The Effect of Fertility on the Evolution of the Post-Reproductive Sex Ratio

Giambattista Salinari, Cristina Giuliani, Marco Breschi, Virginia Zarulli, Claudio Franceschi, Gustavo De Santis

Abstract

At post-reproductive ages (50+), there are typically fewer men than women, whose mortality is lower. This is generally imputed to genes (sex hormones, sex chromosomes and the mitochondrial DNA) and smoking. In this paper, however, we submit and test an alternative hypothesis: high fertility was one of the causes of high female mortality in the post reproductive age span (50 years and over) of past generations, and extra improvements in female survival, as compared to men, could be observed only after, and because of, the fertility transition. To substantiate our claim, we analyze longitudinal (aggregate) data of the cohorts born between the end of the 19th and mid-20th century in 16 Italian regions and 16 European countries: the effects of differential cohort fertility on female survival are evident and become stronger with age. This is in line with some theories of evolutionary medicine, which suggest that increased somatic investment in the reproductive function among women comes at the cost of worse health and lower survival in the post-reproductive period.

1. Introduction

In 1999, the female-to-male ratio (FMR) was a mere 2.5 among Sardinian centenarians, well below the lower bound of what was then considered the "normal" 4-to-7 range (Deiana et al 1999; Koenig 2001). Shortly later, in 2001, the FMR was found to be low also among the centenarians in a few southern Italian regions (Sicily, Calabria, Molise), but high, between 7 and 8, in the north of the country (Friuli, Liguria, Veneto, and Trentino; Passarino et al 2002). As the female post-reproductive survival advantage is a relatively recent phenomenon, observed only in the cohorts born towards the end of the 19th century, both in Italy and elsewhere (Robine et al 2006; Beltran-Sanchez et al 2015), genes, the "first suspect" (Passarino et al 2002), cannot be the only cause.

Neither can be smoking: sure enough, it was almost exclusively a male habit in those cohorts, but it can explain only part of the gender gap in survival, no more than 30% (Preston and Wang 2006). Differential migration is another non-negligible contributor to the gap (Caselli, Battaglini and Capacci 2018). Can others be included? Here we will focus on the link between fertility and sex-specific mortality because several papers (Kirkwood 1977, Jasienska 2017) and evolutionary/life history theories about energy allocations suggest that senescence can be seen as a by-product of the priority that natural selection places on reproduction, sometimes at the expense of later survival.

Indeed, higher parity has been found to be associated with telomere shortening (Pollack et al 2018), higher biological age (Ryan 2018), higher risks of type-2 diabetes (Mueller et al 2013) and higher overall mortality (Gagnon et al 2013; Zeng et al 2016; Bolund et al. 2016). A substantial fertility decline (from, e.g., 5-6 to 1-2 children per woman) may thus improve female post-reproductive longevity and explain gender differences in survival in the post-reproductive ages. To test this hypothesis, however, the effect of the fertility transition on survival must be isolated from other potentially confounding historical processes such as the general process of mortality decline, the spread of smoking and the obesity pandemics.

To this purpose, the Italian regions constitute a particularly promising set of observations because of two main reasons. First, the large regional variation in the onset of the fertility transition, which emerged in last decades of the 19th century in the north, but only in the 1950s or even later in the south (Livi Bacci and Breschi 1990);

Second, the possibility offered by the Italian data to neutralize the effect of the smoking pandemics and other perturbing processes in the analysis of the FMR. However, to exclude that our results may depend on the specific geographic context analyzed, we also replicated part of our analysis on 16 European countries.

2. The Italian and the European dataset

The Italian dataset. We used the data published by the Italian Institute of Statistics (ISTAT) to reconstitute the demographic history of eight cohorts born in the years 1862-66, 1872-76, …, 1932-36 in 16 Italian administrative regions (see Table A1 in the supplementary material). For each region and each cohort, we collected aggregate data on cohort fertility, the probability of survival from age 55-59 to age x (>59), the FMR at various post-reproductive ages and the proportion of ever-married women.

Figure 1 shows, as an example, how we proceeded with the reconstitution of the demographic history of the cohorts born between 1892 and 1896, which we treated as a unique 5-year birth cohort. The decennial censuses of 1951 to 2011 gave us information on the FMR and on the evolution of the proportion of never-married women. The information on the complete fertility of this cohort derives, instead, from two different types of sources. The first are two retrospective surveys on fertility taken in 1931 (not shown) and 1961, which report information on cohort fertility by region. This kind of information, however, is available only for five out of the eight cohorts that we are interested in here (from 1872-76 to 1912-16). Therefore, we resorted to a second source: the general fertility index, I_f, estimated within the European Fertility Project (Coale and Cotts Watkins 1986): in Italy and its regions this could be done every ten years, from 1861 to 1961 (Livi Bacci 1977). This cross-sectional index - which measures the fertility of a given population as a fraction of the highest fertility ever recorded (that of Hutterites) - has been shown to be a good proxy for the total fertility of the cohort born A years before, where A is the mean age at childbearing, typically not too far from 30 years (Ryder 1964). In our data, for instance, the Pearson correlation coefficient between the true cohort fertility and the general fertility index estimated when the cohort was in its 25-29 age interval turned out to be 0.93 (see Figure A2 in the online material). We will then use the Princeton general fertility index to extend our analysis back or forth in time.

For the cohorts 1912-16, …, 1932-36 we also know the lung-cancer mortality rate at 75-79 years, which we will use as a proxy for the exposure of these cohorts to the spread of the smoking pandemics.

As for the European dataset, we collected data on the FMR from the Human Mortality Database (HMD) for the cohorts born between the 1820s and the 1930s in 16 European countries (see the supplementary material for the details). Information on fertility derives, once again, from the European Fertility Project.

3. Fertility and Post-Reproductive Sex-Ratio

The sex ratio $r_x = F_x/M_x$ observed at a given age *x* reflects the entire evolution of male and female mortality and migration from conception to age x. As our analysis focuses only on the post-reproductive period, we decided to use the "double ratio" r_x/r_{55-59}, where $x > 59$, which neutralizes all the sex-specific disturbing factors up to age 59 (maternal mortality, young accident mortality, war mortality, migration etc.). Of this "double ratio" we took the logarithm, a transformation that has several advantages, one of which is that it is not bounded to positive values. In the following, we will refer to this variable as the *relative female survival* (RFS). We estimated this variable for each cohort, region and age in the Italian dataset and we put it in relation with the corresponding (complete) cohort fertility (CF), for different birth cohorts and regions.

In Figure 2, colors represent age classes (x = 75-79, 85-89, 95-99) and points represent the combination of the RFS_x and the CF observed in a given region and cohort at a specific age. Selected points, with labels, will hopefully help our readers: they represent the 1872-76 and 1912-16 cohorts of Liguria (the forerunner of the fertility transition in Italy) and Sardinia (the last region to experience its fertility transition), all four of them at ages x=85-89 years.

Figure 2 highlights two main phenomena. First, the sex-ratio underwent a temporal evolution, and the regions that had high values of this index in 2001 (north of Italy) had much lower values in the past, comparable to those recorded in Sardinia in 2001. In other words, a low FMR at high ages was probably the norm in pre-transitional populations. Second, the influence of fertility on the sex ratio becomes stronger with age. At high ages (95-99 years), fertility alone can explain about three quarter of the overall observed variance in RFS (see the corresponding R^2 value in Figure 2). We obtain very similar results if we use I_f instead of CF in our regressions (see Table 1 – first section "All cohorts").

To be sure, Figure 2 presents a simple correlation, which may be biased by other processes connected to differential survival, such as the general process of mortality decline, the spread of smoking and the obesity pandemics. To disentangle the effect of fertility from these disturbing factors, we run several checks on the association between fertility and RFS_x.

First, we estimated the association between the Princeton index of general fertility (I_f) and RFS only for the two pre-transitional cohorts born in 1862-66 and 1872-76 (Table 1 – second section "Oldest cohorts"). These cohorts are too old to have been exposed to smoking or excess nutrition in any significant way (see Figure A3 in the supplementary material). Even in these cases, fertility (I_f) turns out to exert a significant negative effect on the dependent variable (RFS_x).

Secondly, we estimated the association between I_f and RFS_x in the three youngest cohorts of our dataset, born in the years 1912-16, 1922-26 and 1932-36. These cohorts were exposed to the smoking pandemics, but we know their mortality rate due to lung cancer in the age class 75-79, and, as Preston et al (2010), we can use this as a proxy for smoking (Table 1 – third section "Youngest cohorts"). Despite the addition of these control variables to the regression model, separately for males and females, the effect of I_f on RFS_x remains negative and highly significant. Incidentally, this model explains a large part of the observed variance of RFS_x, up to 84% at 95-99 years.

Thirdly, we repeated the three analyses of Table 1 at the European level. The results of this analysis (see Table B2 in the supplementary material) confirm those presented in Table 1, and they indicate that the negative relationship between fertility and the RFS_x holds in general, and not just in Italy.

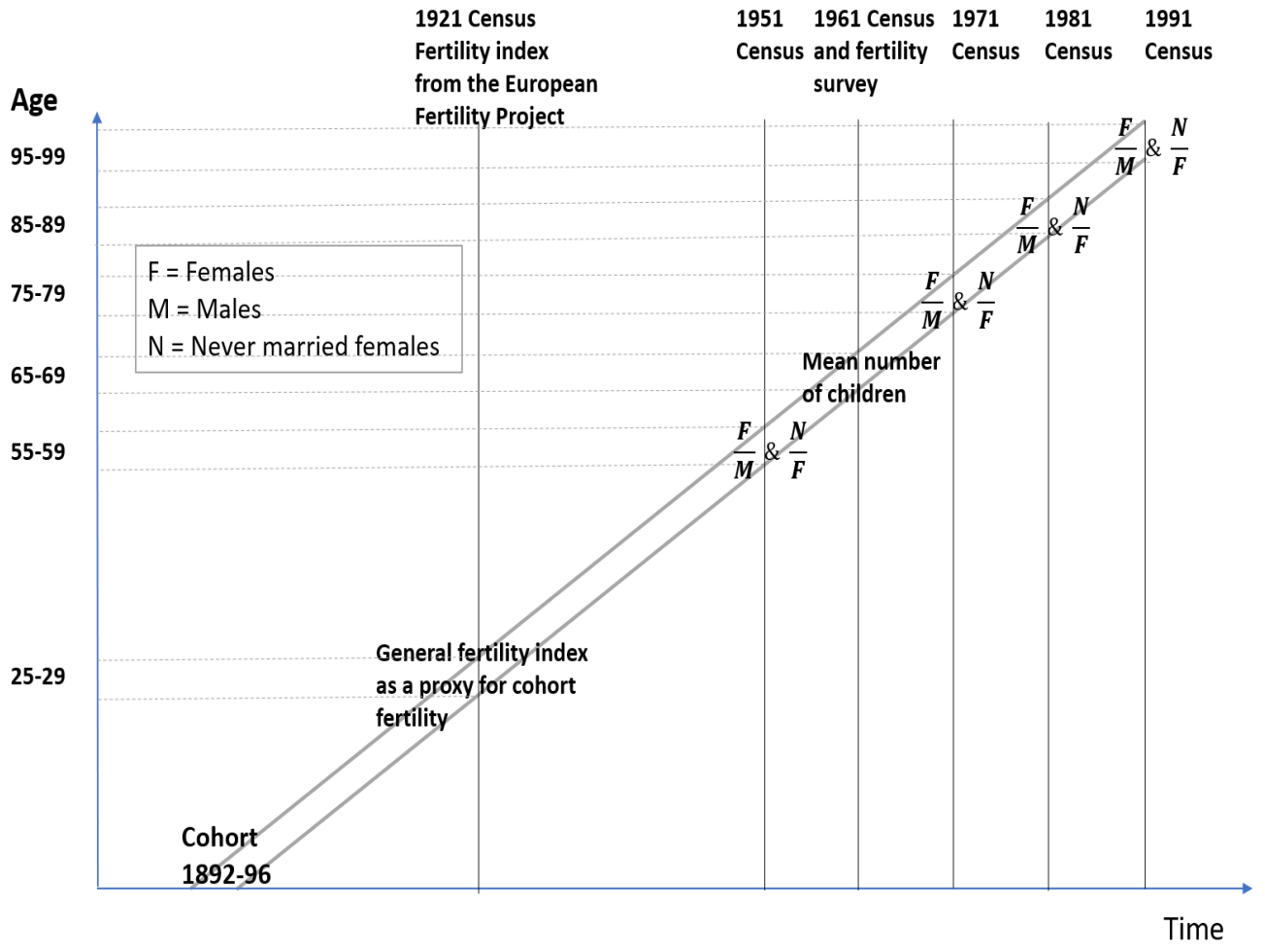
Further support to the fertility hypothesis comes from our fourth check, the analysis of the proportion of never married women at older ages. This condition is generally considered disadvantageous for female survival in the past, because: 1) it was frequently associated with health issues; 2) nulliparous women show increased risks of breast, ovarian and uterine cancer (Sun *et al*. 2016; Fraumeni *et al*. 1969); 3) never-married women rarely had children and therefore lacked support in their oldest years, when they most needed it (Zeng et al 2016; Caselli et al 2013). However, if fertility is negatively associated to women's post-reproductive survival, the disadvantage of never married women should decrease in the cohorts with higher fertility.

This is exactly what we observe in Figure 3, where the ratio between the proportions of never-married women at various old ages x and at age 55-59 (a measure of the differential survival of never vs. ever married women in the post reproductive period) is regressed on cohort fertility. The figure shows that the relationship is positive: high fertility affects negatively the survival of "mothers" (actually: of married women, most of whom had

children) and therefore, at the aggregate level, reduces the disadvantage of never married older women. At high levels of fertility, this even reverts into a survival *advantage* of never married women.

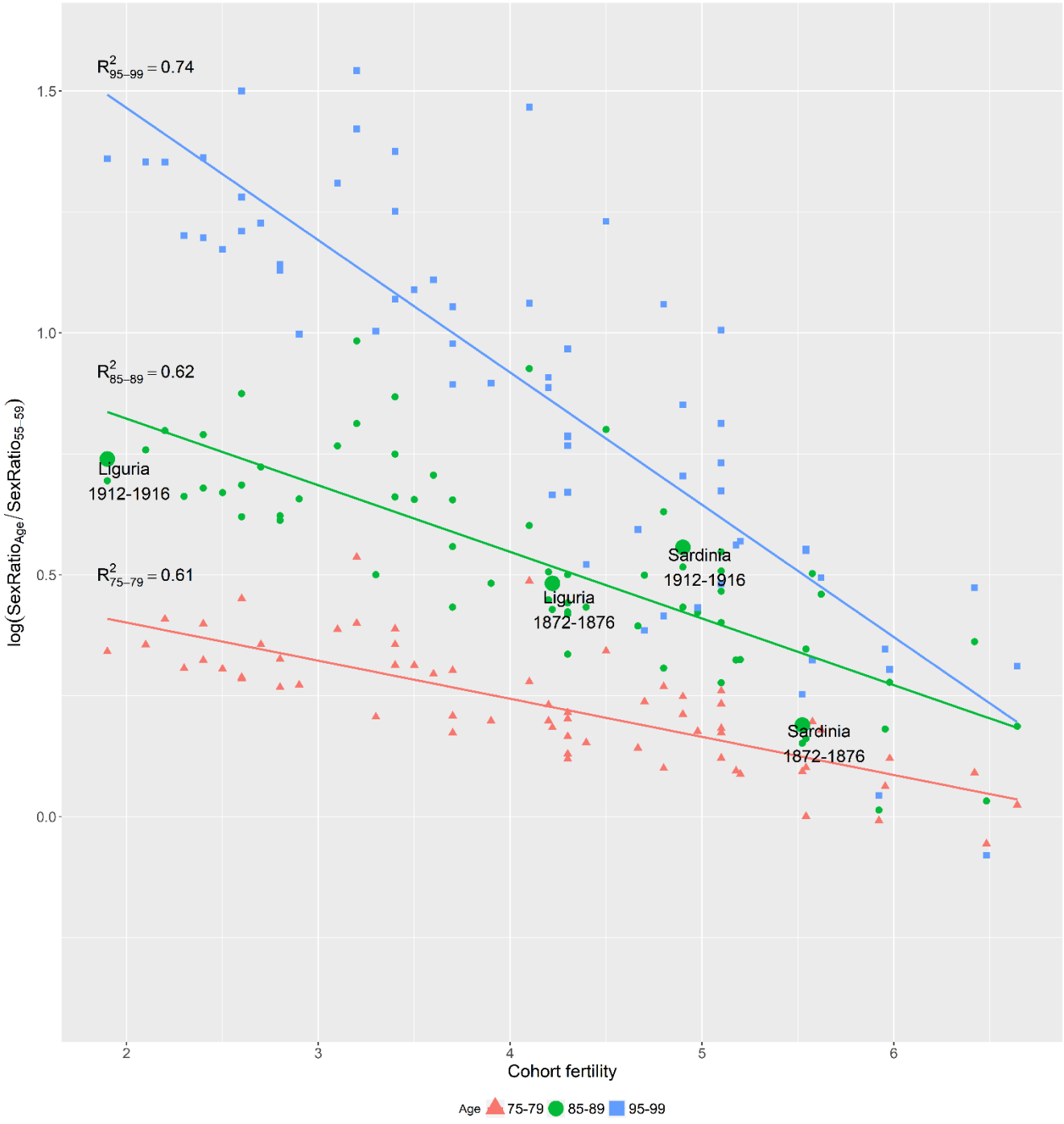
Summing up, the exceptionally low value of the FMR among older Sardinians seems to depend not only on Sardinian exceptional male longevity but also, and perhaps predominantly, on past (high) fertility, because it took Sardinia a very long time to start its fertility transition. Inside Sardinia, Nuoro was the very last province to embark on the fertility reduction trail (see Table A3 in the supplementary material), and, consistently with our hypothesis, it was also characterized by the lowest FMR_{100+} ever recorded: one male centenarian per one female centenarian (Deiana 1999). The individual-level data collected on the Sardinia centenarians also confirm the effect of fertility on female survival. Indeed, Sardinian centenarians have been found to be associated with lower fertility, higher mean age at childbearing and higher age at first and last child compared to non-centenarians belonging to the same birth cohort (Lipsi et al 2015).

Figure 1. *Data collected for the 1892-96 cohort (taken as an example)*



Note: Reconstitution of the demographic history of the cohort born in 1892-96. Information on fertility derives from the 1921 census (Princeton Indexes) and from the retrospective fertility survey of 1961. The decennial censuses from 1951 to 1991 allow us to calculate the FMR_x and the proportion of never married women at older ages x>55. We applied the same procedure to all the other cohorts (not shown here).

Figure 2. *Effect of fertility on the post-reproductive sex ratio in Italy*



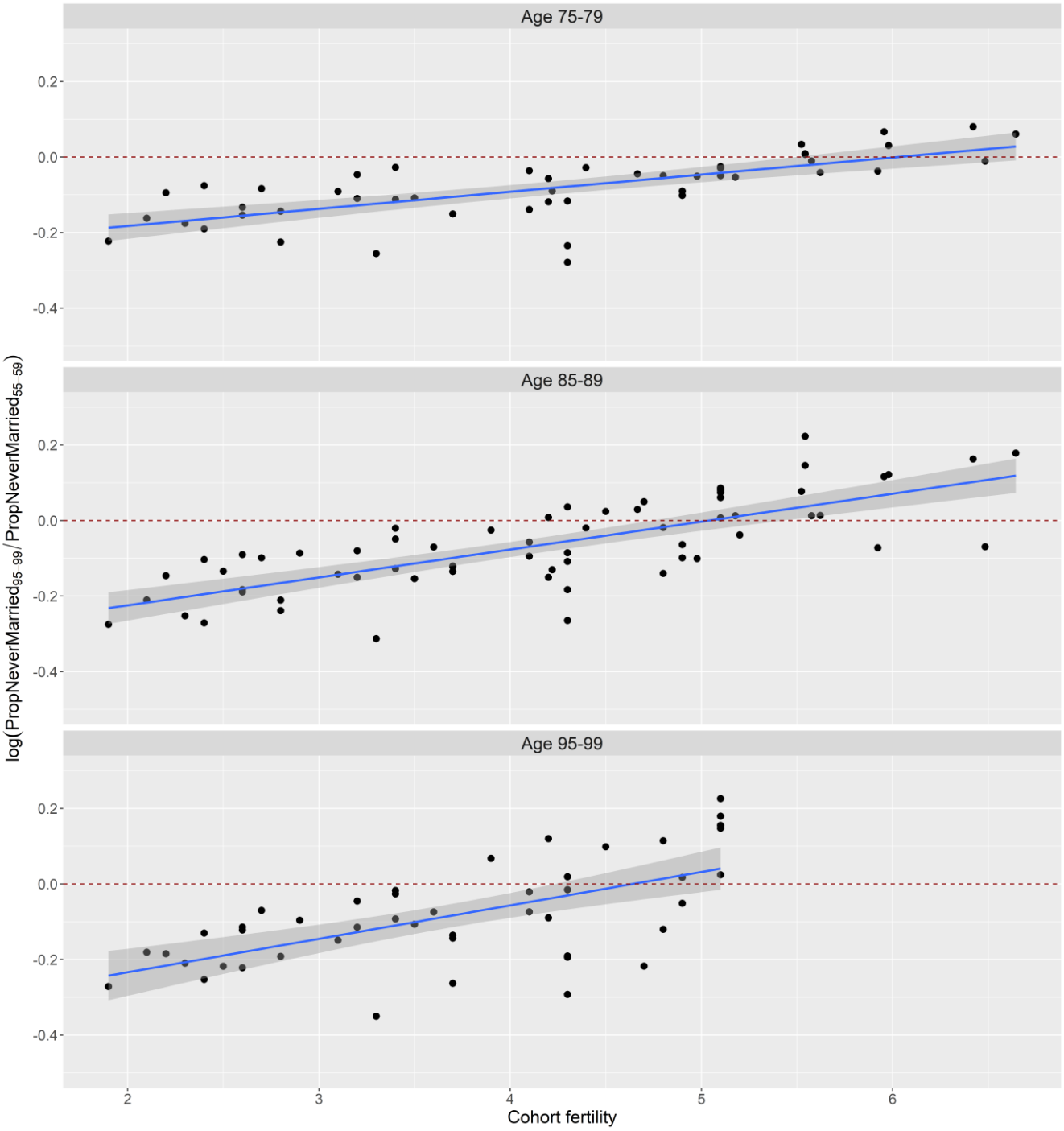
Note: colors represent age classes (x = 75-79, 85-89, 95-99), whereas the points represent the combination of the RFS_x and the CF observed in a given Italian region and cohort at a specific age.

Table 1. Checks on the relation between fertility and post-reproductive sex-ratio

	Dependent var. $RFS_{75\text{-}79}$			Dependent var. $RFS_{85\text{-}89}$			Dependent var. $RFS_{95\text{-}99}$		
All Cohorts (1862-66, ….,1932-36)									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	0.774	0.104	<0.001	1.072	0.114	<0.001	1.713	0.157	<0.001
I_f	-1.112	0.142	<0.001	-1.865	0.256	<0.001	-3.336	0.399	<0.001
Survival	-0.409	0.118	<0.001	-0.104	0.205	0.612	7.734	2.157	<0.001
	df=93	R^2=.46	S=.43	df = 93	R^2=.54	S=.67	df=77	R^2=.70	S=.93
Oldest Cohorts (1862-66, 1872-76)									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	1.081	0.306	0.004	1.366	0.264	<0.001	1.215	0.317	<0.001
I_f	-1.614	0.371	<0.001	-2.377	0.630	<0.001	-1.710	0.812	0.044
Survival	-0.882	0.490	0.095	-1.887	0.955	0.058	-19.193	8.819	0.038
	df=13	R^2=.39	S=.82	df=29	R^2=.26	S=.80	df = 29	R^2=.15	S=0.01
Youngest Coh. (1912-16, …, 1932-36)									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	0.727	0.109	<0.001	0.698	0.193	0.001	0.861	0.254	0.006
I_f	-0.511	0.124	<0.001	-0.669	0.282	0.013	-1.389	0.394	0.002
Survival	-0.685	0.106	<0.001	-0.667	0.300	0.035	4.281	4.158	0.325
Lung cancer mortality									
- males	0.003	0.001	0.001	0.007	0.002	0.002	0.011	0.003	0.004
- female	-0.006	0.003	0.019	-0.009	0.010	0.192	-0.008	0.022	0.361
	df = 43	R^2=.78	S=.09	df = 27	R^2=.71	S=.18	df = 11	R^2=.84	S=.53

Note: This table shows the regression of RFS_x on the Princeton index of general fertility (I_f). The first section refers to *all the cohorts* in our dataset: by comparison with Figure 2, it shows that the association between cohort fertility and the sex ratio at mature and old ages does not depend on how fertility is measured. The second section focuses on the *oldest cohorts* in the dataset, among whom both smoking and excess nutrition were virtually absent. The third section, instead, considers only the youngest cohorts in our dataset. Among these, smoking (predominantly on the part of men) was widespread, but we can partly control for this using lung cancer mortality as a proxy. Standard errors (SE) have been computed with the heteroskedasticity- and autocorrelation-consistent sandwich estimator (Cribari-Neto 2004). The P values for the coefficients of "I_f", "male lung mortality" and "female lung mortality" refer to a one-tail t test. For each regression we show the degrees of freedom (df), the explained variance (R^2) and the P value associated with the Shapiro test on the normality of regression residuals (S), where S < 0.05 signals a probable departure from normality.

Figure 3. Effect of *fertility on the proportion of never married women at older ages*



Note: This figure shows the relationship between cohort fertility and the proportion of never married women by age. We dropped the data on the oldest survivors from the 1971 census because these data were later found to be biased (Ventisette 1980): this is why a truncation appears in the third panel (age 95-95).

4. Conclusion

In the present study we reported demographic data supporting the evolutionary perspective on the relation between human reproduction and ageing (Jasienka et al 2017), with particular attention to the post-reproductive FMR.

A large body of the literature has identified numerous biologic factors linked to sex hormones, mitochondrial DNA, and the sex chromosome that may confer a significant survival advantage to females before and during the reproductive period (Giuliani et al 2018; Marais 2018; Vina 2005). However, a large progeny takes its toll by absorbing a high proportion of the energy reservoir of women, whose ability to survive at later ages is thus reduced (Penn 2007). In this respect, the process of fertility decline seems to represent a major turning point for human biology, because it poses the condition to move energy from reproduction to somatic maintenance, favoring female longevity. This theoretical framework seems consistent with our findings and may explain why the female survival advantage at older ages emerged worldwide only in the cohorts born at the end of the 19th century, i.e. after the start of the fertility transition.

Our results show that fertility may affect survival decades after the end of reproduction. Some biological theories on ageing (antagonistic pleiotropy, disposable soma and energy allocation) may help to explain this phenomenon. All the theories converge on that increased investment in the reproductive function may compromise investment in health, and especially health at later ages, so that faster sexual maturation and higher parity increase the risk of certain diseases later on (e.g., diabetes and cardiovascular diseases) through different biological mechanisms (Jasienka et al 2017; Ryan et al. 2018).

A recent study linking these theories with molecular data on telomere length and epigenetic age acceleration showed that reproduction in women may carry costs in term of accelerated aging through two mechanisms: one reflecting cellular turnover (i.e. telomere length), and one reflecting epigenomic regulation (DNA methylation age calculated according to Horvath epigenetic clock) (Ryan et al, 2018).

Epigenetic changes represent a potential mechanism underlying the correlation between fertility and survival decades after the end of reproduction. DNA methylation within the Developmental Origins of Health and Disease framework generally either explore associations between offspring DNA methylation in early life and later health outcomes (an example in humans is the Dutch Famine Study) (Heijmans et al., 2008, Felix et al, 2018).

This, however, transcends the scope of our article, which was merely meant to suggest that differential fertility levels may, by themselves, explain a large part of the apparently mysterious evolution over time, and geographic distribution at a given time, of the female-to-men ratios at older ages.

References

Beltrán-Sánchez H., Finch C.E., Crimmins E.M. (2014). Twentieth century surge of excess adult male mortality. *PNAS* 112(29): 8993–8998.

Bolund E., Lummaa V., Smith K.R., A. Hanson H.A., Maklakov A.A. (2016). Reduced costs of reproduction in females mediate a shift from a male-biased to a female-biased lifespan in humans. *Scientific Reports* 6:24672.

Coale A. J., Cotts Watkins S. (1986). *The Decline of Fertility in Europe*. Princeton, Princeton University Press.

Caselli G., Battaglini M., Capacci G. (2018). Cohort analysis of gender gap after one hundred years old: the role of differential migration and survival trajectories. *Journal of Ageing Science* 6(3)

Caselli G., Lipsi R. M., Lapucci E., Vaupel J. W. (2013). Exploring Sardinian longevity: women fertility and parental transmission of longevity. *Vienna Yearbook of Population Research* (Vol. 11), pp. 247–266.

Cribari-Neto, F. (2004) Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis* 45(2), pp. 215–233.

Deiana L., Ferrucci L., Pes G.M., Carru C., Delitala G., Ganau A., Mariotti S., Nieddu A., Pettinato S., Putzu P., Franceschi C., Baggio G. (1999). AKEntAnnos. The Sardinia Study of Extreme Longevity. *Aging Clinical Experimental Research* 11:142-149.

Felix JF, Cecil CAM (2018). Population DNA methylation studies in the Developmental Origins of Health and Disease (DOHaD) framework. *J Dev Orig Health Dis*. doi: 10.1017/S2040174418000442. [Epub ahead of print] PubMed PMID: 30101736.

Fraumeni JF, Lloyd JW, Smith EM, Wagoner JK (1969). Cancer mortality among nuns: role of marital status in etiology of neoplastic disease in women. *J Natl Cancer Inst* 1969;42:455–468.

Gagnon A., Smith K.R., Tremblay M., Vézina H., Paré P., Desjardins B. (2009). Is There a Trade-Off Between Fertility and Longevity? A Comparative Study of Women from Three Large Historical Databases Accounting for Mortality Selection. *Am J Hum Biol.* 21(4): 533–540.

Giuliani C., Garagnani P., Franceschi C. (2018). Genetics of Human Longevity Within an Eco-Evolutionary Nature-Nurture Framework. *Circulation Research* DOI: 10.1161/CIRCRESAHA.118.312562.

Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE,Lumey LH (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 105(44):17046-9. doi: 10.1073/pnas.0806560105. Epub Oct 27. PubMed PMID: 18955703; PubMed Central PMCID: PMC2579375.

Jasienska G, Bribiescas RG, Furberg AS, Helle S, Núñez-de la Mora A. (2017) Human reproduction and health: an evolutionary perspective. *Lancet* 390(10093):510-520. doi: 10.1016/S0140-6736(17)30573-1. Epub 2017 Jul 27. Review. PubMed PMID: 28792413.

Lipsi R.M., Caselli G., Pozzi L., Baggio G, Carru C., Franceschi C., Vaupel J. W. Deiana L. (2015). Demographic characteristics of Sardinian centenarian genealogies: Preliminary results of the AKeA2 study. *Demographic Research* 32(37): 1049-1064.

Livi Bacci M. (1977). A History of Italian Fertility During the Last Two Centuries, Princeton University Press.

Livi Bacci M., Breschi M. (1990). Italian Fertility: An Historical Account. *Journal of Family History* 15(4):385-408.

Koenig R. (2001). Sardinia's Mysterious Male Methuselahs. *Science* 291(5511):2074-2076.

Kirkwood TB (1977). Evolution of ageing. *Nature* 270:301–304.

Marais G.A.B., Gaillard J.-M., Vieira C., Plotton I., Sanlaville D., Gueyffier F., Lemaitre J.-F. (2018). Sex gap in aging and longevity: can sex chromosomes play a role? *Biol Sex Differ* 9(1):33.

Mueller N.T., Mueller N.J., Odegaard A.O., Gross M.D., Koh W.P., Yuan J.M., Pereira M.A. (2013). Higher parity is associated with an increased risk of type-II diabetes in Chinese women: the Singapore Chinese Health Study. *Epidemiology* 120:1483–1489.

Passarino G, Calignano C, Vallone A, Franceschi C, Jeune B, RobineJM, Yashin AI, Cavalli Sforza LL, De Benedictis G. (2002). Male/female ratio in centenarians: a possible role played by population genetic structure. *Exp Gerontol*. 37:1283–1289.

Pollack A.Z., Rivers K., and Ahrens K.A. (2018). Parity associated with telomere length among US reproductive age women. *Human Reproduction* 33(4):736-744.

Preston S. H., Wang H. (2006). Sex Mortality Differences in the United States: The Role of Cohort Smoking Patterns. *Demography* 43(4):631-646.

Preston S.H., Glei D.A., Wilmoth J.R. (2010). Contributions of Smoking to International Differences in Life Expectancy. In Crimmins E.M., Preston S.H., Cohen B. (Eds.), International Differences in Mortality at Older Ages: Dimensions and Sources. Washington D.C., The National Academies Press.

Penn, S. (2007). Differential fitness costs of reproduction between the sexes. PNAS 104(2): 553–558.

Robine J.-M., Caselli G., Rasulo D., Cournil A. (2007). Differentials in the femininity ratio among centenarians: Variations between northern and southern Italy from 1870. *Population Studies: A Journal of Demography* 60(1): 99-113.

Robine J-M, Caselli G., Rasulo D., Cournil A. (2006). Differentials in the femininity ratio among centenarians: Variations between northern and southern Italy from 1870. Population Studies, 60(1):99/113.

Ryan CP, Hayes MG, Lee NR, McDade TW, Jones MJ, Kobor MS, Kuzawa CW, Eisenberg DTA (2018). Reproduction predicts shorter telomeres and epigenetic age acceleration among young adult women. *Sci Rep.* Jul 23;8(1):11100. doi: 10.1038/s41598-018-29486-4. PubMed PMID: 30038336; PubMed Central PMCID: PMC6056536.

Ryder N.B. (1964). The Process of Demographic Translation. *Demography* 1(1):74-82.

Sun X, Nichols HB, Tse C-K, Bell MB, Robinson WR, Sherman ME, Olshan AF, Troester MA (2016). Association of parity and time since last birth with breast cancer prognosis by intrinsic subtype. *Cancer Epidemiol Biomark Prev* 25:60–67.

Ventisette M. (1980). Su alcuni dati dell'undicesimo censimento demografico italiano (24 Ottobre 1971). *Collana Quaderni del Dipartimento Statistico*, Università degli studi di Firenze n 8.

Vina J, Borras C, Gambini J, Sastre J, Pallardo FV (2005). Why females live longer than males? Importance of the upregulation of longevity-associated genes by oestrogenic compounds. *FEBS Lett*. 579:2541–2545.

Zeng Y., Ni Z., Liu S., Gu X., Huang Q., Liu J., Wang Q. (2016). Parity and All-cause Mortality in Women and Men: A Dose-Response Meta-Analysis of Cohort Studies, *Scientific Reports* 6(19351):1-11.

Description of the Sources Used in the Analysis

We prepared two different datasets, one for Italy (16 administrative regions), described in section A, and one for Europe (16 countries), Section B. In each case, we inserted a few additional analyses, to reinforce the results presented in the main section.

A. The Italian Dataset

We used four different types of sources: a) censuses; b) two retrospective fertility surveys; c) the data collected by Massimo Livi Bacci within the European Fertility Project (births and census data); d) data on mortality by cause.

The Italian censuses: Censuses in Italy have traditionally been taken every ten years, in the years ending in 1, since 1861, with a few exceptions (not in 1891 and in 1941; there was an extra one in 1936, but we exploited it only marginally, and we will see shortly). For the present analysis, we used the censuses of the period 1911-2011. Census data report, among other things, the number of men, women and never married women in specific age intervals, by administrative regions. Table A1 shows the data subset we used for each of our cohorts, by age and year. For instance, the 1911 census gives us information on the 1862-66 cohort when they were 45-49 years old. The demographic history of some cohorts is incomplete (truncation), in part because we lack the 1941 census and in part because by 2011 the youngest cohorts had not yet reached the ending age bracket that we use in our analysis (95-99 years). As most of the Italian censuses of the past century adopted a final open class 100+ years, we cannot extend our analysis to the centenarians. This, however, saves us the trouble of correcting for the age distortions that are typically observed at very old ages, especially after the psychological threshold of 100 years.

The censuses between 1911 and 1931 refer to the *de facto* population; after that, we refer instead to the *de jure* population: when the comparison was possible, we verified that only trivial differences emerged in the age classes selected for our analysis.

Some of the 20 regions that exist today in Italy were not always administrative units and could not be considered in our analysis: three are in the north (Valle d'Aosta, Trentino-Alto Adige and Friuli-Venezia Giulia, created only in 1951), and one in the center (Molise, data on which started to be published separately only in 1971).

Table A1. *Demographic history of selected cohorts captured by the Italian censuses*

	Census										
Cohort	**1911**	**1921**	**1931**	**1941**	**1951**	**1961**	**1971**	**1981**	**1991**	**2001**	**2011**
1862-1866	45-49	55-59	65-69	-	85-89	95-99	-	-	-	-	-
1872-1876	-	45-59	55-59	-	75-79	85-89	95-99	-	-	-	-
1882-1886	-	-	45-49	-	65-69	75-79	85-99	95-99	-	-	-
1892-1996	-	-	-	-	55-59	65-69	75-79	85-89	95-99	-	-
1902-1906	-	-	-	-	45-49	55-59	65-69	75-79	85-89	95-99	-
1912-1916	-	-	-	-	-	45-49	55-59	65-69	75-79	85-89	95-99
1922-1926	-	-	-	-	-	-	45-49	55-59	65-69	75-59	85-89
1932-1936	-	-	-	-	-	-	-	45-49	55-59	65-69	75-79

Fertility surveys: To reconstitute the complete fertility of the cohorts under scrutiny we exploited two retrospective surveys carried out with the 1931 and 1961 censuses. These surveys report for each administrative

region the average number of children by (ever married) women's age. To guarantee comparability between the two surveys, we restricted the observation to women who had been married only once. Table A2 shows that we can determine the average number of children of the cohort 1912-1916, thanks to the 1961 survey when this cohort had 45-49 years. For the 1892-1896 cohort, instead, the mean number of children has been determined when the cohort was in the 65-69 age-range. In the case of the 1872-1876 cohort, whose fertility is known through the 1931 survey, the decennial age classes published by Istat (50-59 years) overlap only partially with those that we needed here (55-59 years), but we used them anyway.

As complete cohort fertility is known at different ages for the various cohorts, selection bias is possible, at least in theory, because of the higher mortality associated with higher fertility. In Figure A1 we present an analysis of the sex ratio similar to that presented in Figure 2: in this case, however, the relative female survival is computed as r_x/r_{45-49}, (i.e. relative to a younger age class). This analysis is probably less protected from the effect of migration, but it allows us to show that fertility begins to have a measurable influence on the RFS (relative female survival) only after age 65-69. In short, only the 1892-1896 cohort, whose complete fertility is known at ages 65-69 years, can be affected by a selection bias, and very marginally so. This bias, however, if at all relevant, runs counter our hypothesis, because it tends to eliminate from the observation women with higher fertility and higher mortality.

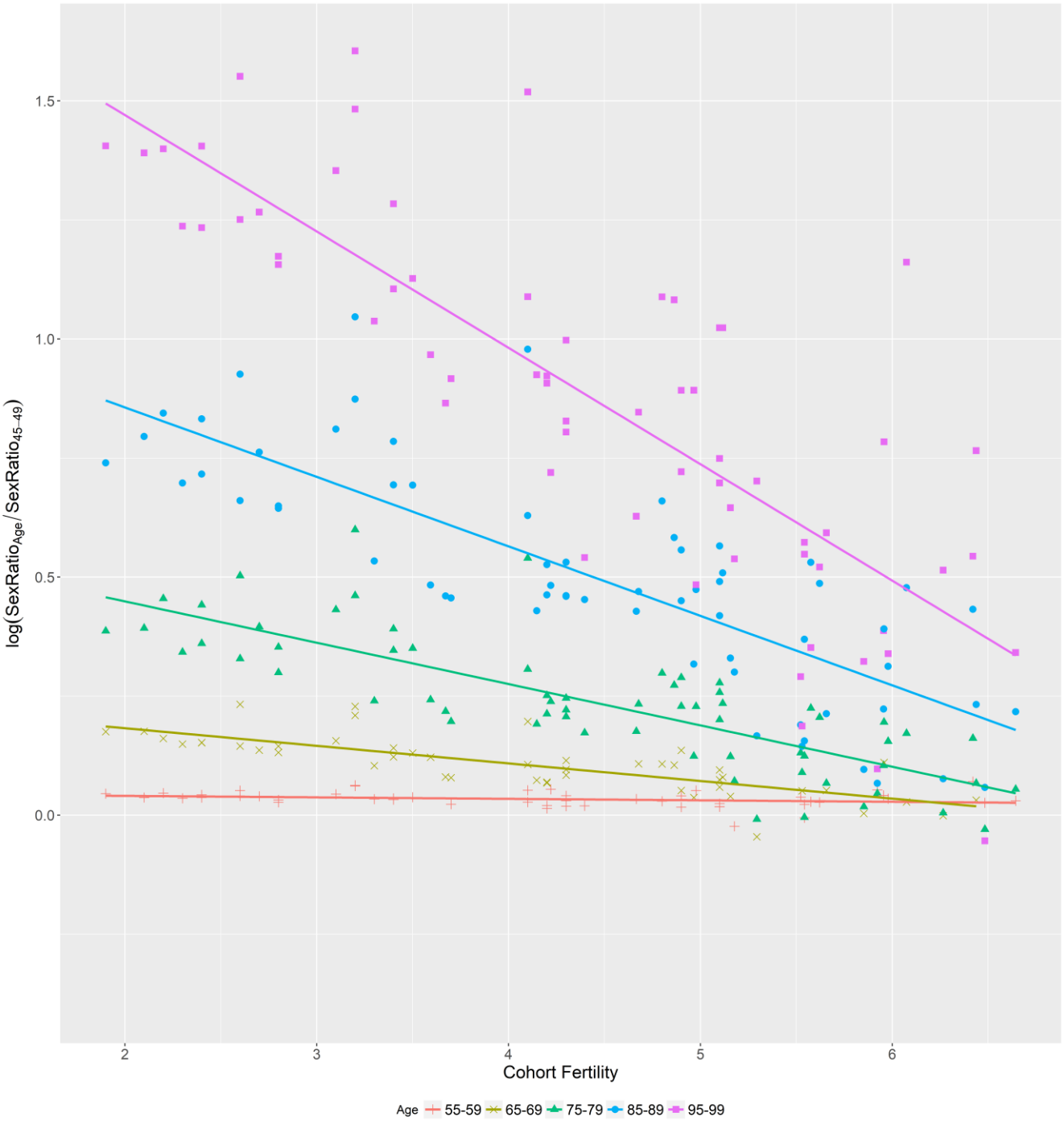
In Table A3 we show the evolution of the marital fertility index (I_g) in the Italian regions and in three Sardinian provinces. These data show that Sardinia was the latest Italian region to start its fertility transition, and that inside this island, Nuoro was the latest province.

Table A2. *Sources of information on fertility and lung cancer mortality for selected Italian cohorts*

Cohort	Survey 1931	Survey 1961	EFP (I_f)	HA (Lung Cancer)
1862-1866	-	-	1891	-
1872-1876	50-59	-	1901	-
1882-1886	45-49	-	1911	-
1892-1996	-	65-69	1921	-
1902-1906	-	55-59	1931	-
1912-1916	-	45-49	1936	1991
1922-1926	-	-	1951	2001
1932-1936	-	-	1961	2011

Note: "Survey 1931" and "Survey 1961" refer to the Italian retrospective fertility surveys carried out in 1931 and 1961. In these columns, the ranges refer to the age classes for which information on cohort complete fertility exists. "EFP (I_f)" indicates the general fertility index defined in the framework of the European Fertility Project. Each cell in this column indicates the year when I_f was computed and the cohort to which this index has been associated in this research. "HA (Lung cancer)" indicates the lung cancer mortality at 75-79 years, derived from the ISTAT "Health for All" database. Each cell indicates the year when cancer mortality has been estimated.

Figure A1. *Effect of fertility on the post-reproductive sex ratio in Italy: age from 55-59 to 95-99*



Note: colors represent age classes (x = 55-59, …, 95-99), and points represent the combination of the RFS_x and the CF observed in a given Italian region and cohort at a specific age.

Table A3. *Evolution of the index of marital fertility in the Italian regions and in three Sardinian provinces*

	Index of marital fertility (Ig)			
Italian regions	**1870-72**	**1900-02**	**1930-32**	**1950-52**
Piedmont	631	592	299	209
Liguria	630	551	294	189
Lombardy	657	683	407	286
Veneto	682	743	527	358
Emilia	597	638	386	237
Tuscany	662	595	341	235
Marche	620	632	475	299
Umbria	600	590	439	256
Lazio	610	528	465	314
Abruzzi	638	609	553	362
Campania	667	634	647	496
Puglia	701	677	643	501
Basilicata	645	606	624	474
Calabria	639	573	595	496
Sicilia	696	618	531	416
Sardegna	631	647	662	589
Italy	646	633	471	344

Sardinian provinces	**1870**	**1911**	**1931**	**1951**
Cagliari	620	698	656	603
Sassari	649	676	638	531
Nuoro	649	657	710	634

Sources: Livi Bacci and Breschi (1990), Coale and Cotts Watkins (1986).

The European Fertility Project: The European Fertility Project (EFP) (Coale and Watkins 1986) aimed at describing the process of the fertility transition in Europe: its timing and its geographical distribution. Data were collected for twenty-seven European countries, including Italy, at the national, regional and provincial level. Various measures were computed, but in this paper we considered only the general fertility index I_f, calculated every ten years (with a couple of exceptions) in the period 1861-1961, by administrative region. This is "the ratio of the births that women in a given population actually have to the number they would have if subject to a maximal well-recorded age-specific fertility schedule (that of the Hutterites)". As Table A2 (EFP column) shows, we imputed to each cohort the value of I_f recorded 25-29 years after its birth (e.g., the I_f of 1891 was considered a proxy of the completed fertility of the 1862-66 cohort). The 1912-16 cohort represents a partial exception, because it was to the value of I_f recorded in 1936, due to the lack of the 1941 census). For each region and for the cohorts born between 1872 and 1916, Figure A2 shows the (close) association between I_f (occasionally used as a proxy of cohort fertility in this paper) and the actual complete cohort fertility, derived from the two surveys of 1931 and 1961.

Figure A2. *Association between cohort fertility and the general fertility index I_f in Italy*

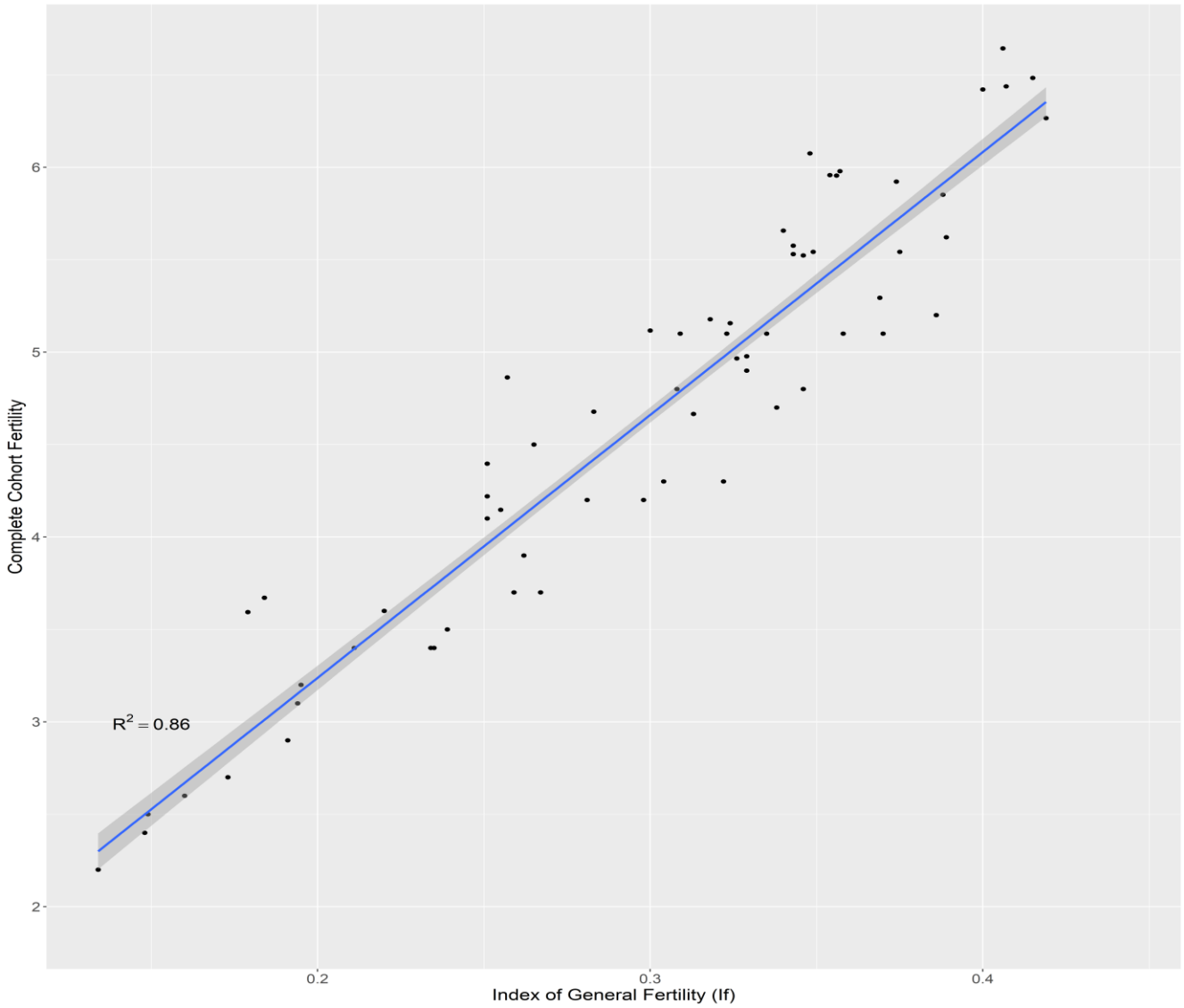


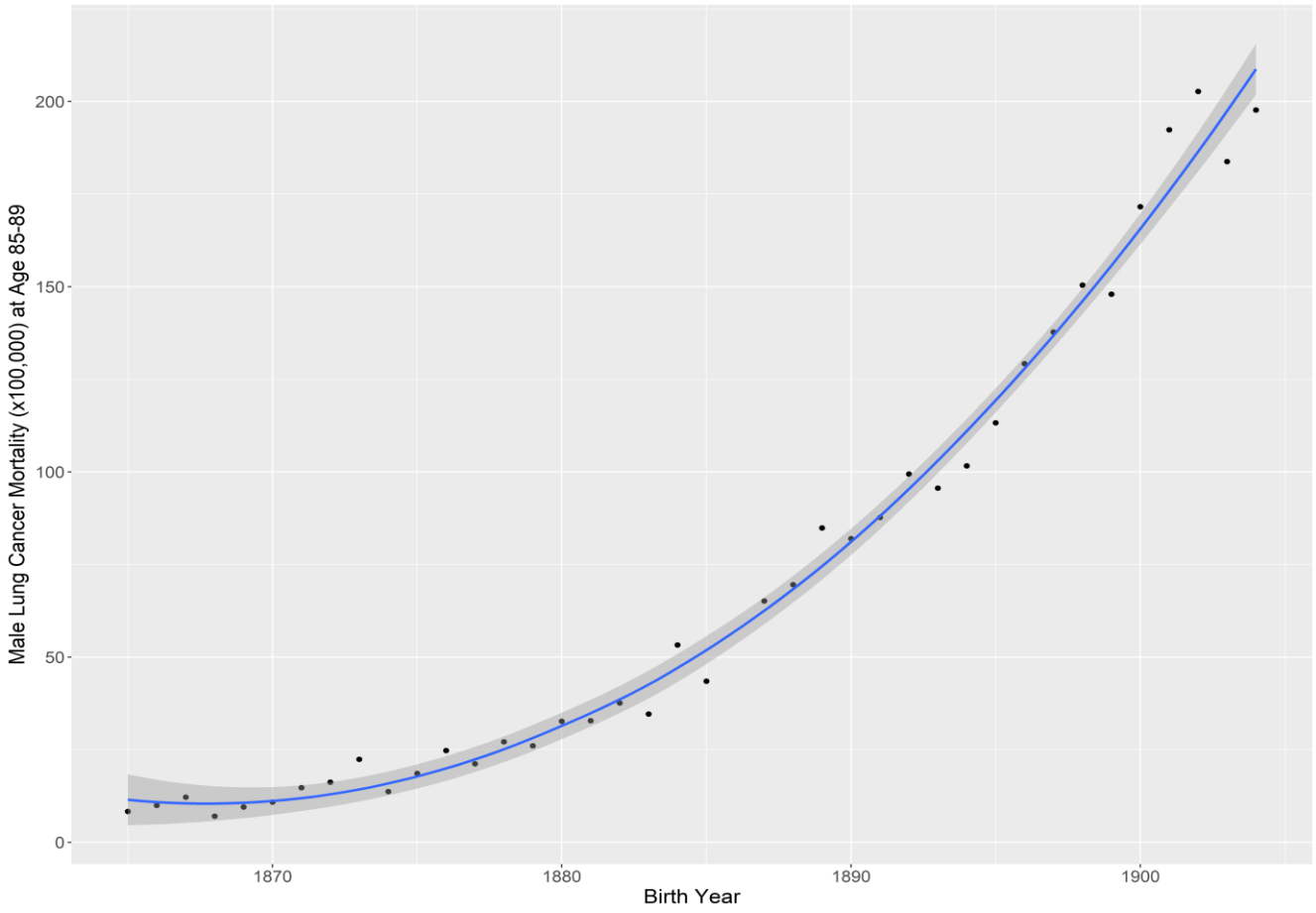
Figure 2A shows that I_f can be a good proxy of cohort fertility, when both are observable.

Mortality by cause: Following Preston et al (2010), in our analysis we use lung cancer mortality as a proxy for the damage provoked by smoking, by cohort (i.e., a combination of prevalence and intensity). Data on lung cancer mortality in Italy in the period 1951-2016 were derived from the World Health Organization Mortality Database (https://www.who.int/healthinfo/statistics/mortality_rawdata/en/). The statistics on lung cancer mortality at regional administrative level were instead derived from ISTAT's "Health for All database" (https://www.istat.it/it/archivio/14562) and cover the period 1990-2015. To make the various disease classifications (from ICD-7 to ICD-10) comparable over time, we followed the IARC recommendation (http://www-dep.iarc.fr/WHOdb/WHOdb.htm) in order. Table A2 (column HA – Lung mortality) shows that the three last cohorts in our dataset, born in 1912-16, 1922-26, 1932-36, were associated to the lung cancer mortality observed at age 75-79 respectively in the years 1991, 2001 and 2011.

For Italy as a whole, Figure A3 shows that an increase in lung cancer mortality among males aged 85-89 years begins with the cohorts born after 1875 (here we used the 85-89 age class because it is the only age class which

allows us to get information on the oldest cohorts in our database, given that the WHO series starts in 1951). The data of Figure A3 are consistent with the historical knowledge on the diffusion of smoking, which identifies in the First and the Second World Wars the beginning of the mass diffusion of this phenomenon (cigarettes were given to soldiers because of their alleged soothing properties; Pampel 2010).

Figure A3. *Evolution over time of the age-specific rate of lung cancer mortality (age 85-89): Italian male cohorts*



B. The European Dataset

To build the European dataset we used three different types of sources: a) The Human Mortality Database; b) The European Fertility Project; d) data on mortality by cause.

The Human Mortality Database (HMD): We used the HMD to get information on the population (counts) by age and sex. These data present several differences from those employed in the analysis of Italy. First, the HMD provides the intercensal reconstruction of the age structure of the population, so we are not confined to just a few points in time, as in the case of Italy, to calculate our indexes. Second, in most cases the HMD reconstruction is performed only at country level, with no further geographical detail. Third, the population counts are provided by the HMD by single year of age, whereas in Italy the computation of the sex ratio was possible only for 5-years age classes. The HMD, however, lacks information on the number of never married women, and therefore we cannot replicate this part of our analysis at the European level.

The European Fertility Project: In the case of the European dataset, we do not have information on cohort fertility: the only comparable source of information derives from the European Fertility Database (see the A section in the supplementary material). As for Italy, for our analysis we decided to focus on the (cross-sectional) general fertility index I_f.

*Mortality by cause: W*e used the lung cancer mortality observed in the age class 75-79 as a proxy for the exposure to smoking. The information on lung cancer mortality at country level was derived from the World Health Organization Mortality Database (see the A section in the supplementary material).

As an example, Figure B1 shows how we proceeded to the reconstruction of the demographic history of the 1880 Belgian cohort. We used the HMD data to get information on the sex ratio observed at ages 57, 77, 87 and 97. We selected these ages because they represent the middle of the age classes 55-59, 75-79, 85-89, 95-99 employed in the case of Italy. Each cohort in the dataset was associated to the general fertility index I_f measured in the year when the cohort had 30 years, that is in the vicinity of the mean age at childbearing. We thus imputed the value of the general fertility index estimated in 1910 to the Belgian cohort born in 1880.

Overall, the data collected for Europe cover 16 countries and 120 cohorts born over the period 1801-1931. Table B1 shows for each country and for each age the number of cohorts available for the analysis.

Figure B2 shows the association between the relative female survival (RFS_x) observed at different ages x and the general fertility index. This analysis shows that the same pattern observed in Italy (see Figure 2 in the main text) can be identified also in Europe. The case of France is worth noticing. For reasons which are not yet entirely understood, in the 19th century France had the lowest fertility in Europe. Consistently with our thesis, the 1851 French cohort presents a remarkable high value of the RFS_{87}, which Italy reached only fifty years later, with its 1901 birth cohort.

In Table B2 we present the same checks we performed for Italy on the association between fertility and the RFS_x. These checks are intended to rule out the risk that a spurious correlation between fertility decline and other historical processes may bias our results. The potential sources of bias are: mortality decline, the spread of smoking and the obesity pandemics. As the nine regressions of table B2 show, the effect of I_f on RFS_x (relative female survival) remains negative and significant in all cases, even after the inclusion of (proxies for) these disturbing processes.

Figure B1. *Data collection process for the Belgian 1880 cohort*

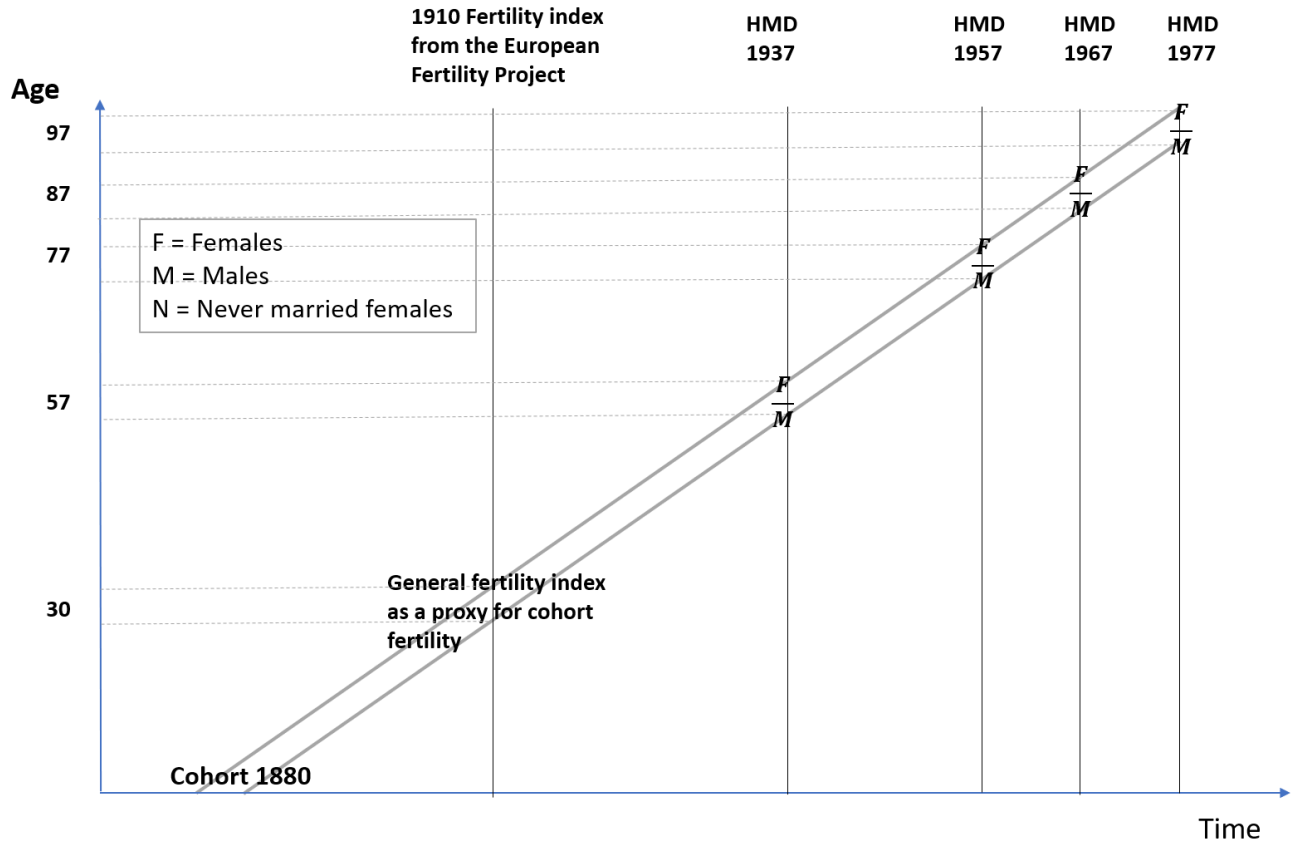


Table B1. *Cohorts in the European dataset*

Country	Cohorts	N Cohorts Age = 77	N Cohorts Age = 87	N Cohorts Age = 97
Belgium	1880-1940	5	4	4
Bulgaria	1904-1926	2	1	1
Denmark	1822-1900	9	9	7
Finland	1835-1930	11	10	4
France	1801-1931	19	18	18
England-Wales	1821-1931	10	9	9
Hungary	1930-1930	1	1	-
Italy	1831-1931	11	10	9
Netherland	1829-1930	9	9	8
Norway	1845-1930	6	5	5
Poland	1930-1930	1	1	-
Portugal	1900-1930	4	3	2
Scotland	1831-1931	9	8	7
Spain	1870-1930	7	7	6
Sweden	1850-1930	5	5	4
Switzerland	1830-1930	11	11	7
Total		120	111	91

Figure B2. *Effect of fertility on the post-reproductive sex ratio in Europe*

Note: colors represent age classes (x = 75-79, 85-89, 95-99), whereas points represent the combination of the RFS_x and the CF observed in a given European country and cohort at a specific age.

Table B2. *Checks on the relation between fertility (I_f) and the post-reproductive sex-ratio (RFS, Relative Female Survival) in Europe*

	Dependent var. RFS_{77}			Dependent var. RFS_{87}			Dependent var. RFS_{97}		
All Cohorts									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	0.694	0.086	<0.001	1.224	0.113	<0.001	1.920	0.175	<0.001
I_f	-1.267	0.182	<0.001	-2.251	0.333	<0.001	-3.154	0.577	<0.001
Survival	-0.193	0.092	0.038	-0.292	0.219	0.184	-0.682	2.931	0.816
	df=117	R^2=.38	S=0.04	df = 108	R^2=.40	S=.67	df=88	R^2=.37	S=.79
Oldest Cohorts (<1880)									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	0.666	0.107	<0.001	1.132	0.176	<0.001	1.916	0.257	<0.001
I_f	-0.812	0.281	0.003	-1.721	0.493	<0.001	-3.022	0.799	<0.001
Survival	-0.573	0.277	0.002	-1.728	0.817	0.039	-18.255	25.516	0.478
	df=56	R^2=.24	S=.53	df=56	R^2=.23	S=.32	df = 46	R^2=.19	S=.53
Youngest Cohorts									
	Estimate	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
Intercept	0.724	0.085	<0.001	0.918	0.119	<0.001	1.433	0.194	<0.001
I_f	-0.794	0.249	0.002	-1.372	0.421	0.001	-2.595	0.597	<0.001
Survival	-0.601	0.108	<0.001	-0.632	0.175	0.001	3.218	4.465	0.478
Lung cancer mortality									
- males	2.4e-04	5.0e-05	<0.001	4.9e-04	9.1e-05	<0.001	1.2e-04	1.7e-04	0.249
- female	9.2e-05	1.8e-04	0.305	-2.8e-04	3.6e-04	0.219	0.005	0.001	<0.001
	df = 32	R^2=.62	S=.62	df = 27	R^2=.61	S=.18	df = 22	R^2=.68	S=.43

Note: This table shows the regression of RFS_x (relative female survival) on the general fertility index I_f. The first section refers to *all the cohorts* in our dataset. The second section focuses only on the *oldest cohorts*, who were not significantly exposed to smoking and excess nutrition. The third section, instead, considers only the youngest cohorts for which information on lung cancer mortality in the age class 75-79 is available. Standard errors (SE) have been computed with the heteroskedasticity- and autocorrelation-consistent sandwich estimator (Cribari-Neto 2004). The P values for the coefficients of "I_f", "male lung mortality" and "female lung mortality" refer to a one-tail t test. For each regression, we indicated the degrees of freedom (df), the explained variance (R^2) and the P value associated with the Shapiro test on the normality of regression residuals (S), where S < 0.05 signals a probable departure from normality.

References

Censuses:

Ministero di Agricoltura, Industria e Commercio (1914). *Censimento della popolazione del Regno d'Italia al 10 giugno 1911, Volume II, Popolazione presente classificata per sesso, età stato civile ed istruzione*. Roma, Tipogradia Nazionale di G. Bertero E C.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Abruzzo e Molise*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Basilicata*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Calabrie*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Campania*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Emilia*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Lazio*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Liguria*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Lombardia*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Marche*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Piemonte*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Puglia*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Sardegna*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921. Sicilia*. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921*. Toscana. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921*. Umbria. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica (1927). *Popolazione del Regno d'Italia al 1° dicembre 1921*. Veneto. Roma, Stabilimento Poligrafico per l'Amministrazione dello Stato.

Istituto Centrale di Statistica del Regno d'Italia (1933). *VII Censimento Generale della Popolazione al 21 Aprile 1931-IX, Risultati Sommari Italia Settentrionale*. Roma, Istituto Poligrafico dello Stato Libreria.

Istituto Centrale di Statistica del Regno d'Italia (1934). *VII Censimento Generale della Popolazione al 21 Aprile 1931-IX, Risultati Sommari Italia Centrale, Meridionale e Insulare*. Roma, Istituto Poligrafico dello Stato Libreria.

Istituto Centrale di Statistica (1956). *IX Censimento Generale della Popolazione 4 novembre 1951, Volume III, Sesso Età Stato Civile Luogo di Nascita*. Roma Soc. Abete.

Istituto Centrale di Statistica (1968). *10° Censimento Generale della Popolazione 15 Ottobre 1961, Volume V, Sesso Età Stato Civile Luogo di Nascita*. Napoli Tip. STAGRAME.

Istituto Centrale di Statistica (1974). *11° Censimento Generale della Popolazione 24 ottobre 1971*, Volume V, Sesso Età Stato Civile. Roma.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Abruzzo*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Basilicata*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Calabria*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Campania*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Emilia-Romagna*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Lazio*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Liguria*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Lombardia*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Marche*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Molise*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Piemonte*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Puglia*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Sardegna*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Sicilia*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Toscana*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Umbria*. Roma, Abete Grafica spa.

Istituto Centrale di Statistica (1985). *Censimento Generale della Popolazione 25 Ottobre 1981, Volume II, dati sulle caratteristiche strutturali della popolazione e delle abitazioni. Veneto*. Roma, Abete Grafica spa.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Abruzzo*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Basilicata*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Calabria*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Campania*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Emilia-Romagna*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Lazio*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Liguria*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Lombardia*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Marche*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Molise*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Piemonte*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Puglia*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Sardegna*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Sicilia*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Toscana*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Umbria*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (1994). *13° Censimento Generale della Popolazione e delle Abitazioni 20 ottobre 1991. Fascicolo Regionale Veneto*. Roma, Istituto Poligrafico e Zecca dello Stato.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Abruzzo*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Basilicata*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Calabria*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Campania*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Emilia-Romagna*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Lazio*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Liguria*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Lombardia*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Marche*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Molise*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Piemonte*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Puglia*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Sardegna*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Sicilia*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Toscana*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Umbria*. Roma, System Graphic S.r.l.

Istituto Nazionale di Statistica (2006). *14° Censimento Generale della Popolazione e delle Abitazioni 21 ottobre 2001. Popolazione Residente e Abitazioni nelle Regioni Italiane. Veneto*. Roma, System Graphic S.r.l.

The data from the 2011 census can be retrieved at http://dati.istat.it/.

Human Mortality Database https://www.mortality.org/

Surveys on fertility:

Istituto Centrale di Statistica del Regno d'Italia (1936). *VII Censimento Generale della Popolazione, Volume VI, Indagine sulla Fecondità della Donna*. Roma, Tipografia I. Failli.

Istituto Centrale di Statistica (1974). Indagine sulla Fecondità della Donna. *Note e Relazioni* n. 50

https://opr.princeton.edu/archive/pefp/

Mortality by Causes:

https://www.who.int/healthinfo/statistics/mortality_rawdata/en/

https://www.istat.it/it/archivio/14562