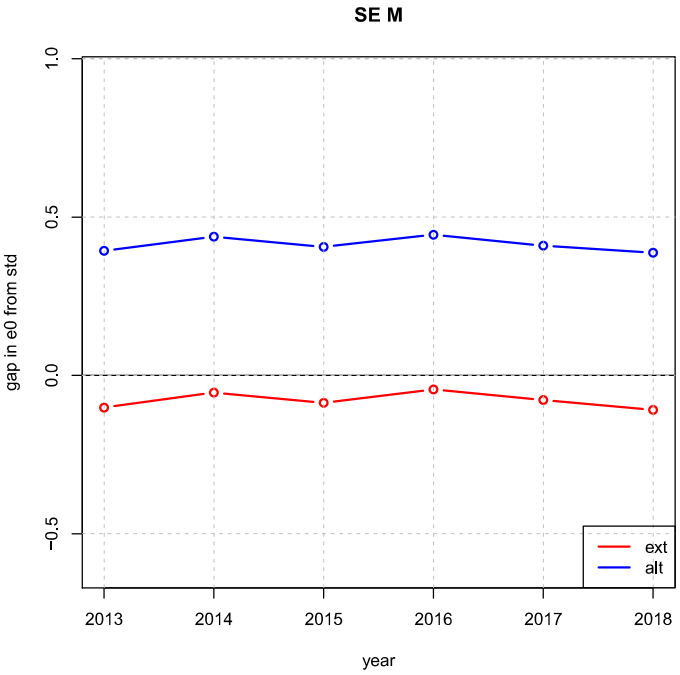

gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

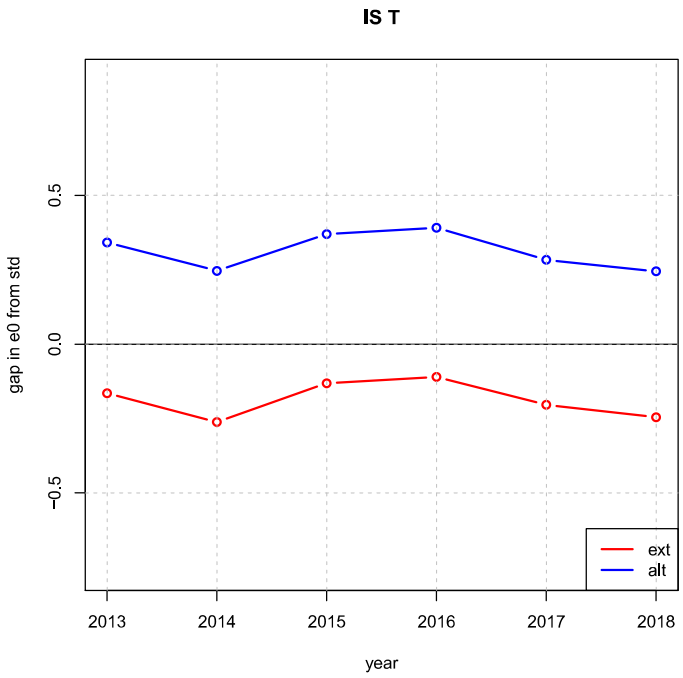
FI M

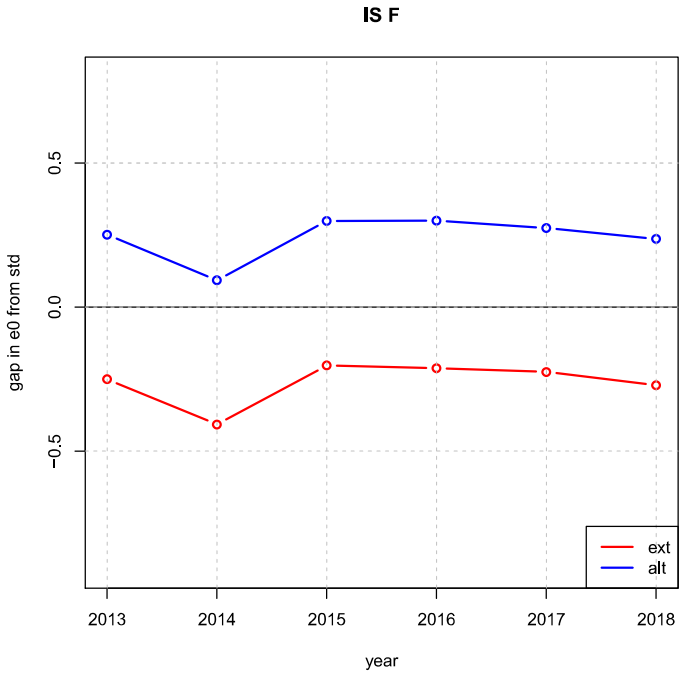

gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

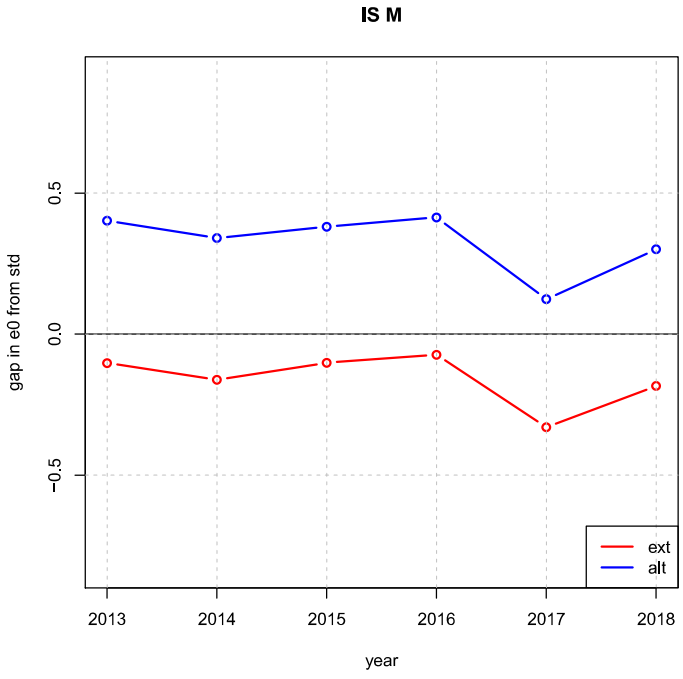

SE T
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt


SE F
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

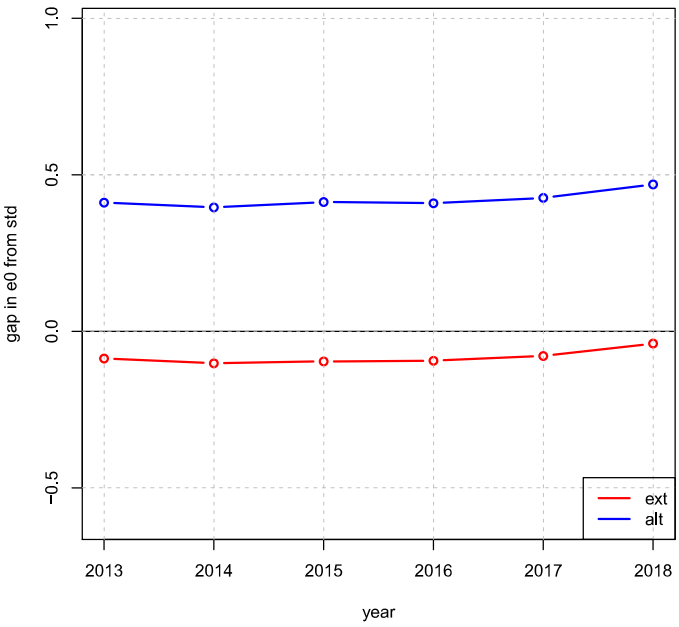

SE M
gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt


IS T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

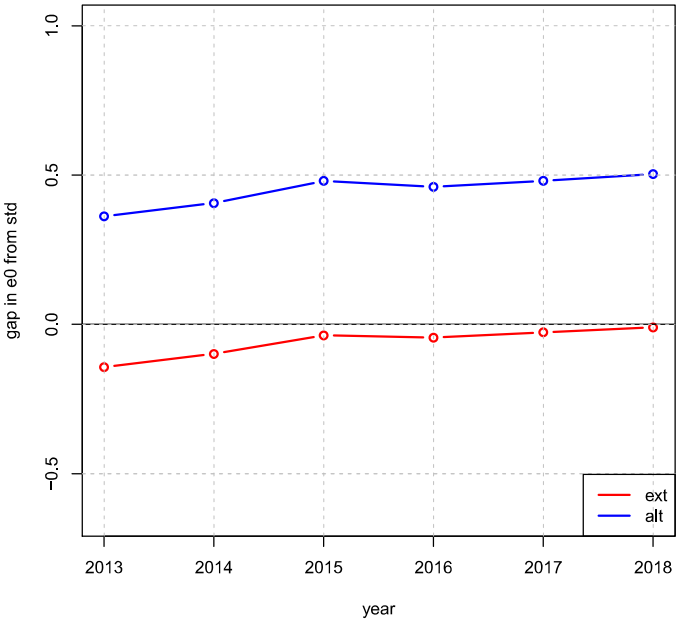

IS F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year


IS M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year

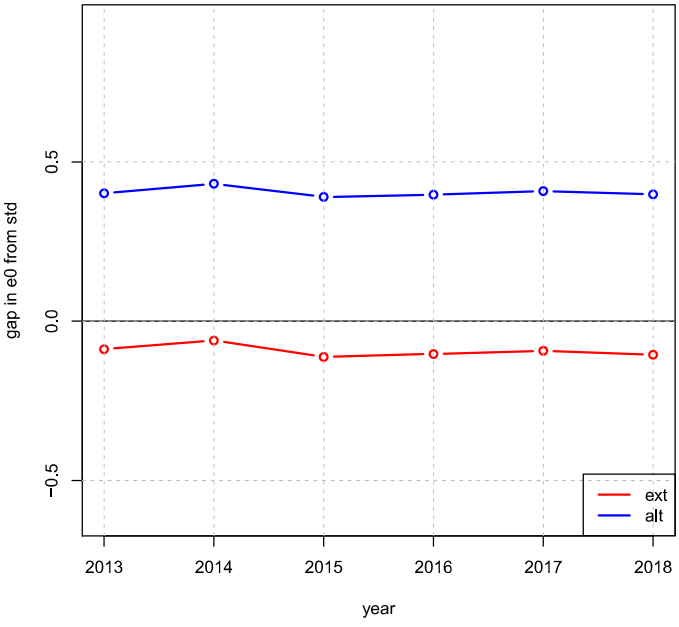
NO T

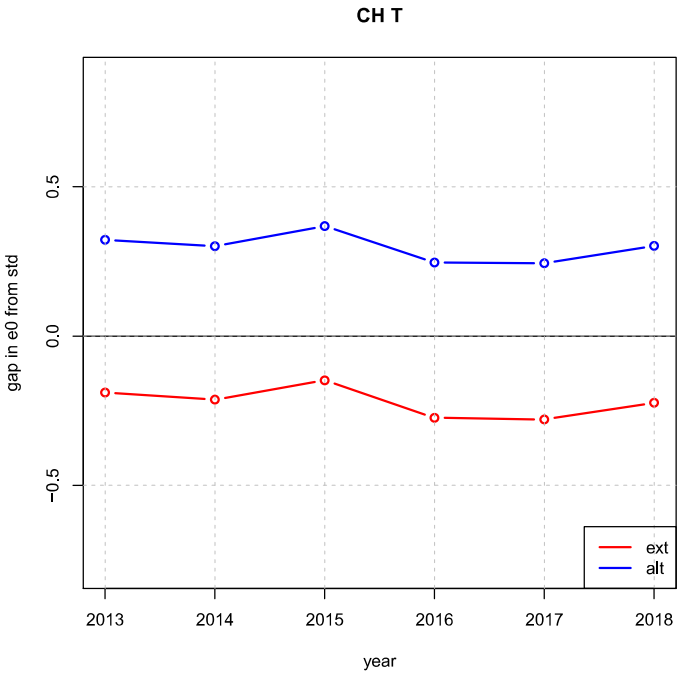

gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

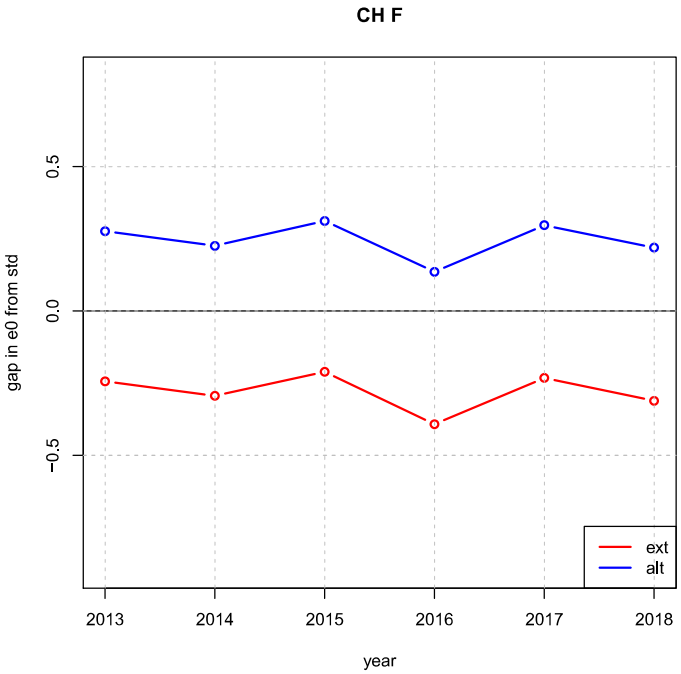
NO F

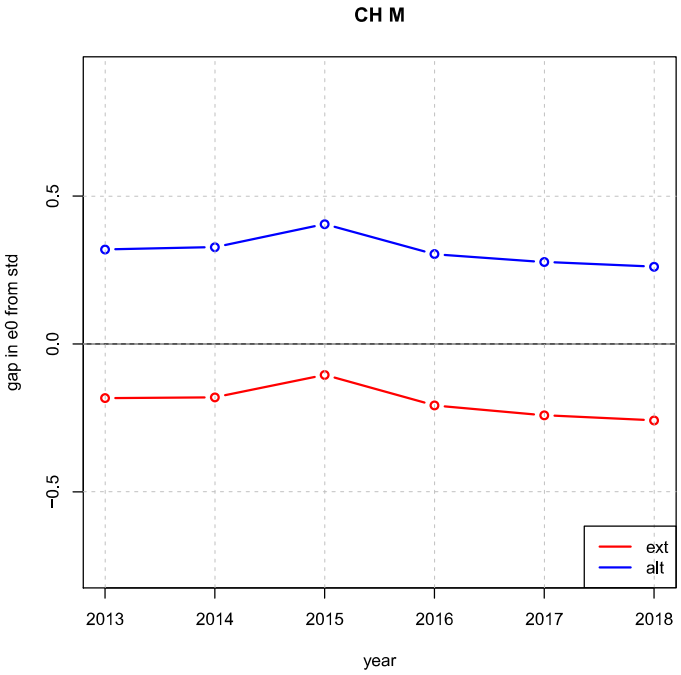

gap in e0 from std
1.0
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt

NO M


gap in e0 from std
0.5
0.0
−0.5
2013
2014
2015
2016
2017
2018
year
ext
alt


CH T
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year


CH F
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year


CH M
gap in e0 from std
0.5
0.0
−0.5
ext
alt
2013
2014
2015
2016
2017
2018
year