EUROPEAN POPULATION CONFERENCE
STOCKHOLM, 13-16 JUNE 2012

Session 70:

Assortative mating: Trends, causes, and consequences

Friday 15 June, 14:00 -15:30

TOWARDS AN ANNUAL MEASURE OF PREVALENCE FOR INTERMARRIAGES IN EUROPEAN COUNTRIES*

Giampaolo LANZIERI

Statistical Office of the European Union (Eurostat)

Draft version of 1 June 2012

PLEASE DO NOT CITE WITHOUT AUTHOR'S PERMISSION

* This paper is released to inform interested parties about research work. The views expressed are exclusively those of the author and not necessarily represent the views of the European Commission / Eurostat.

Introduction

The important migratory flows that have entered in Europe in the recent years and the perspective of immigration as partial solution to ageing populations have brought the issue of the integration of migrants to the attention of policy-makers within the European Union (EU). Two sets of measures, the MIPEX indicators (Huddleston *et al.*, 2011) and the Zaragoza indicators (Eurostat, 2011a), are already important sources of information on several dimensions of migrants' integration, but in none of them is yet included an indicator based on intermarriages. Since very long time, intermarriages are cited in the sociological literature as the litmus test of assimilation of migrants. The first studies on intermarriages date back to the beginning of the past century and researches in this domain were initially motivated by the Great European Migration to America and Australia, where intermarriages analyses then flourished (e.g., Drachsler, 1921). Intermarriages have been defined using various concepts (faith/religion, race, ethnicity, country of birth, citizenship, etc.), but scarce data availability has limited the number and extent of comparative analysis, especially on those European countries where the phenomenon of immigration has become more relevant only in the latest decades. With the growing stocks of migrants of first and second generation, and regardless of the general trends about marriages, intermarriages are now re-gaining scientific interest in Europe under various perspectives (cf. Kalmijn and van Tubergen, 2006; Schuh, 2006; Wray, 2006; van Tubergen and Maas, 2007; Gaspar, 2008; Trilla *et al.*, 2008; Timmerman, 2008; Lucassen and Laarman, 2009; Van Ham and Tammaru, 2011; or the recently started project on "Toward a European Society: Single Market, Binational Marriages, and Social Group Formation in Europe – EUMARR").

Various measures of intermarriage have been already proposed in literature (cf. Lanzieri, 2011), but most of the time focussing on the number of mixed events, and therefore as measures of incidence/flows. As pointed out by earlier scholars (Price and Zubrzycki, 1962a, 1962b), the prevalence of intermarriages is as well a suitable measure for the analysis of migrants' assimilation, although this is a controversial view which several recent studies have confuted. In the more than centenarian tradition of studies on intermarriages, measures of stocks are mainly derived from censuses, which have however a wide time span and are probably not fit for the monitoring of migrants' integration in rapidly evolving situation like in the nowadays EU. In lack of information from population registers or alternative exploitable administrative sources, annual large sample surveys may be a precious source of information. In the current study, two important sample surveys are explored to derive annual measures of prevalence of intermarriages.

The contribute of the study is then primarily in the examination of new data sources for intermarriage analyses. Additionally, it intends to make a first test of few selected measures,

potential candidates to the role of indicators of migrants' social integration. The study also contributes by providing for the first time a comparative overview of the diffusion of intermarriages in European countries.

After a description of the characteristics of the surveys and of the relevant definitions, I justify the use of prevalence measures for intermarriages and I propose four indexes for a first explorative analysis. Results from the two data sources are then compared and a choice is made for the final dataset. Values of the measures are reported for 30 European countries and some potential developments are indicated. A discussion closes the paper.

Coverage, Data and Definitions

The countries included in this study are the 27 Member States of the EU plus 3 countries member of the European Free Trade Association (EFTA), here listed according to the EU protocol order[1] (in brackets their code): Belgium (BE), Bulgaria (BG), the Czech Republic (CZ), Denmark (DK), Germany (DE), Estonia (EE), Ireland (IE), Greece (EL), Spain (ES), France (FR), Italy (IT), Cyprus (CY), Latvia (LV), Lithuania (LT), Luxembourg (LU), Hungary (HU), Malta (MT), the Netherlands (NL), Austria (AT), Poland (PL), Portugal (PT), Romania (RO), Slovenia (SI), Slovakia (SK), Finland (FI), Sweden (SE), the United Kingdom (UK), Iceland (IS), Norway (NO) and Switzerland (CH). The analysis is restricted to the national level.

Data are taken from two important sample surveys carried out annually in the EU: the Survey on Income and Living Conditions (SILC) and the Labour Force Survey (LFS). The former survey is the reference source for comparative statistics on income distribution and social inclusion in the EU and includes a cross-sectional survey of about 130,000 households and 270,000 persons aged 16 years and over (Wolff *et al.*, 2010); the latter, based on a larger sample, has as main statistical objective to divide the population of working age (15 years and above) into three mutually exclusive and exhaustive groups - persons in employment, unemployed persons and inactive persons - and to provide descriptive and explanatory data on each of these categories (Eurostat, 2012).

From the LFS has been selected the variables[2] "Sex" (SEX), "Marital status" (MARSTAT), "Nationality" (NATIONAL) and "Country of birth" (COUNTRYB); from SILC the variables[3] "Sex" (PB150), "Marital status" (PB190), "Consensual union" (PB200), "Country of birth" (PB210) and "Citizenship" (PB220A). Due to the explorative purposes of this study, the variable "Age" has

[1] The EU order of protocol for the Member States is alphabetical, based on the original written form of the short name of each country. The same applies for the EFTA countries.

[2] See Eurostat (2010a) for a detailed description of the variables available in LFS. In brackets the codes used in the survey.

[3] See Eurostat (2010b) for a detailed description of the variables available in SILC. In brackets the codes used in the survey.

not been included. Besides that, another variable of interest could have been the year of residence in the country, which is only available in LFS. No date of marriage is available to identify the newlyweds, i.e. persons whose marriage occurred recently (in relative terms).

For the geographical coverage here considered, some data for the selected variables are not available, either because the survey has not been carried out in that year or because (some of) those variables were not included/elaborated. In particular, over the period 2005-10 and considering the 30 countries listed above, 28 combinations country/year out of 180 are not available in LFS and only 6 in SILC. Extending the period in the past does not bring additional data, and therefore the time period used for the analysis goes from 2005 to 2010. In particular, LFS covers only the EU member States, while SILC covers the whole set of 30 countries; there are also a few missing combinations in both data collections: in LFS[4], data are not available for Bulgaria (2005), Denmark (2005-09) and Sweden (2005-08); in SILC, data are not available for Bulgaria (2005), Romania (2005-06) and Switzerland (2005-07).

Derived variables

Using the variables listed above, from both LFS and SILC have been computed the number of persons for each combination of sex (SEX), legal marital status (LMS), country of birth (COB) and citizenship (CTZ). Each of those variables is split in two categories (plus non responses). Further, each survey has an identifier for the person recognized as partner, regardless of his/her legal marital status. By using that partner identifier, for each respondent have been added the variables related to the sex (P_SEX), marital status (P_LMS), country of birth (P_COB) and citizenship (P_CTZ) of the partner. The first one is necessary because some countries formally recognise also same-sex unions. The marital status of the partner is necessary as well because cases of legally married persons whose partner is not legally married are possible. In fact, the status of married in these two variables (LMS of the respondent and LMS of the partner) do not necessarily implies that the two persons in question are married each other, but this is of course a reasonable assumption.

Features of the two surveys

Between the two surveys there is a clear disparity in terms of sampling size. The Table 7 shows the sample sizes of the two surveys: on average across countries and years, the LFS sample size is about ten times that of SILC, ranging from more than 20 times in Germany and Ireland to about 1-2 times in Estonia, Luxembourg and Slovenia. The much larger sampling size of the LFS, which in principle allows also for finer disaggregation of the variables, would push in favour of its adoption as only data source. However, SILC has a wider geographical coverage, a specific identification of

[4] In LFS, data for Germany classified as non responses on the country of birth have been reclassified as foreign-born for the whole period 2005-10. This assumption has permitted to recover this country from those missing in the analysis by

the consensual unions[5] and offers the possibility of analysis linking intermarriages to income, social exclusion, etc.

Both surveys have a complex sampling design which makes it difficult to get a precise estimate of the sampling variability. SILC has also a longitudinal sample and most of the countries apply an integrated design where part of the sample is interviewed in more than one survey. Some countries adopt a rotational sample also in LFS. However, longitudinal information on the selected variables is not available, and therefore those two surveys will be used only as source of cross-sectional data.

For the sake of simplicity and for the only purpose of a purely explorative analysis of the statistical significance of the differences of the results between the two surveys, I will here assume that those samples are drawn according to a simple random design, therefore neglecting the impact of the design effect on the variance of the estimators[6].

Last point is about the population of reference, which is the population aged 16+ living in private households for SILC, and almost the same for LFS, where the age range is in most of the cases 15+, or 15-74 (e.g., see Eurostat, 2012; Table 1 on p.63). Although in almost all European countries the minimum marriageable age is 18 years old, there are exceptions which allows the marriage between persons of minor age, in most cases down to 16 years old (Eurostat, 2003; p.74). Therefore the population of reference of the two surveys corresponds by and large to the one in marriageable age (slightly more precise in SILC). For sake of simplicity, I will not implement here corrections for the single age(s) of difference between LFS and SILC, also considering the intrinsic nature of approximation of the sample surveys.

Further, data from LFS are in fact from quarterly samples. Data should then be interpreted as person-years of permanence in the status (intermarried, etc.), rather than as number of persons at a given moment. For the sake of simplicity, I consider data from SILC, whose surveys are carried out in (and refer to) different dates of the year across countries, as an estimate as well of the person-years. This conceptual distinction is relevant whether those data are used to estimate the exposure to the risk of marriage.

Definitions

Between the definition of intermarriage based on country of birth or citizenship, I here privilege the former because it is a permanent characteristic of the individual, while citizenship may change over time and it is largely affected by the naturalisation practices in the various countries, especially

country of birth.

[5] In fact, this could be indirectly estimated also in LFS, assuming that all those persons for whom there is not a partner identification number are single, and taking the difference from the total population. This procedure would be possible also in SILC, which however has a specific variable collected during the survey. Thus I consider SILC as better source for consensual unions.

[6] The EU regulation on SILC expresses the required sample sizes in terms of "minimum effective sample sizes", meaning the size required if the survey was based on simple random sampling (see also Eurostat, 2010b; p.24).

considering that one important channel of acquisition of citizenship is by transfer (i.e., following the marriage with a national)[7].

In consideration of the level of aggregation of the information about country of birth, mostly available without reference to the specific foreign country of birth, the measures can not be focussing on specific foreign group(s), an approach often taken in the literature. The perspective must be changed to one looking to the native-born persons and to their social interaction (in terms of marriage) with those who are not native-born, regardless of their specific country of birth. By doing so, there is definitely a loss of precious information as dynamics can be very different according to the groups in play, but it is a price to pay to gain an overall view.

One of the peculiarities of the studies on marriages is the identification of the proper unit of analysis, which can be based either on couples or on individuals. In the present study, a married person is identified as someone who is classified as married in the legal marital status and has a partner whose legal marital status is also "married". This is different from simply taking the number of married persons classified from the only variable "legal marital status", as it cleans that figure from all cases where the married person is in relation with somebody who is not the "originally married" partner, or from the cases where the married person has not anymore a relationship, although being still formally married. By using this definition, there is a straightforward relation between the number of individuals and the number of couples. However, such relation may not hold when disaggregated by sex, because in some European countries same-sex marriages are formally allowed. In the measures presented below involving the breakdown by sex, it is assured the consistency with the individuals involved in same-sex relationships.

Therefore, in this study, intermarried (or outmarried) persons are those legally married native-born persons living with a legally married foreign-born partner (same or opposite sex) in private households.

Measures

Prevalence vs. incidence

In the scientific literature, there are several scholars who criticise the use of prevalence measures for analyses on intermarriages. Among others, Kalmijn (1998) states that incidence measures are preferable for the analysis of trends; van Tubergen and Maas (2007) clarify that the problem of using cross-sectional data is that they include marriages occurred abroad, do not include marriage dissolutions and there is an inherent difficulty to assess causality, because information is usually available as at the time of the survey and not at the time of marriage. Trilla *et al.* (2008) stress that

[7] A more detailed description of the peculiarity of intermarriage data based on citizenship is given in Lanzieri (2011, 2012).

the prevalence nature of census data does not allow the interpretation of the results according to the prevalent theories on intermarriages, such as the classical or segmented assimilation and the social exchange. It must be noted that there are also scholars who promote the use of prevalence measures as best practice (Price and Zubrzycki, 1962a, 1962b) or that consider them as valid as incidence measures (e.g., Schwartz and Mare, 2005)

All the above criticisms are certainly proper whether the purpose is to understand which factors are influencing the occurrence of intermarriages. However, what a prevalence measure says is a different part of the story. As stated by Kalmijn (1998:397), the sociological interest towards intermarriage it is not only in their function of indicator of boundaries between groups in the society, but also in their "…potential of cultural and socioeconomic change." An assessment of such potential can not exclude the extent of diffusion of this phenomenon in the society of interest, regardless of what are the current trends. A prevalence measure informs about the cumulative effect of past trends on the current population composition (by marriage). For the purpose of a view at such potential of change, it is not that relevant if a marriage has taken place in the country or abroad[8]: what it matters, it is the role that the (intermarried) couples (and their descendants) play everyday in the society. It is neither a problem if an intermarried couple dissolved in the meanwhile because, as noted by Kalmijn (1998), if mixed marriages are more likely to break up then conclusions about openness of social groups based on occurrences may need to be revised. Although it may not tell us *why* they intermarried, a prevalence measure tells us how many intermarried persons are acting at that moment in the society[9]. Thus, if one intends to look at a measure for the *monitoring* (and not necessarily for the *understanding*), prevalence can be also informative. From this point of view, restricting the monitoring to the only events may be misleading, as there is no information about the future permanence of the newlyweds in the country of occurrence. Further, it must also be noted that van Tubergen and Maas (2007) find that the use of stock data, which include older marriages, does not affect the conclusions on the individual and contextual predictors.

Another issue using prevalence data is that those residents who are in endogamous marriage may actually have never been exposed to the risk of intermarriage in that society. This concerns more the foreign- than the native-born persons (e.g. a foreign-born who immigrated being already married in his/her country of birth), especially in countries with long history of immigration. Those persons would not be part anyway of the estimated population at risk of marriage (i.e., the unmarried) and

[8] This issue is instead considered a drawback in those studies where the variable of interest is the endogamy of the foreign-born persons (i.e., the extent of the intra-marriages within specific foreign-born populations), because those foreign-born may have never been exposed to the risk of intermarriage.

[9] They could be also used to explore the *consequences* of intermarriage, although for certain hypotheses it may be difficult to disentangle the cause from the effect using those data. For an example of study on the effect of intermarriage on economic assimilation (intermarriage premium), see Meng and Gregory (2005).

thus there would be virtually no consequences on a *rate* of intermarriage[10], while the problem exists if those persons are included in the denominator of a *proportion*, as they have an inflating effect. From the point of view of the (current) social influence of potential factors of change, it may be irrelevant whether those persons have or have not had in the past the opportunity of intermarry: the matter is simply that they are not intermarried. However, from the perspective of migrants' integration, it must be recognised that a given proportion of intermarriage has a different meaning if those persons at the denominator have chosen not to intermarry, or if they never had the chance. Therefore, a drawback still remains if one uses only prevalence data, but those stock data may be fruitfully used as input for more complete indexes of intermarriages incorporating flow data as well, as it will be shown below.

Four simple measures

Several measures for intermarriages incidence are analysed and proposed in Lanzieri (2011), which could be as well used for intermarriage prevalence. I here select few of them whose results are of easier communication to a wider audience.

The first measure is n_m^o / n_m , the proportion of intermarried persons n_m^o (where the exponent *o* stands for out-married) on the total number of married persons n_m, which can be interpreted also as the number of intermarried couples c_m^o on the total number of married couples c_m:

$$p_1 = \frac{n_m^o}{n_m} = \frac{n_m^o / 2}{n_m / 2} = \frac{c_m^o}{c_m} \qquad [1]$$

The measure [1] is independent from the prevalence of marriages/married persons in the population. However, a high proportion of intermarried persons would have a different meaning if marriage is still a popular and widespread form of relation or if instead it is an institution of decreasing importance. Incorporating this latter factor in the measure [1] gives the following:

$$p_2 = \frac{n_m^o}{n_m} \cdot \frac{n_m}{n_{16+}} = \frac{n_m^o}{n_{16+}} \qquad [2]$$

where n_{16+} is the number of individuals in marriageable age living in private households. Therefore, for a given proportion of intermarried among married persons, the higher the proportion of married persons in the population, the more important the mixed marriages, and vice versa. While the measure [1] gives the extent of intermarriages focussing on those who have made this important life choice, the measure [2] incorporates the extent of diffusion of marriage in the population of reference. The former could be seen as a general indicator of migrants' integration, the latter is more an indicator of the extent of potential forces of social and cultural change in play at a given moment. Both proportions p_1 and p_2 take values in the interval [0, 1], the maximum value being reached for

[10] I consider here rate in demographic sense, i.e. as occurrence/exposure.

the former when all couples are intermarried and for the latter it is necessary the additional condition that everybody in the population of marriageable age is married. In practical terms, p_2 will never reach the maximum theoretical value and it will always holds $p_2 < p_1$.

Both the previous measures are still affected by several important factors which influence the opportunities of marriage. The age and sex structure of a population may or may not be favourable to (opportunities of) marriages, an effect which is generally known as marriage squeeze, and adding the split between host population and immigrants introduces a further potential "squeezing" factor.

Neglecting for the sake of simplicity the age breakdown, each of the above measure can be expressed with reference to a specific sex *s* and/or population group *g* (e.g., native-born/foreign-born or national/foreigner, where foreign-born or foreigner may also refer to a selected country):

$$s_1^{s,g} = \frac{n^o_{m,s,g}}{n_{m,s,g}} \qquad [3]$$

and

$$s_2^{s,g} = \frac{n^o_{m,s,g}}{n_{16+,s,g}} = \frac{n^o_{m,s,g}}{n_{m,s,g}} \cdot \frac{n_{m,s,g}}{n_{16+,s,g}} = s_1^{s,g} \cdot \frac{n_{m,s,g}}{n_{16+,s,g}} \qquad [4]$$

each of which can be computed for the four combinations of sex and group.

To get a synthetic measure from those sets, the easier way is to compute their average over sex and group, as done in Lanzieri (2011) for a parallel measure on events:

$$s_1 = \frac{1}{4} \cdot \sum_{\forall s,g} s_1^{s,g} \qquad [5]$$

and

$$s_2 = \frac{1}{4} \cdot \sum_{\forall s,g} s_2^{s,g} \qquad [6]$$

Like the previous proportions, also the measures in [5] and [6] take values in the range [0, 1]: the higher their value, the higher the proportion of intermarriages across the specific subgroups. If the averages in [5] and [6] would be weighted by the respective proportions of the subgroups on the total married or total population, the outcome would be respectively the indicators in [1] and [2].

Let's now focus on the first of them. The minimum value is reached when there are no intermarriages in any subpopulation, and this is straightforward. The maximum value is instead reached when all married persons within each subpopulation are intermarried and "corresponding" subgroups are of equal size[11]. In fact, a man of the group 1 who intermarries must have one "corresponding" woman in the group 2, and vice versa. Therefore, as to get an average value of one it is necessary that all components $s_1^{s,g}$ are equal to one as well, there must be a perfect matching between groups and sexes: all married men of group 1 are intermarried to all married women of

[11] This condition applies for opposite-sex marriages. It is here assumed that same-sex marriages have a negligible impact on the overall number of marriages.

group 2, and all married men of group 1 are intermarried with all married women of group 2. In formula:

$$s_1 = 1 \quad iff \quad \left(n^o_{m,1,1} = n_{m,1,1} = n^o_{m,2,2} = n_{m,2,2}\right) \cap \left(n^o_{m,2,1} = n_{m,2,1} = n^o_{m,1,2} = n_{m,1,2}\right) \qquad [7]$$

When the population of reference is the married persons, this is quite intuitive. Let's now move to the second measure, s_2. Also in this case the maximum value of the measure is reached if and only if each of its components is equal to 1. Adapting the equation above accordingly, it gives:

$$s_2 = 1 \quad iff \quad \left(n^o_{m,1,1} = n_{16+,1,1} = n^o_{m,2,2} = n_{16+,2,2}\right) \cap \left(n^o_{m,2,1} = n_{16+,2,1} = n^o_{m,1,2} = n_{16+,1,2}\right) \qquad [8]$$

which means that the size of the foreign-born group in one sex must equal the size of the native-born group of the opposite sex. Therefore, the measure s_2 can reach the maximum only when the relative size unbalances are not present in the population. The sex ratio can instead still be a squeezing factor within a group, because it may be $n_{16+,1,g} >> n_{16+,2,g}$, and vice versa. To be precise, these contextual factors (relative sizes and sex ratios) are relevant for the occurrence of the events (marriage) and could still be considered as explicative elements for newlyweds. They are in principle less constraining for older marriages, because the context may have changed over time.

If the target is to get an indicator of migrants' integration, an alternative could be to consider only the proportions referring to foreign-born, such as:

$$s_1^* = \frac{1}{2} \cdot \left(\frac{n^o_{m,1,2}}{n_{m,1,2}} + \frac{n^o_{m,2,2}}{n_{m,2,2}} \right) \qquad [9]$$

which would take the maximum value of 1 if both foreign-born married men and foreign-born married women are all intermarried[12]. This measure would not incorporate the effect of different relative sizes of the group of married persons (immigrants and host population), but in small communities of immigrants may rise quite rapidly to high values, which may be misleading. Let's consider the case of a country A where the few immigrants have no choice but to intermarry (if they want to get married), and a country B where the population of immigrants has a much more relevant size (in comparison to the host population), and the immigrants all intermarry as well. The two examples are resumed in the following Table 1:

[12] A parallel measure referring to the total population in marriageable age as indicator of factors of change may be of less interest, because such forces should act in both groups and the native-born side would not be included.

Table 1: examples of distribution of married couples by country of birth of the partners							
		Women					
		Country A			Country B		
		Native-born	Foreign-born	Total	Native-born	Foreign-born	Total
Men	Native-born	998	1	**999**	500	250	**750**
	Foreign-born	1	0	**1**	250	0	**250**
	Total	**999**	**1**	**1000**	**750**	**250**	**1000**

Using the measure [9], it would be $s_1^*(A)=1$ and $s_1^*(B)=1$ as well; using the measure [5], $s_1(A)=0.50$ and $s_1(B)=0.67$. Which one of the two measures better fits the needs of integration monitoring and comparative analyses should be matter of further thoughts. For the time being, I will limit the study to the measures presented in [5] and [6].

Comparison of the results from the two surveys

Taking into account the remarkable difference of the respective sample sizes, LFS has a competitive advantage on SILC and it is therefore initially considered as main data source. The Figure 1 shows the differences between the estimated proportion of intermarried couples on total number of couples between LFS and SILC. Results for Bulgaria, Denmark, Romania and Sweden refers to the years available in both surveys, and the EFTA countries have been excluded from this comparison for the same reason. For 10 countries (up to Italy starting from the left side of the graph), the average difference is positive, meaning that results from LFS indicates a higher proportion of intermarried even up to 4 percentage points (p.p.) more. However, the differences do not point always to the same direction. For instance, in Latvia the indicator p_1 was estimated to be higher in SILC than in LFS by about 1 p.p. in 2005, but (much) lower from 2007 on. Obviously, a difference of – let's say – 1 p.p. has a diverse importance for countries in which the estimated proportion is low or even close to zero than in those cases where the indicator p_1 takes values of 10-20%. Under the simplifying assumption of a simple random sampling design and without any correction for multiple comparisons, a simple statistical test on the differences between estimated proportions from LFS and SILC results significant in many cases – but not all, which is remarkable considering the large difference between their sampling sizes.

Occasionally, the indicator takes suspicious values, which do not fit in the time trends of the countries, or there is a clear break in series. For instance, in SILC, in the Netherlands and in Portugal the proportion double from one year to another, in the latter case remaining stable on that higher level afterwards, change which does not appear in LFS; or, on the other side, in LFS there is a sudden drop in one year for Belgium, with no correspondence in SILC.

Looking at the indicator p_2, the differences between data sources are narrower than for the previous measure. The set of countries on the left side of the Figure 2, from Denmark to Poland, have on average higher proportions of intermarried on the total population from LFS than from SILC, and vice versa for those from Romania rightwards. However, for all countries but Denmark and Germany (which are on the two extremes of the Figure 2), such average differences are within a range of ±2 p.p. This is to be expected, because the difference between p_1 and p_2 is only in the denominator, larger for the latter measure. Again, under the same very simplifying assumptions used above for the statistical tests[13], the differences are not always statistically significant.

Moving to the two measures based on average proportions across sex and group (native-born/foreign-born) composition of the population, makes a change. The differences between surveys are larger, and look particularly problematic for Germany and Romania, where they are also higher than 15 percentage points. More caution is therefore here necessary before proceedings to the use of SILC as replacement of LFS. The same applies when the average proportions are referred to the total sub-populations, regardless of their marital status, where the differences between surveys are however less important (see Figure 4).

The final dataset is then composed by the estimates from LFS, while those from SILC are used in case those from LFS are missing or judged unreliable. Further, in order to keep the time series coherence, it may be as well convenient to use SILC as data source whether LFS has only spare values available for that country. For all indicators here considered, the data for Denmark, Sweden, Iceland, Norway and Switzerland are all taken from SILC, so that the only data still missing are Bulgaria 2005 and Switzerland 2005-07 out of 180 combinations country/year. Other doubtful cases (e.g., Belgium 2007, Norway 2006) are left untouched.

In principle, it is not necessary the dataset of reference to be same for all measures, although this reduces the influencing factors to be considered when making comparisons between indicators. Further, similar comparisons can be made using discriminatory variables other than country of birth, such as citizenship, or country of birth *and* citizenship.

[13] It is important to keep in mind that the statistical tests under those assumptions have a pure explorative nature, and incorporating the design effect and/or correcting for multiple comparisons may take to different conclusions.

Figure 1: difference of the estimated p_1 between LFS and SILC by country (sorted by average difference).

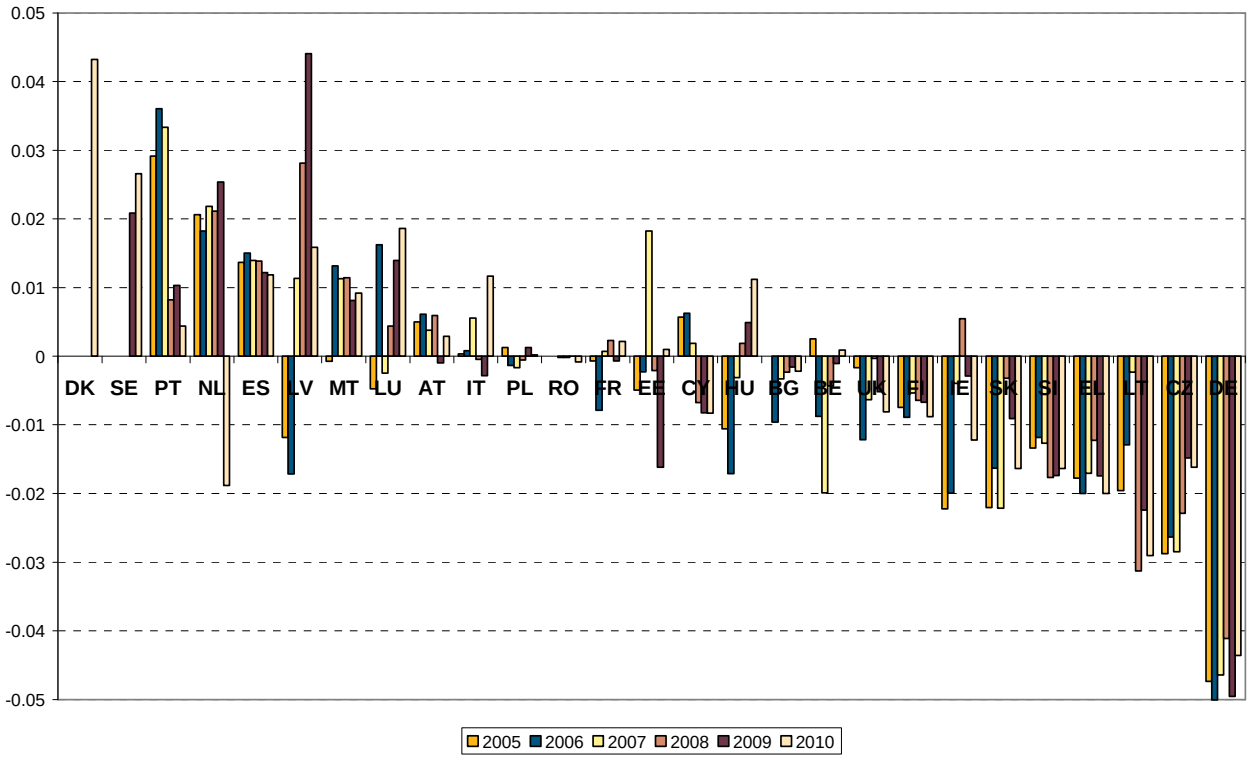


Figure 2: difference of the estimated p_2 between LFS and SILC by country (sorted by average difference).

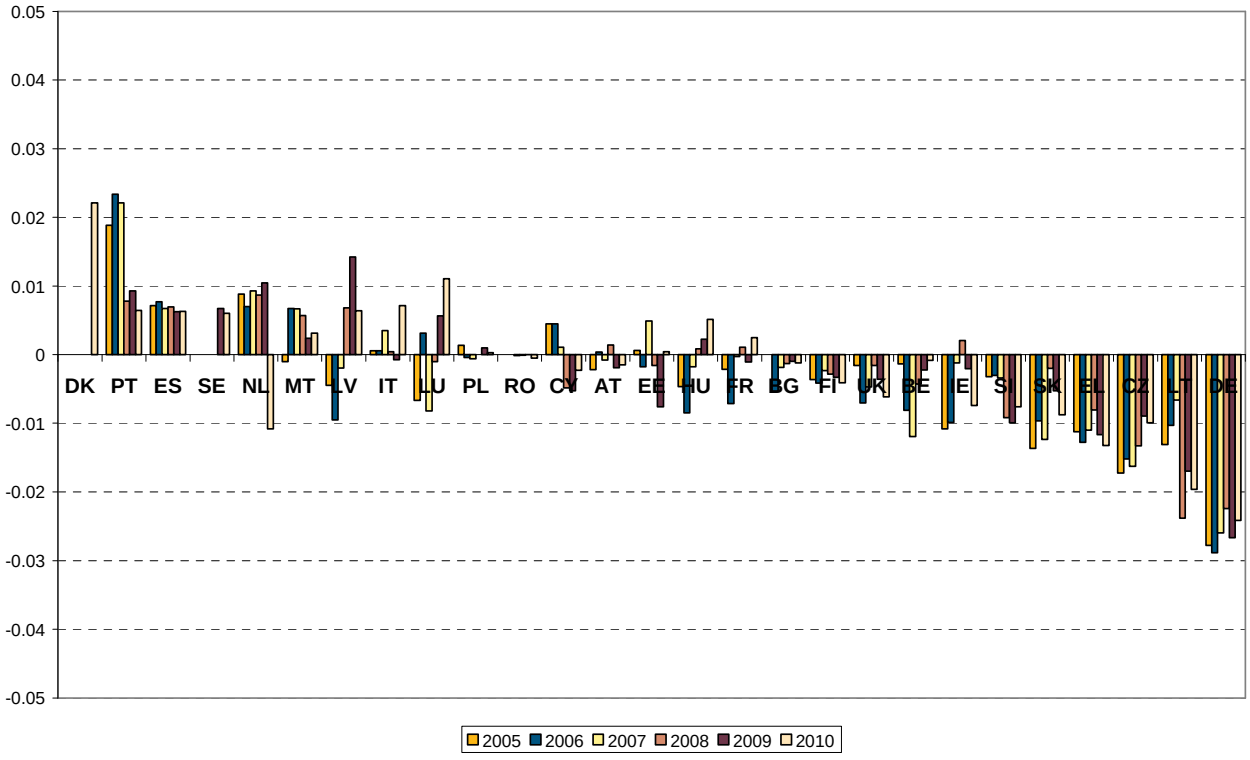


Figure 3: difference of the estimated s_1 between LFS and SILC by country (sorted by average difference).

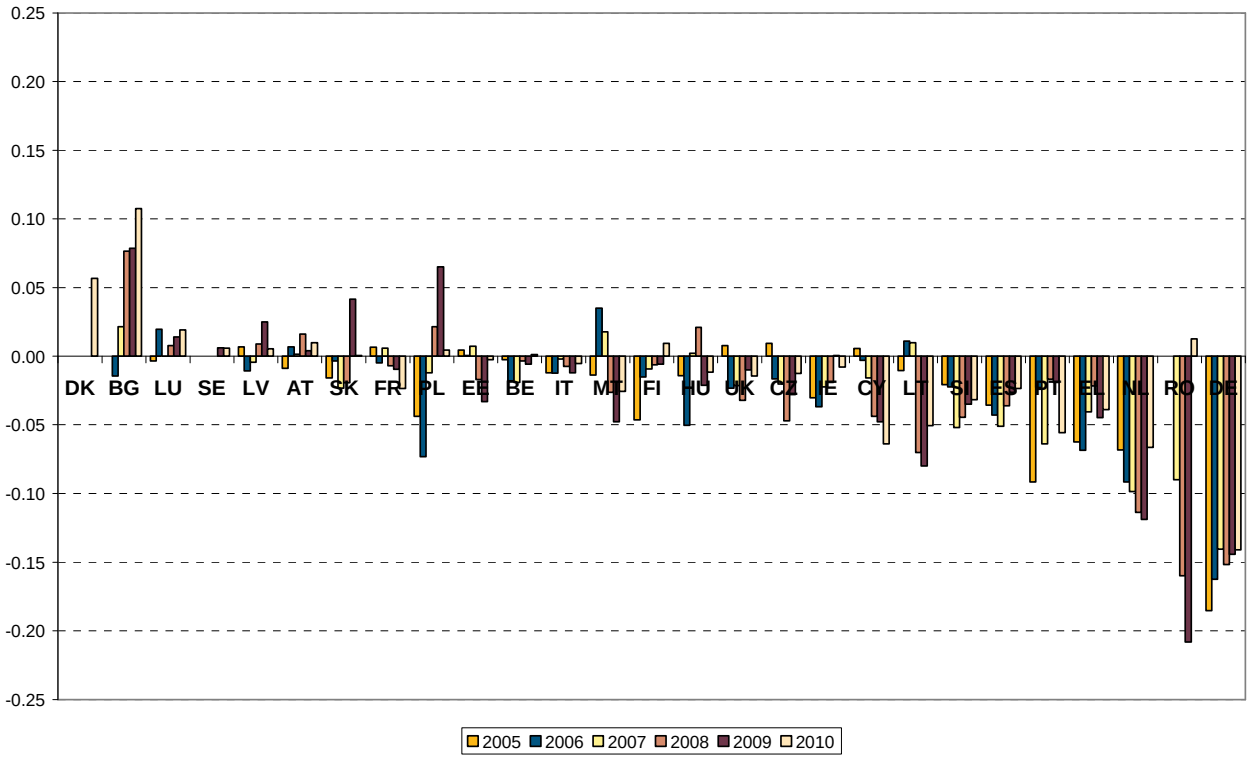
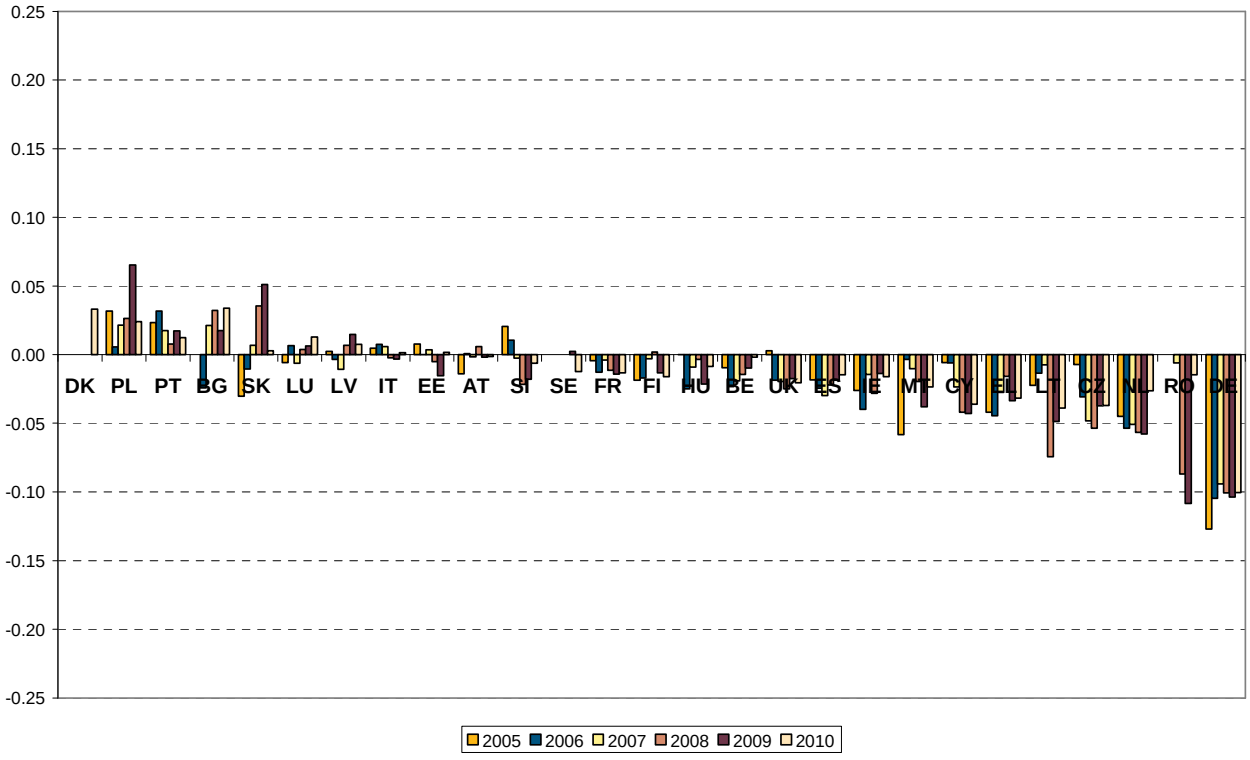


Figure 4: difference of the estimated s_2 between LFS and SILC by country (sorted by average difference).



Results

The Figure 5 shows the average proportions of intermarried couples (or persons) on the total number of couples (or persons) in Europe over two 3-year periods. All countries but Bulgaria, Estonia, Lithuania, the Netherlands and Poland record an increase over time. In Switzerland and Latvia about one couple every five is mixed; other two countries have more than 15% of mixed couples: these four countries are characterized by relatively small population sizes. In the further six countries where the percentage of intermarriages is higher than 10%, there are France and Germany, among the most populous European countries. In eight countries, mostly located in Eastern Europe, such percentage is still below 5%. All in all, one out of twelve couples in Europe is mixed.

When the proportion is related to the total population in marriageable age, the percentage of intermarried persons is about 10% in Switzerland and Luxembourg (see Figure 6), against an overall average of about 4%. In comparison to the previous measure, there are a number of swaps between sorted countries, due to the different countries proportions of married persons, which are incorporated in the indicator p_2. As that proportion of married is slightly above 50% on average across countries, by and large the value of the measure p_2 is approximately half of the measure p_1.

By controlling for the sex and group composition of the population, it turns out that – on average across subpopulations - one out of four married persons is intermarried. In fact, the picture which pops up from the countries comparison is definitely different from what obtained from the previous measures. The Figure 7 shows the ranking of the countries using the indicator s_1: among other things, Bulgaria, Poland and Slovakia which were previously classified at the bottom of the list of countries sorted by proportion of intermarried appear there as the European countries where intermarriage is more widespread; Switzerland is now in the middle of the ranking; Latvia looses a few positions but Luxembourg is towards the bottom of the list (right side of the Figure 7), where the Mediterranean countries (Italy, Spain and Greece) are now located.

To understand how such a change of ranking may occur, let consider the cases of Slovakia, Spain and Luxembourg, as reported in the Table 2. The overall proportion of intermarried persons on the total number of married persons is rather low in Slovakia, where it takes values about 1-2%. However, when the number married persons resident in the country who are foreign-born is taken into account, then it emerges that the share of intermarried is actually much higher, up to values close to 90%, probably due to the (estimated) limited presence of immigrants in the country. Averaging across subgroups allows then to incorporate this information in a synthetic measure of intermarriages.

Spain is a country which has recorded important immigration flows in the latest years and the presence of migrants has risen up to 14.2% at the end of 2010. There the proportions of intermarried persons among married foreign-born persons are again higher than for the native-born ones (and with important differential by sex), but less than one third of the proportions in Slovakia. The measure s_1 takes this diversity into account. Finally, the case of Luxembourg shows the behaviour of the average proportion s_1 in countries where the share of foreign-born persons is more important (32.5% at the end of 2010). Here the indicator s_1 takes values not much different from the overall proportion p_1, again with interesting differential by sex.

Table 2: percentage of intermarried persons in selected countries by population of reference and year

Slovakia	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Married	**2.1%**	**1.7%**	**1.3%**	**2.0%**	**2.0%**	**1.5%**
Native-born married men	1.1%	0.8%	0.6%	0.9%	1.1%	0.8%
Foreign-born married men	84.5%	82.7%	77.8%	84.4%	89.9%	87.9%
Native-born married women	1.1%	0.8%	0.6%	1.1%	0.9%	0.7%
Foreign-born married women	84.4%	83.4%	78.5%	84.1%	91.8%	88.5%
Average sub-groups	**42.8%**	**42.0%**	**39.4%**	**42.6%**	**45.9%**	**44.5%**
Spain	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Married	**4.4%**	**4.6%**	**5.3%**	**5.5%**	**5.2%**	**5.6%**
Native-born married men	2.9%	3.2%	3.6%	3.7%	3.4%	3.7%
Foreign-born married men	21.0%	18.5%	19.9%	19.1%	17.5%	18.1%
Native-born married women	1.9%	2.0%	2.4%	2.6%	2.5%	2.7%
Foreign-born married women	29.1%	26.4%	27.4%	24.8%	22.5%	23.2%
Average sub-groups	**13.7%**	**12.5%**	**13.3%**	**12.5%**	**11.5%**	**11.9%**
Luxembourg	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Married	**16.0%**	**17.3%**	**17.2%**	**18.4%**	**18.2%**	**18.1%**
Native-born married men	14.6%	16.4%	16.0%	15.8%	18.8%	19.3%
Foreign-born married men	17.6%	18.0%	17.5%	20.5%	18.1%	16.1%
Native-born married women	11.6%	11.4%	11.9%	12.8%	13.8%	14.4%
Foreign-born married women	22.4%	26.2%	25.2%	26.4%	22.5%	22.4%
Average sub-groups	**16.5%**	**18.0%**	**17.7%**	**18.9%**	**18.3%**	**18.0%**

Similar considerations apply for the measure s_2, the average proportions on the total population subgroups, with the notice that the proportion of married persons may be different across subgroups. Results for this last indicator are shown in the Figure 8. It may be noted that in these latter two measures there is not the general tendency to increase over time as observed in the two first proportions. Detailed values of the various measures are given in the tables in appendix.

Figure 5: average p_1 by country (sorted by descending order over the period 2008-10)

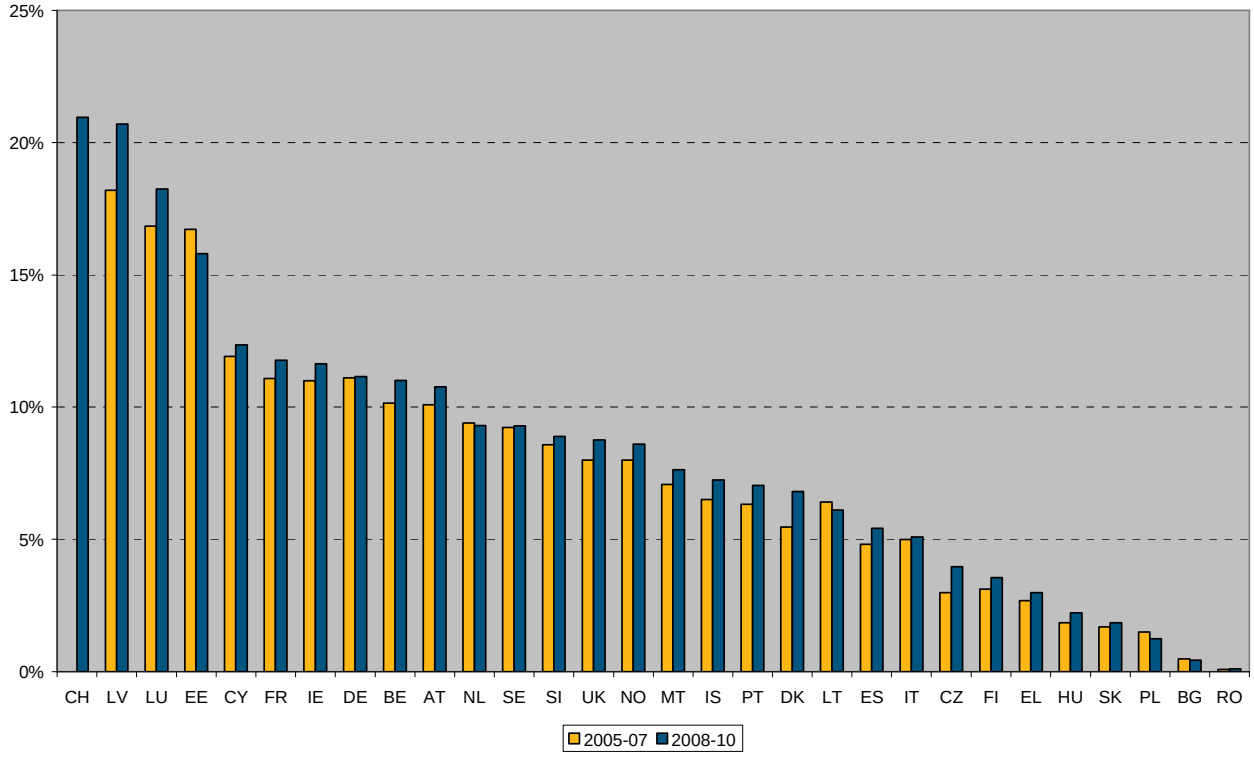


Figure 6: average p_2 by country (sorted by descending order over the period 2008-10)

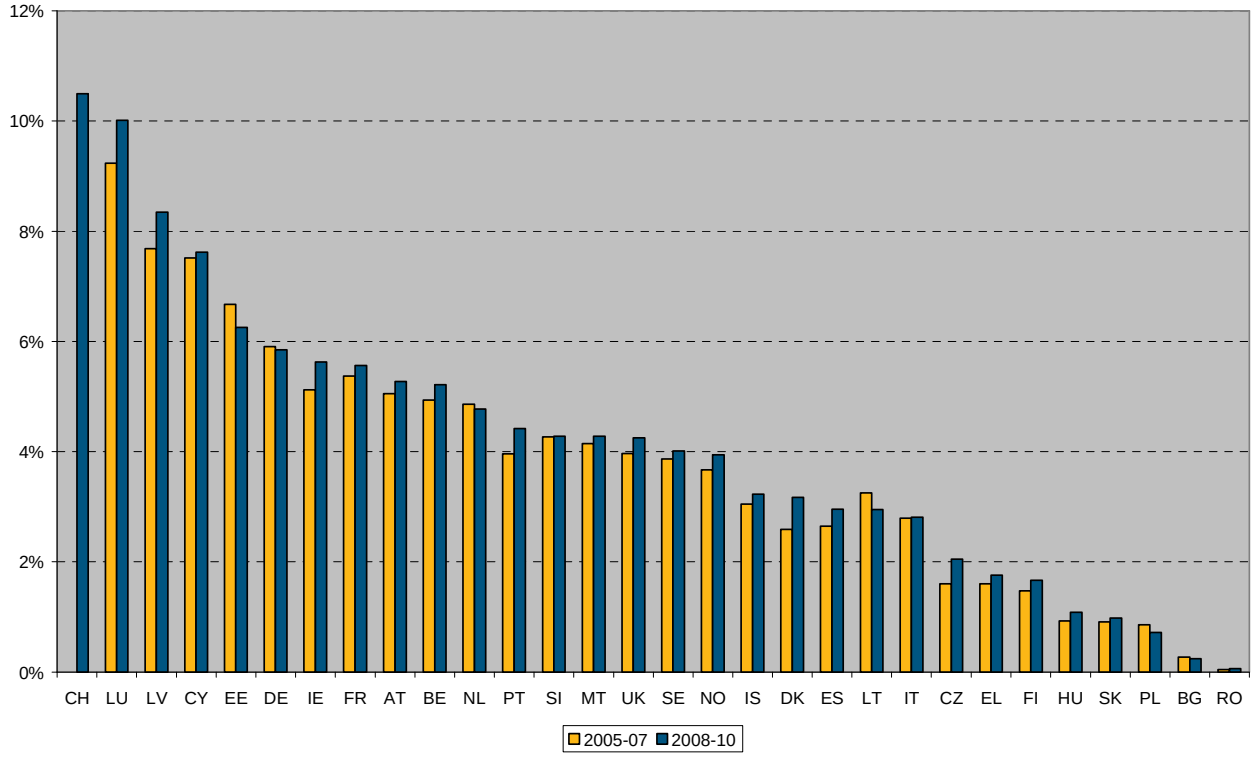


Figure 7: average s_1 by country (sorted by descending order over the period 2008-10)

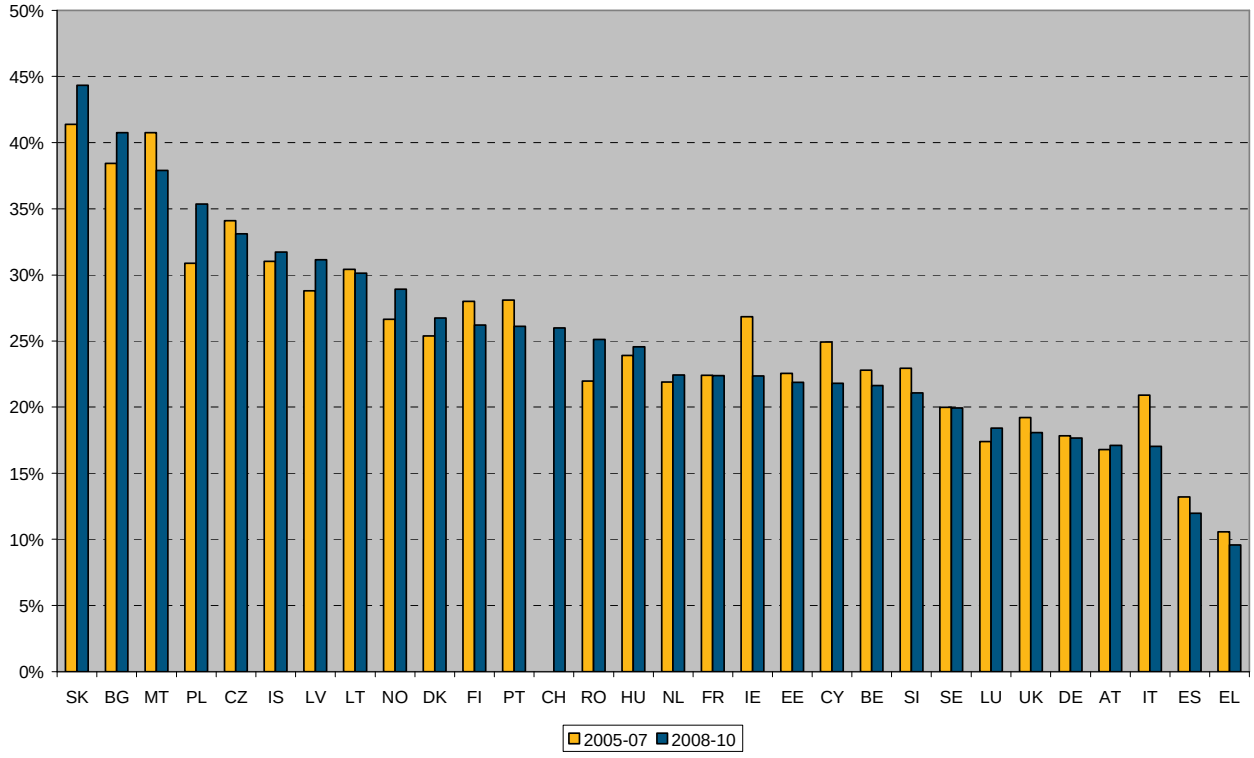
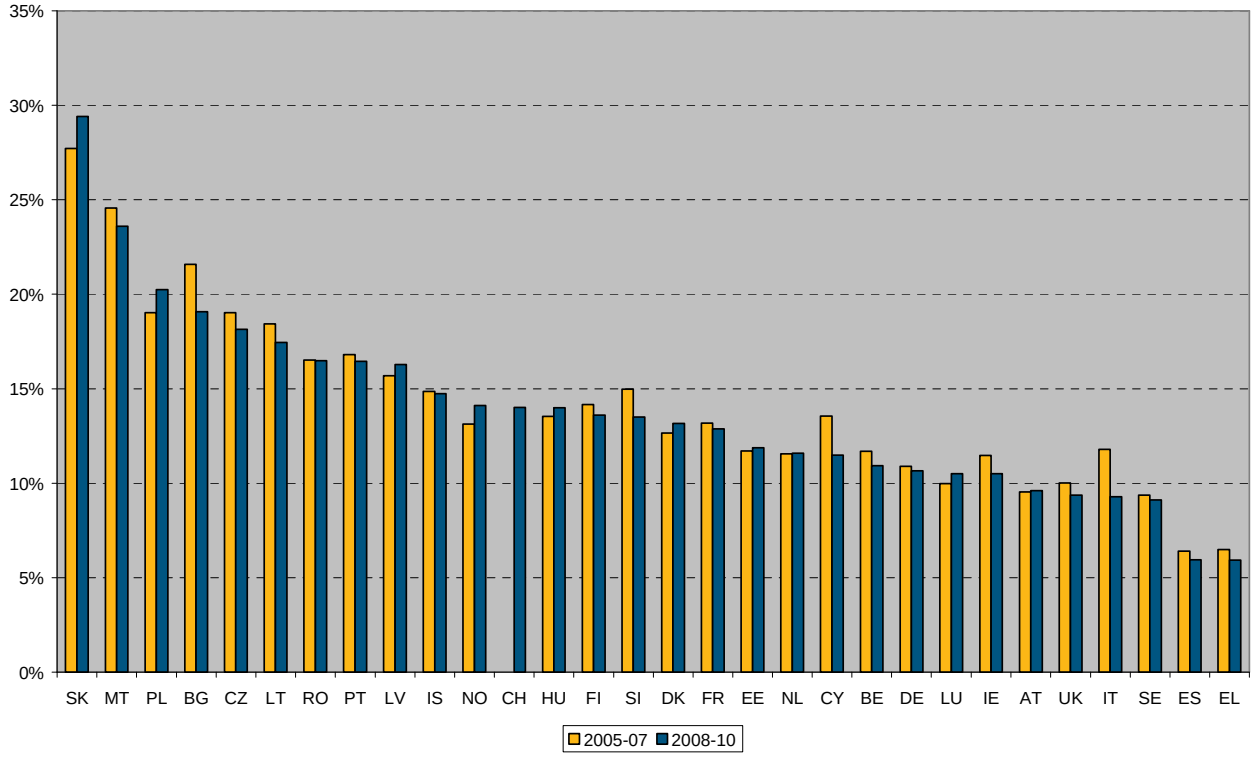
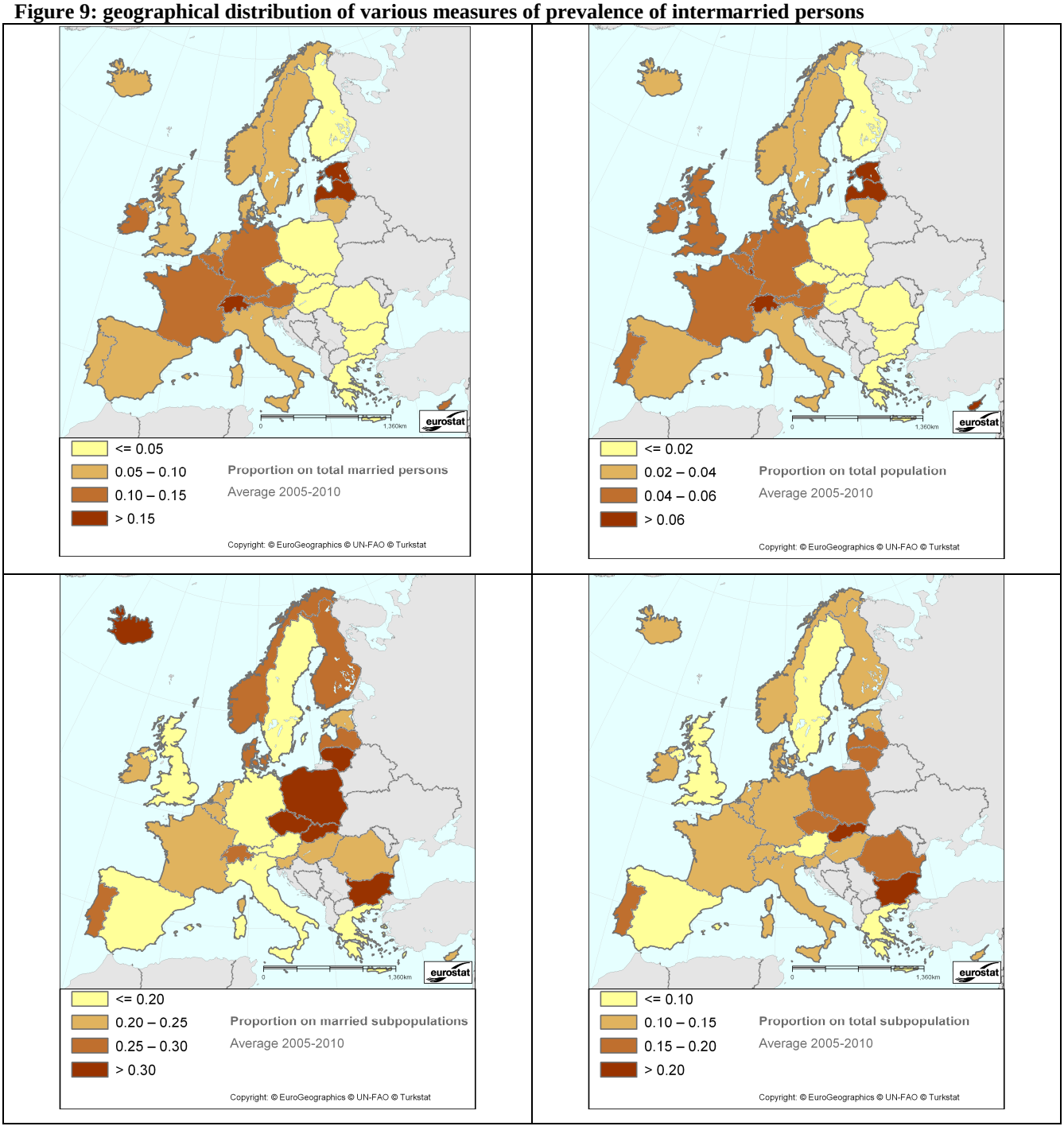


Figure 8: average s_2 by country (sorted by descending order over the period 2008-10)



As for the geographical distribution, according to the first two measures, p_1 and p_2, there is a clear East-West divide, the Eastern European countries being those with lower proportions, either on the total number of married persons or on the total population. Baltic countries are an exception, probably due to the important share of foreign-born persons there present. However, when the structure of the population in terms of sex and group (native-born/foreign-born) is taken into account, the geographical pattern changes remarkably, the Eastern Europe countries recording the higher prevalence of intermarriage.

Figure 9: geographical distribution of various measures of prevalence of intermarried persons



Comparisons with other sources

Very important sources of information on the stock of intermarried persons (by country of birth or other permanent characteristics) are the population censuses. However, they are usually carried out every ten years, the data processing is complex and often requires lengthy periods, and the information on intermarriages is generally not part of standard tabulations. The access to census microdata offered by the IPUMS database[14] is a concrete option, as proved in Esteve *et al.* (2011). It must be noted that those data are samples, though their size is relevant[15], and thus their results are also subject to random variability. Unfortunately, there are no data available from LFS and/or SILC which overlap with the last census data, and the only possibility of direct comparison will then be when the new census 2011 data will be available.

In a preliminary review of the literature of intermarriages in European countries, it was not possible to find data comparable to those here elaborated for the period 2005-2010. Indirectly, this shows that these two surveys may be interesting sources.

Annual migration statistics for the EU and EFTA countries, now harmonized and collected according to an EU regulation, are improving in quality and availability. Although there is no information about the marital status of the immigrants, they may be used in the future to improve the estimate of the population of reference in some of the measures presented above.

Further developments

Consensual union vs. legal marital status

It is sometimes commented that in several European countries the institution of marriage is loosing importance and that any analysis of families based exclusively on the legal marital status would miss important segments of the population, perhaps even giving a biased picture of the society. The survey SILC has a specific variable which allows identifying persons in consensual unions, regardless of their legal marital status. Those data are here compared with the information on marriages as from the same data source. It must be noted that, given the definition of married person here adopted, those in *de facto* relationships[16] are simply the difference between the persons in consensual unions and those married.

[14] Minnesota Population Center. Integrated Public Use Microdata Series – International. Available at: http:/international.ipums.org/international

[15] Among the available European countries, only for the Netherlands the sample in IPUMS is about 190 thousand for the census 2001, against an LFS sample of about 330 thousand in the period 2005-10.

[16] For the sake of precision, these *de facto* relationships include also the registered partnerships which, although not being a formal marriage, do have a legal basis and therefore should be considered as *de jure* relationships. Actually the "Consensual union" variable of SILC allows to distinguish between consensual union with legal basis and consensual union without legal basis, the former being composed by marriages and registered partnership. In this study, these two categories have been merged.

A first quality check[17] has revealed a few problems for data on consensual unions from Germany (2005, 2006), Lithuania (all years), Luxembourg (2007), Romania (all years) and Slovakia (2005, 2006, 2007), which have been thus excluded from the comparison. The total number of remaining available cases is thus 157 on 28 countries.

The Figure 10 shows the average differences between indicators measured on consensual unions and those on marriages. It can be noted that the measures which incorporate the extent of the unions (either consensual or legal) are more sensitive to the change to the concept of consensual union. In fact, recalling the equations [2] and [5], as the marriages are here a subset of the consensual unions, if the estimated proportions p_1 and s_1 are equal for the two concepts, then the proportions p_2 and s_2 would be higher using the consensual union concept. If instead there would be no additional cases of mixed union in the *de facto relationships*, then the measures p_2 and s_2 would remain unchanged. In fact:

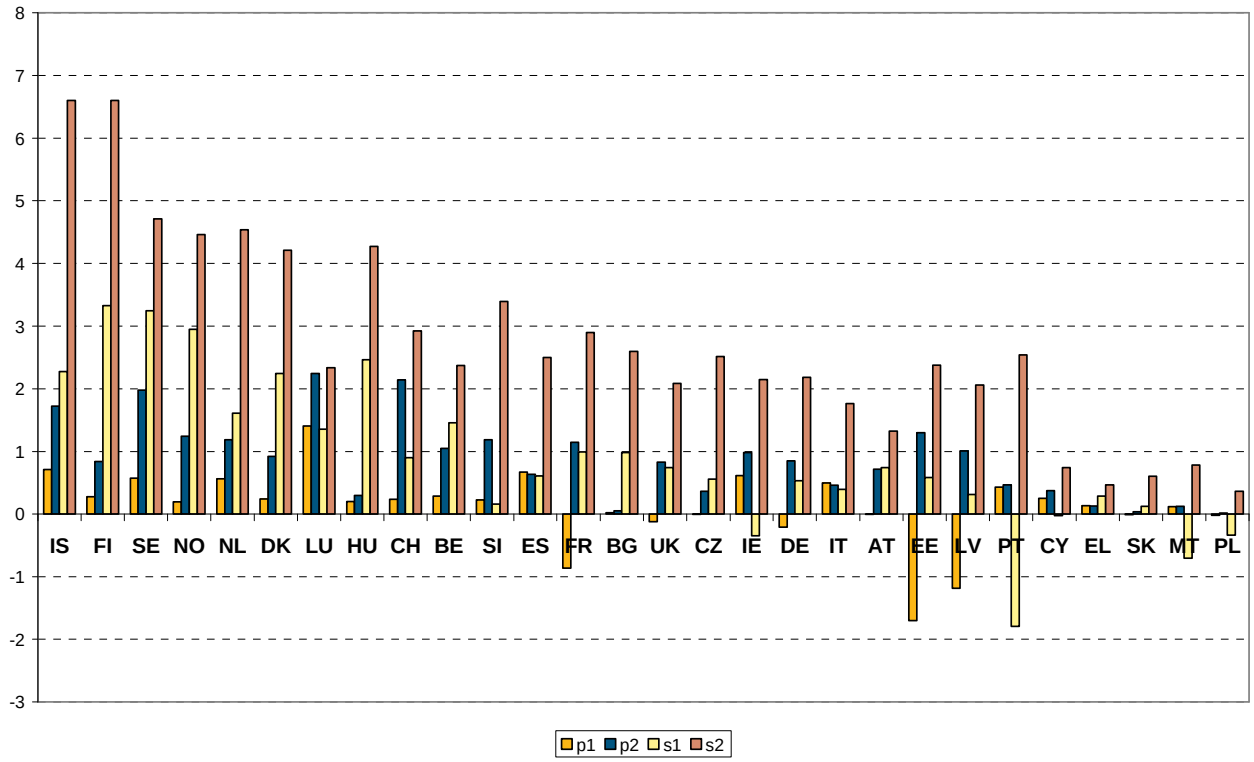
$$p_2^c = \frac{n_m^o + n_f^o}{n_m + n_f} \cdot \frac{n_m + n_f}{n_{16+}} \qquad [10]$$

where n_f is the number of individuals in *de facto* relationship and n_f^o the number of them in mixed union. If $n_f^o = 0$, then $p_2^c = p_2$; the same applies, *mutatis mutandis*, to s_2. Therefore, in theory, moving from the legal marital status to the consensual union concept can not cause a decrease of the measures p_2 and s_2.

On average over the period 2005-10, in the majority of the countries the proportion of persons in mixed union is higher when including the *de facto* relationships; however, that increase is always – apart Luxembourg – less than 1 percentage point, and several big countries (France, Germany and the United Kingdom) have actually a negative sign. When the sex and group structure of the population is taken into account like in the measure s_1, those differentials become wider and may reach also the 3 p.p. When the extent of diffusion of the unions in the population are included in the picture (measures p_2 and s_2), the differences between consensual unions and marriages become relevant especially in the Northern Europe countries, where *de facto* relationships are indeed more common.

[17] For instance, the number of persons in consensual unions was not higher than those married, or the values were very low, much different from those in other years for the same country. Norway 2006 has also an abnormal value for the number of couples (either in marriage or in consensual union), but their proportions to the total population are consistent with the data series and therefore has not been excluded.

Figure 10: average difference 2005-10 (in p.p.) of various measures of mixed unions between consensual unions and marriages (countries sorted by average across measures)



Focussing on immigrants by generation

The crossing of country of birth and citizenship may open further possibilities in the intermarriage studies, identifying specific subpopulations of interest as shown in the Table 3. For instance, if the purpose is to monitor the integration of first generation immigrants, then the selection of the category "*Foreigner Foreign-born*" could better identify the group of interest, removing all cases where marriages between persons belonging to different groups are actually between persons sharing certain characteristics[18]. As citizenship is subject to change over time, it may be useful to consider as well the years of residence in the host country.

Table 3: example of classification of the population by country of birth (rows) and citizenship (columns)

	National	*Foreigner*
Native-born	Natives	Foreign origins (2^{nd} generation immigrants)
Foreign-born	Native origins (descendants of natives)	Immigrants (1^{st} generation immigrants)

[18] Both surveys LFS and SILC have as well an identifier for the father and the mother, which in principle could be used to define the national/foreign background of a person. However, to be interviewed, those father and mother should be resident in the same household, which obviously limits the representativeness of any variable targeting the background.

Adopting a quite restrictive approach, intermarriages could be defined as events occurring between a native and an immigrant (of first generation). Likewise, intra-marriages could be defined as events occurring between natives or between immigrants. However, these strict definitions would leave out all those cases in which there are shared origins. In the Table 4 it is shown the classification in inter- and intra-marriages deriving from the classification of the population in the Table 3. The cells in gray are cases whose classification is not straightforward. To be all precise, also the marriages between foreign-born foreigner persons could be not classified as intra-marriages, because their country of birth and/or citizenship may actually be different, in common being only the fact that they are not native-born and not national. The choice will depend on the hypotheses and the focus of the study.

Table 4: typology of marriages/unions using country of birth and citizenship

	country of birth	*Native-born*	*Native-born*	*Foreign-born*	*Foreign-born*
country of birth	**citizenship**	*National*	*Foreigner*	*National*	*Foreigner*
Native-born	*National*	Intra			Inter
Native-born	*Foreigner*				
Foreign-born	*National*				
Foreign-born	*Foreigner*	Inter			Intra

To show an example on how the cross-classification may turn out to be relevant, I report the application of a single measure (the proportion p_1) on a single country (Austria), considering as intermarriage strictly those between a native-born national person with a foreign-born foreigner person. Results are in the Table 5.

Table 5: comparison between the proportion of intermarriages on the total married couples for the total immigrants and the immigrants of first generation in Austria

	2005	**2006**	**2007**	**2008**	**2009**	**2010**
Total immigrants	10.1%	10.0%	10.1%	10.6%	10.4%	11.3%
Immigrants of first generation	2.0%	1.9%	2.1%	2.2%	2.2%	2.5%

A measure for annual events

While the estimation of the four measures above presented is of minor interest for intermarriage analyses based on the citizenship (due to its non-permanent nature), the same data sources can

provide an important input for the computation of indicators linked to the events. Schoen (1986) has proposed an index *Z* which takes into account the age, sex, marital status and group composition of the population. As the data requirements for such measure are quite demanding, he suggested a reduced version that makes use only of the prime marriage ages instead than of the full range of ages. In Lanzieri (2011) it is proposed a further approximated version which requires only the disaggregation of the events and of the population at risk by sex and group, and this latter measure is applied to the European countries using the breakdown of the marriage events by citizenship (Lanzieri, 2012). The two surveys LFS and SILC allow adding precision to the estimation of the persons at risk of marriage, by identifying those unmarried. If the events are classified as in the Table 6:

Table 6: number of events by citizenship of the spouses

Marriages in selected year(s)		Citizenship of the bride/spouse 2		
		National	Foreigner	Total
Citizenship of the groom/ spouse 1	National	n_{NN}	n_{NF}	n_{N+}
	Foreigner	n_{FN}	n_{FF}	n_{F+}
	Total	n_{+N}	n_{+F}	n_{++}

then the approximated index *Z** is:

$$Z^* = \frac{\frac{n_{NF}}{P_N^{m,u}} + \frac{n_{FN}}{P_N^{w,u}} + \frac{n_{FN}}{P_F^{m,u}} + \frac{n_{NF}}{P_F^{w,u}}}{\frac{n_{N+}}{P_N^{m,u}} + \frac{n_{+N}}{P_N^{w,u}} + \frac{n_{F+}}{P_F^{m,u}} + \frac{n_{+F}}{P_F^{w,u}}} \qquad [11]$$

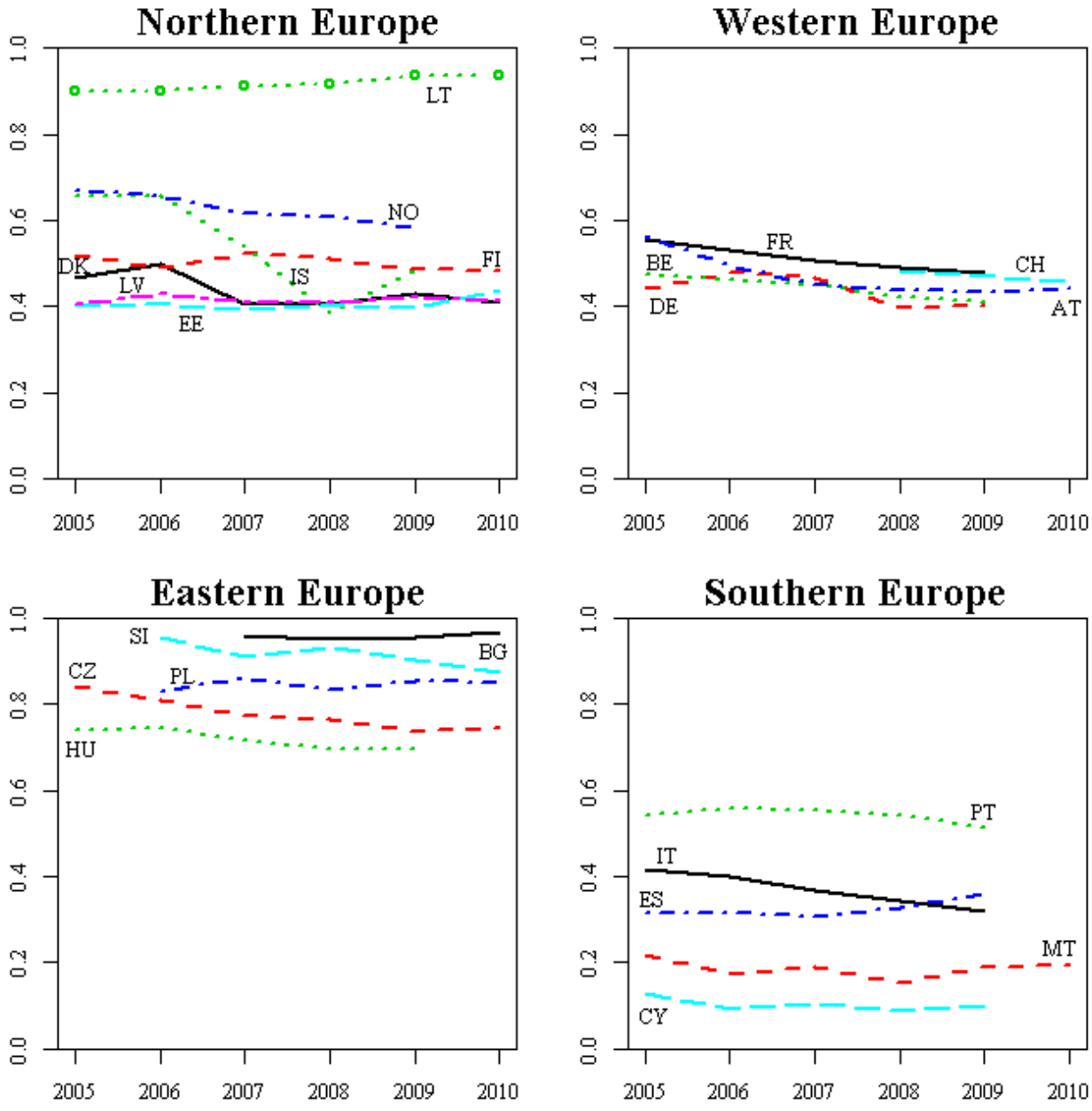
where at the denominators there are the male and female unmarried population in the two groups (nationals and foreigners). Using the data on events from Lanzieri (2012) and the estimation of the population at risk from LFS and SILC, the index *Z** has been calculated for the years 2005-2010 for 24 European countries. The computation of such an index is useful also to spot quality data problems, especially when calculating the ratios referring to the foreign population at risk. This has been the case for Bulgaria (2005, 2006), Latvia[19] (2005), Romania (2005, 2006, 2008, 2009, 2010) and Slovakia (all years), where the estimates of the unmarried foreign persons appear to be too low in comparison to the number of events which should involve them. Those cases have been excluded from the computation of *Z**, thus finally getting available 165 observations on 22 countries[20].

The Figure 11 shows the estimated values of the indicator Z* by group of countries. It can be noted how, once taken into account the population structure (as for sex, citizenship and legal marital status) and the relative size of the groups, the measures of intermarriages are higher in Eastern Europe and Lithuania, where the share of immigrants is relatively small. The Southern Europe

[19] Values for Latvia 2005-07 has been then taken from SILC.
[20] The only valid observation available for Romania has been dropped as well.

countries, which have experienced large immigration flows in recent years, are on lower levels than the countries where immigration is historically present, like in Central and Northern Europe. It is interesting to note that countries which record high proportions of intermarried persons (and events) like Switzerland do not outstand when using the index *Z**. A preliminary comparison with the values in Lanzieri (2012) over the available common data (120 combinations country/year over the period 2005-09) shows that the use of more precise estimations of the population at risk of marriage in the index causes an average increase of the same index of about 5 p.p. For a discussion on the results from the application of a simplified version of the *Z* index on European countries, see Lanzieri (2011, 2012).

Figure 11: trends of the index *Z based on citizenship**



Summary and Discussion

This study starts from the consideration that the analysis only of the occurrence of the events (marriages) is a partial view at the social integration of migrants, of which intermarriages are usually considered an important indicator (although with some criticisms, see Song 2009). Such an analysis does not inform about the diffusion of those cross-boundaries relationships in the society and thus at the (cumulative) extent of merging of the groups as identified by the selected characteristic(s) (country of birth, ethnicity, religion, etc.). Besides various drawbacks reported by several scholars of intermarriages, a possible reason for neglecting the prevalence dimension (i.e. the number of *intermarried* persons) in favour of the incidence (i.e. the number of *intermarrying* persons) may also be the difficulty to find relevant data. In fact, data sources have been so far limited to the censuses or to specific surveys: the former are available only at wide time spans (and with delay), the latter are often limited to a single country and/or exercise, which make it difficult to develop comparative analyses. As migration is nowadays an important item on the political agenda in the EU, the monitoring of their socio-economic integration requires regular and comparable data across European countries. Efforts in this direction have been already undertaken, but the use of the ordinary outcomes from two major sample surveys in the EU, the LFS and the SILC, opens further possibilities. Both surveys allow identifying intermarried persons, either by country of birth or by citizenship, and to link this information to a wealth of additional statistical data, more focussed on labour issues the former survey, on income and social inclusion the latter. Therefore, the choice of one of these two surveys as data source depends as well on the interests of the researcher. If the access to microdata is the preferred option, statistical linking may offer additional insights (whether its technical complexity does not hamper the benefits). Between LFS and SILC there is an evident disparity of sample sizes, but not always this takes to significant differences when the variable of interest is a much aggregated measure, such as general intermarriage indicators.

Privileging the concept of country of birth for intermarriage analyses based on prevalence measures, in this study I have considered two very simple measures of intermarriage, one referring to the overall population of reference and the other discounting the sex and group (native-born/foreign-born) composition of the same population. The former is a simple proportion, the latter an average proportion across subpopulations. The population of reference is either the married persons or the persons in marriageable age. Measures based on the latter concept incorporate the impact of the diffusion of the institution of marriage in the given society. In total, I computed four basic measures of intermarriage, any of them of quite immediate understanding also for non-technical users.

In consideration of the important sampling sizes differentials, I have chosen LFS as main data source. However, whether the information is not available for a specific country and/or year, SILC can be a valuable substitute, so to get a wider geographical and temporal coverage. For the years on

which relevant information is currently available from these data sources (2005-10), only 4 cases out of 180 are missing in the final dataset (Bulgaria 2005 and Switzerland 2005-07). Their regular implementation provides new annual estimates within one year from the year of reference, thus with a timing comparable to other related data sources (e.g., demographic and migration statistics). Further, their execution is ensured and framed by EU regulations, and therefore less influenced by occasional constraints (e.g., funding) or by the voluntary nature as other surveys. These data sources may thus be considered also for future analyses. In particular, both surveys allow access to microdata under specific conditions.

Turning to the results, the mixed couples are a phenomenon of some relevance in Europe: on average across countries, one out of twelve married persons has a foreign-born partner. If this is reported to the total population of marriageable age, on average the share of intermarried person is one out of 25. This is an important engine of change of the societies, although there is a large variability across countries, reflecting also their immigration history. Further, the prevalence in specific population subgroups may be much higher and averaging across countries and across subpopulations gives back a level of intermarriage as high as one quarter (referring to married population). The importance of the application of the measures of intermarriage to relevant subpopulations, such as by sex and country-of-birth group, emerges when comparing the results across countries. The picture obtained by using a weighted or not weighted version of the proportions may be quite different, especially for those countries where the relative sizes of the subpopulations are much different. Are those measures (p_1 and p_2 vs. s_1 and s_2) incompatible between them? The answer would be no, and it is also related to their purposes. The former couple of measures seem of easier communication to policy-makers and in general to the large public, which in the case of an overall look at the composition of our societies may be an advantage. However, when it comes to the finding of the reasons of the cross-countries variability, or to an international comparison unaffected by the population structures, the latter measures probably provide a better service. Although the arithmetic mean is a rather elementary way of synthesising in one single measure the differences between subgroups, it has the advantage of keeping transparent the process of estimation of the value.

A further question may be about the effective necessity of an *annual* measure of intermarriage prevalence. As for the events (mixed marriages) there is evidence of fluctuations over years, it may be interesting as well to see how this impacts on the stock of intermarriages, the same way as it goes for many other characteristics on annual basis (e.g., migrants stocks and migration flows). Further, as shown above, these data are also an input to more complex indexes of intermarriages, as they provide an estimation of the population at risk. However, the final word is up to the experts in the field, which may consider the use of such measures as indicator on time span larger than one

year (e.g., a 3-year average could fit). Annual statistical availability does not translate necessarily in annual production and use.

Further analyses definitely need to be carried out before a final decision could be taken as for what concern an indicator of migrants' social integration based on intermarriages. The current study is a simple preliminary exploration of two potential data sources, hopefully drawing the attention of interested researchers to the analytical possibilities they offer. This analysis also highlights the fact the measures should not be adopted blindly, and that the use of one or another may completely change the picture. Before looking at the explanatory variables, it is fundamental to quantify properly the phenomenon of interest. This study was just an attempt of a little step in that direction.

References

Drachsler, J. (1921): "Intermarriage in New York City: a statistical study of the amalgamation of European peoples". Columbia University.

Esteve-Palos, A., J. Garcia and R. McCaa (2011): "Comparative perspectives on Marriage and International Migration, 1970-2000: findings from IPUMS-International census microdata samples". Paper for the IUSSP Seminar on Global Perspectives on Marriages and International Migration, Seoul, 20-21 October 2011.

Eurostat (2003): "Demographic Statistics: Definitions and Methods of Collection in 31 European Countries". *Eurostat Methodologies and Working Papers*. Publications Office of the European Union, Luxembourg.

Eurostat (2010a): "EU Labour Force Survey database – User Guide". Eurostat document version November 2010.

Eurostat (2010b): "Description of target variables: Cross-sectional and Longitudinal – 2010 operation". Eurostat document version February 2010.

Eurostat (2011): "Indicators of Immigrant Integration - A Pilot Study". *Eurostat Methodologies and Working Papers*. Publications Office of the European Union, Luxembourg.

Eurostat (2012): "Labour force survey in the EU, candidate and EFTA countries – Main characteristics of national surveys, 2010". *Eurostat Methodologies and Working Papers*. Publications Office of the European Union, Luxembourg.

Gaspar S. (2008): "Towards a definition of European intra-marriage as a new social phenomenon". CIES e-Working Paper N.46/2008. Centre for Research and Studies in Sociology, Lisbon, Portugal.

Huddleston, T., J. Niessen, E. Ni Chaoimh and E. White (2011): "Migrant Integration Policy Index". British Council and Migration Policy Group, Brussels. Available at www.mipex.eu

Kalmijn, M. and F. van Tubergen (2006). Ethnic intermarriage in the Netherlands: confirmations and refutations of accepted insights. *European Journal of Population*, 22:371-397.

Lanzieri, G. (2011): "A comparison of recent trends of international marriages and divorces in European countries". Paper for the IUSSP Seminar on Global Perspectives on Marriages and International Migration, Seoul, 20-21 October 2011.

Lanzieri, G. (2012): "Mixed Marriages in Europe, 1990-2010". In D.-S. Kim (ed.): "Cross-Border Marriage: Global Trends and Diversity", p.81-121. Korea Institute for Health and Social Affairs (KIHASA), Seoul (*forthcoming*).

Lucassen, L. and C. Laarman (2009): "Immigration, intermarriage and the changing face of Europe in the post war period". *History of the Family*, 14:52-68.

Meng, X. and R.G. Gregory (2005): "Intermarriage and the Economic Assimilation of Immigrants". *Journal of Labor Economics*, 23(1):135-174.

Price, C. A. and J. Zubrzycki (1962a): "The Use of Inter-Marriage Statistics as an Index of Assimilation". *Population Studies* 16(1):58-69.

Price, C. A. and J. Zubrzycki (1962b): "Immigrant Marriage Patterns in Australia". *Population Studies* 16(1):123-133.

Schoen, R. (1986): "A methodological analysis of intergroup marriage". *Sociological Methodology*, 16:49-78.
Schuh, U. (2006): "Mixed Marriages in the EU". Research Note of the Institute for Advanced Studies, Vienna.
Schwartz, C.R. and R.D. Mare (2005): "Trends in educational assortative marriage from 1940 to 2003". *Demography*, 42(4):621-646.
Song, M. (2009): "Is Intermarriage a Good Indicator of Integration?". *Journal of Ethnic and Migration Studies*, 35(2):331-348.
Timmerman, C. (2008): "Marriage in a 'Culture of Migration'. Emirdag Marrying into Flanders". *European Review*, 16(4):585-594.
Trilla, C.C., A. Esteve and A. Domingo (2008): "Marriage Patterns of the Foreign-Born Population in a New Country of Immigration: The Case of Spain". *International Migration Review*, 42(4):877-902.
Van Ham, M. and T. Tammaru (2011): "Ethnic Minority-Majority Unions in Estonia". *European Journal of Population*, 27:313-335.
Van Tubergen, F. and I. Maas (2007): "Ethnic intermarriage among immigrants in the Netherlands: An analysis of population data". *Social Science Research*, 36:1065-1086.
Wolff, P., F. Montaigne and G. Rojas-González (2010): "Investing in statistics: EU-SILC". In A.B. Atkinson and E. Marlier (eds.): "Income and living conditions in Europe". *Eurostat Statistical Books*. Publications Office of the European Union, Luxembourg.
Wray, H. (2006): "An Ideal Husband? Marriages of Convenience, Moral Gate-keeping and Immigration to the UK". *European Journal of Migration and Law*, 8:303-320.

Table 7: achieved sample sizes of individuals in LFS and SILC by country and year (in thousand)

	LFS						SILC						Ratio LFS/SILC					
Country	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
BE	86.8	90.4	88.4	84.4	83.2	81.6	10.0	11.3	12.3	12.2	11.8	11.8	8.7	8.0	7.2	6.9	7.1	6.9
BG		106.4	104.0	108.0	110.4	106.4		10.3	10.4	10.4	13.2	14.5		10.3	10.0	10.4	8.4	7.3
CZ	196.0	199.6	197.2	190.0	185.2	183.2	8.6	14.9	19.4	22.8	19.8	18.2	22.8	13.4	10.2	8.3	9.4	10.1
DK						83.2	11.9	11.4	11.6	11.5	11.8	11.7						7.1
DE	516.8	538.4	525.2	525.6	529.6	527.2	25.0	26.0	26.3	24.3	23.8	23.5	20.7	20.7	20.0	21.6	22.3	22.4
EE	14.4	16.8	19.2	18.4	16.0	16.4	9.6	13.0	12.0	10.9	11.3	11.2	1.5	1.3	1.6	1.7	1.4	1.5
IE	263.6	248.0	236.4	211.6	198.8	177.6	12.0	11.5	10.9	10.1	9.9	8.8	22.0	21.6	21.7	21.0	20.1	20.2
EL	241.6	231.6	224.4	222.4	226.0	229.6	12.4	12.6	12.3	14.1	15.0	14.8	19.5	18.4	18.2	15.8	15.1	15.5
ES	470.0	487.6	505.2	508.0	516.8	518.4	30.4	28.1	28.7	30.1	30.8	31.0	15.5	17.4	17.6	16.9	16.8	16.7
FR	248.8	245.6	255.6	253.6	302.4	362.0	18.8	19.5	20.4	20.1	20.2	21.1	13.2	12.6	12.5	12.6	15.0	17.2
IT	535.2	517.6	508.8	503.6	492.0	492.0	47.3	46.0	44.6	44.3	43.1	40.4	11.3	11.3	11.4	11.4	11.4	12.2
CY	29.2	28.0	29.2	30.4	30.8	33.2	9.0	8.8	8.5	8.1	7.6	9.1	3.2	3.2	3.4	3.8	4.1	3.6
LV	17.6	15.2	30.0	30.8	32.4	29.6	7.9	9.1	9.3	10.9	12.2	13.0	2.2	1.7	3.2	2.8	2.7	2.3
LT	37.6	27.6	49.6	48.8	52.8	53.2	9.9	10.2	10.9	10.5	11.2	11.6	3.8	2.7	4.6	4.6	4.7	4.6
LU	17.2	16.0	16.0	10.8	14.4	14.4	7.5	7.8	7.9	7.6	8.6	10.2	2.3	2.1	2.0	1.4	1.7	1.4
HU	244.8	245.6	240.4	227.6	226.0	221.6	14.8	16.5	18.5	18.7	21.0	20.7	16.5	14.9	13.0	12.2	10.8	10.7
MT	16.8	17.2	22.4	21.6	20.4	20.8	8.2	8.3	8.3	7.9	8.5	8.7	2.0	2.1	2.7	2.7	2.4	2.4
NL	355.6	360.0	333.6	330.4	315.2	298.8	17.9	17.4	19.6	19.5	18.3	19.1	19.9	20.7	17.0	16.9	17.2	15.6
AT	153.6	151.2	151.6	147.2	142.0	139.2	10.4	12.0	13.4	11.0	11.1	11.5	14.8	12.6	11.3	13.4	12.8	12.1
PL	173.6	167.2	160.0	160.4	164.4	321.2	37.7	36.6	34.9	33.8	31.7	30.8	4.6	4.6	4.6	4.7	5.2	10.4
PT	146.0	137.6	132.4	129.2	126.0	123.2	10.7	10.1	9.9	10.1	11.1	11.4	13.6	13.6	13.4	12.8	11.4	10.8
RO	215.6	206.0	199.6	192.4	192.0	196.4			17.0	16.5	16.3	16.2			11.7	11.7	11.8	12.1
SI	58.4	57.2	51.6	53.6	66.0	52.0	23.9	27.1	24.7	25.0	25.4	25.2	2.4	2.1	2.1	2.1	2.6	2.1
SK	90.8	89.2	87.6	87.6	83.6	84.0	12.9	12.6	12.6	14.1	13.8	14.1	7.0	7.1	7.0	6.2	6.1	6.0
FI	143.6	139.2	137.6	138.8	136.4	133.6	23.0	22.1	21.8	21.1	20.1	21.7	6.2	6.3	6.3	6.6	6.8	6.2
SE					200.0	260.0	12.2	13.6	14.2	14.9	14.9	14.3					13.4	18.2
UK	356.8	348.0	347.6	342.4	329.6	318.4	20.1	18.6	17.5	16.8	15.6	15.1	17.8	18.7	19.9	20.4	21.1	21.1
IS							6.7	6.5	6.6	6.6	6.5	6.8						
NO							11.9	11.4	11.7	10.9	10.6	10.2						
CH										13.2	14.3	14.7						

Annex

	Values of p_1						Values of p_2					
Country	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
BE	0.107	0.107	0.090	0.106	0.111	0.113	0.052	0.052	0.044	0.051	0.053	0.053
BG		0.005	0.005	0.004	0.005	0.004		0.003	0.003	0.002	0.003	0.002
CZ	0.032	0.030	0.027	0.037	0.041	0.041	0.018	0.016	0.014	0.019	0.021	0.021
DK	0.058	0.054	0.052	0.062	0.072	0.070	0.027	0.026	0.025	0.029	0.033	0.033
DE	0.115	0.111	0.107	0.114	0.110	0.112	0.062	0.058	0.057	0.060	0.058	0.058
EE	0.155	0.168	0.179	0.158	0.151	0.165	0.064	0.066	0.070	0.063	0.059	0.065
IE	0.107	0.110	0.114	0.115	0.118	0.116	0.050	0.051	0.052	0.054	0.058	0.057
EL	0.029	0.025	0.027	0.031	0.030	0.028	0.018	0.014	0.016	0.018	0.018	0.016
ES	0.044	0.046	0.053	0.055	0.052	0.056	0.025	0.026	0.029	0.030	0.028	0.030
FR	0.114	0.107	0.111	0.119	0.115	0.119	0.056	0.052	0.053	0.056	0.055	0.056
IT	0.049	0.050	0.050	0.050	0.050	0.053	0.028	0.028	0.028	0.028	0.028	0.029
CY	0.110	0.121	0.127	0.127	0.123	0.121	0.071	0.076	0.079	0.078	0.076	0.074
LV	0.174	0.176	0.196	0.207	0.222	0.191	0.077	0.076	0.077	0.084	0.088	0.079
LT	0.057	0.065	0.071	0.058	0.061	0.064	0.029	0.033	0.035	0.028	0.030	0.031
LU	0.160	0.173	0.172	0.184	0.182	0.181	0.089	0.095	0.093	0.103	0.100	0.098
HU	0.021	0.018	0.017	0.021	0.023	0.023	0.010	0.009	0.009	0.010	0.011	0.011
MT	0.068	0.075	0.069	0.071	0.078	0.080	0.040	0.044	0.041	0.041	0.043	0.044
NL	0.099	0.091	0.092	0.091	0.094	0.094	0.052	0.047	0.047	0.046	0.047	0.050
AT	0.101	0.100	0.101	0.106	0.104	0.113	0.051	0.051	0.050	0.052	0.051	0.055
PL	0.018	0.014	0.013	0.014	0.013	0.011	0.011	0.008	0.007	0.008	0.008	0.006
PT	0.060	0.063	0.066	0.071	0.071	0.070	0.038	0.040	0.042	0.044	0.044	0.044
RO	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.001	0.001	0.000
SI	0.090	0.086	0.082	0.085	0.090	0.092	0.045	0.042	0.041	0.041	0.043	0.044
SK	0.021	0.017	0.013	0.020	0.020	0.015	0.012	0.009	0.007	0.011	0.011	0.008
FI	0.027	0.033	0.034	0.034	0.036	0.037	0.013	0.016	0.016	0.016	0.017	0.017
SE	0.099	0.090	0.088	0.093	0.093	0.093	0.042	0.038	0.037	0.041	0.039	0.041
UK	0.079	0.079	0.082	0.087	0.086	0.089	0.040	0.039	0.040	0.043	0.042	0.042
IS	0.059	0.061	0.075	0.074	0.071	0.073	0.027	0.029	0.035	0.033	0.032	0.033
NO	0.078	0.079	0.083	0.083	0.087	0.089	0.035	0.036	0.038	0.038	0.040	0.040
CH				0.208	0.211	0.209				0.104	0.105	0.106

	Values of s_1						Values of s_2					
Country	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
BE	0.236	0.232	0.216	0.222	0.215	0.212	0.121	0.117	0.112	0.111	0.109	0.107
BG		0.378	0.390	0.392	0.407	0.424		0.214	0.218	0.186	0.198	0.189
CZ	0.352	0.340	0.331	0.321	0.332	0.340	0.207	0.189	0.174	0.181	0.179	0.184
DK	0.265	0.257	0.240	0.265	0.269	0.268	0.141	0.123	0.116	0.124	0.137	0.134
DE	0.173	0.184	0.178	0.174	0.177	0.178	0.105	0.113	0.109	0.105	0.106	0.108
EE	0.215	0.228	0.233	0.212	0.211	0.233	0.114	0.116	0.122	0.115	0.113	0.127
IE	0.292	0.267	0.246	0.229	0.225	0.216	0.136	0.107	0.101	0.098	0.107	0.109
EL	0.113	0.101	0.103	0.109	0.094	0.084	0.070	0.061	0.063	0.068	0.057	0.053
ES	0.137	0.125	0.133	0.125	0.115	0.119	0.067	0.062	0.063	0.062	0.056	0.060
FR	0.229	0.221	0.222	0.231	0.222	0.219	0.136	0.131	0.129	0.133	0.127	0.126
IT	0.218	0.210	0.200	0.183	0.166	0.162	0.123	0.118	0.113	0.099	0.091	0.088
CY	0.246	0.252	0.250	0.240	0.219	0.195	0.130	0.140	0.137	0.125	0.117	0.103
LV	0.279	0.290	0.295	0.312	0.323	0.300	0.154	0.165	0.151	0.165	0.167	0.156
LT	0.288	0.307	0.317	0.274	0.307	0.323	0.179	0.182	0.192	0.157	0.176	0.190
LU	0.165	0.180	0.177	0.189	0.183	0.180	0.096	0.103	0.100	0.111	0.103	0.101
HU	0.251	0.234	0.232	0.241	0.246	0.249	0.143	0.133	0.130	0.131	0.137	0.151
MT	0.392	0.418	0.413	0.383	0.370	0.384	0.234	0.256	0.247	0.234	0.234	0.239
NL	0.224	0.218	0.215	0.219	0.219	0.235	0.121	0.113	0.112	0.110	0.112	0.125
AT	0.173	0.166	0.164	0.170	0.167	0.176	0.099	0.095	0.092	0.095	0.094	0.099
PL	0.321	0.296	0.308	0.350	0.369	0.341	0.203	0.181	0.186	0.199	0.218	0.190
PT	0.278	0.286	0.278	0.272	0.260	0.251	0.165	0.171	0.169	0.167	0.166	0.161
RO	0.224	0.224	0.211	0.265	0.292	0.197	0.186	0.165	0.144	0.181	0.191	0.124
SI	0.237	0.240	0.211	0.212	0.215	0.206	0.160	0.153	0.137	0.135	0.137	0.133
SK	0.428	0.420	0.394	0.426	0.459	0.445	0.284	0.273	0.274	0.291	0.307	0.284
FI	0.255	0.304	0.282	0.257	0.264	0.265	0.127	0.153	0.145	0.140	0.137	0.131
SE	0.223	0.188	0.188	0.201	0.195	0.202	0.104	0.090	0.087	0.096	0.092	0.086
UK	0.202	0.189	0.186	0.182	0.181	0.180	0.106	0.099	0.096	0.095	0.094	0.092
IS	0.320	0.306	0.305	0.319	0.316	0.317	0.151	0.141	0.153	0.153	0.146	0.144
NO	0.265	0.250	0.285	0.301	0.270	0.297	0.131	0.125	0.138	0.147	0.130	0.146
CH				0.261	0.266	0.253				0.141	0.140	0.139