Investigaciones de Historia Económica - Economic History Research 13 (2017) 166–179

Asociación Española de Historia Económica

Investigaciones de Historia Económica - Economic History Research

www.elsevier.es/ihe

Article

Gender equality in human capital and fertility in the European regions in the past

Ralph Hippe[a,*,1], Faustine Perrin[b]

[a] European Commission, Joint Research Centre (JRC), Directorate for Growth and Innovation, Human Capital and Employment Unit, Ispra, Italy
[b] Department of Economic History, School of Economics and Management, Lund University, Sweden

ARTICLE INFO

Article history:
Received 30 September 2016
Accepted 8 February 2017
Available online 24 March 2017

JEL classification:
N33
N93
O18

Keywords:
Human Capital
Gender equality
Fertility
European regions

ABSTRACT

Gender inequality in human capital has been shown to be an important indicator of economic development, but has remained unexplored in a European history perspective. Using a new and large historical database on male and female literacy rates, new evidence is presented on the distribution of gender inequality in human capital in European regions in 1900 and 1960. An analysis is made of the distribution of fertility rates to determine the relationship between gender equality in education and the demographic transition. The results show a reversal of educational fortunes and regional fertility rates. Regions with lower fertility rates in the past tend to display higher rates today.

Igualdad de género en capital humano y fecundidad en las regiones europeas en el pasado

Códigos JEL:
N33
N93
O18

Palabras clave:
Capital humano
Igualdad de género
Fecundidad
Regiones europeas

RESUMEN

La igualdad de género en capital humano es un indicador importante del desarrollo económico, pero no ha sido explorada en una perspectiva europea. Utilizando una nueva y extensa base de datos históricos de alfabetización masculina y femenina, mostramos nuevos resultados de la igualdad de género en capital humano en la regiones europeas en 1900 y 1960. Así, exploramos la distribución de las tasas de fecundidad para entender mejor la relación entre la igualdad de género y la transición demográfica. Los resultados muestran cambios importantes a través del tiempo.

1. Introduction

Human capital theory, as originally advanced by Becker (e.g., Becker, 1960; Becker and Tomes, 1986; Becker et al., 1990) and more recently by the Unified Growth Theory (e.g., Galor and Weil, 1996; Galor, 2005, 2012), has emphasized the importance of human capital in explaining the demographic transition and economic growth. In particular, the renowned quantity–quality (QQ) trade-off between the number and the education of children has attracted the attention of many researchers during the last decades and particularly during the last years (e.g., Hanushek, 1992; Guinnane, 2008; Bleakley and Lange, 2009; Becker et al., 2010, 2012; Diebolt et al., 2015, 2016; De la Croix and Perrin, 2016).

* Corresponding author.
E-mail address: Ralph.HIPPE@ec.europa.eu (R. Hippe).

[1] The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

https://doi.org/10.1016/j.ihe.2017.02.001
1698-6989/

The European fertility decline was a predecessor for similar transitions in many continents in the world in the past and in the present. The European Fertility Project (EFP) carried out at Princeton University in the 1960s and 1970s was probably the most important endeavor in this area and had the ambition to provide answers to the major questions surrounding the European demographic transition.[1] This major initiative significantly contributed to improve the general understanding of the phenomenon of demographic transition in Europe and its determinants.

The EFP did not place particular emphasis on the role played by human capital in Europe. Instead, human capital was considered as having only contributed to a minor extent to the fertility decline, with the exception of a few isolated cases (e.g. Switzerland, see Van de Walle, 1980). The specific role of female education was even more largely neglected by the historical literature. This left aside aspect is more than surprising as "education, particularly of girls, has been shown to be the factor most closely related to fertility decline, by delaying marriage and first births" (Kirk, 1996, p. 377).

Since the end of the EFP, the impact of human capital has remained unexplored at the European level. A few analyses have recently been conducted at the county-level for Prussia (e.g., Becker et al., 2010; Becker et al., 2012, and Becker et al., 2013 focusing more specifically on the role of women's education) and France (Murphy, 2015; Bignon and García-Peñalosa, 2016; as well as Diebolt et al., 2015, 2016; and De la Croix and Perrin, 2016, accounting for a gender dimension in their analyses) but none yet explored the situation in a European comparative perspective.

Hippe and co-authors have recently provided a European perspective on regional human capital levels (e.g., Hippe and Baten, 2012; Hippe, 2012, 2013; Diebolt and Hippe, 2016a,b), however, their analyses do not account for gender differences and fertility. Thus, despite the important implications for the demographic transition and long-run economic growth (Diebolt and Perrin, 2013a,b), the difference between the human capital levels of men and women has remained unexplored in a European regional perspective.

Why is gender parity in education important? Gender parity in education represents the power balances between the sexes in general and between husband and spouse in particular (De Moor and van Zanden, 2010, p. 3). As shown by Diebolt and Perrin (2013a,b), changes in gender relations toward greater equality would be at the origin of the fertility transition and contributed to engage the take-off to modern economic growth. But gender equality is not only a relevant issue for history; it has also huge economic implications today. For example, a recent McKinsey Global Institute study shows that "if all [worldwide] countries were to match the progress toward gender parity of the best performer in their region, it could produce a boost to annual global GDP of as much as $12 trillion in 2025" (McKinsey Global Institute, 2016, p. 1). Such numbers emphasize once again the relevance of improving our knowledge on gender equality, and notably on the distribution of gender equality in human capital in the past and present.

Therefore, this paper aims at providing a comprehensive description of the geographical distribution of gender differences in human capital and fertility regimes in the past and present. For this purpose, we construct a new and large historical database on male and female literacy rates and match it with data on fertility and nuptiality by Coale and Watkins (1986). Our analysis focuses on literacy and fertility rates in the European regions in the past. For data availability reasons, we decide to focus on two specific years: 1900 and 1960. We separately explore the evolution and patterns of regional fertility rates and gender equality in human capital in Europe using correlation matrices and scatter plots. We also use OLS regression analyses to provide further econometric evidence. The empirical investigation of the relationships between gender equality, human capital, and fertility confirms the existence of: (i) a positive and significant association between gender equality and human capital; and (ii) a negative significant association between gender equality and fertility in 1900, which then becomes insignificant in 1960.

We also note that there is a reversal between educational fortunes and regional fertility rates in Europe. While men had mostly higher basic human capital than girls in the past, we observe the opposite situation today (i.e., on average, girls possess a higher educational background). In addition, we find that regions with higher fertility rates in 1900 tend to present lower rates nowadays.

The paper is structured as follows. Section 2 provides a summary of the literature on the European fertility transition and on its links with endowments in human capital. Section 3 presents the data and approach used to conduct our analysis. Section 4 gives an overview of the stylized facts of the fertility decline in Europe and beyond. Section 5 discusses the geographical distribution of gender inequality in human capital in the past. Next, Section 6 investigates the relationship between gender equality, human capital, and fertility in the past. Finally, Section 7 summarizes our findings and opens the discussion on future research.

2. Related literature on fertility and human capital

The concept of demographic transition is rather new and was first defined by Frank Notestein in 1945 (Notestein, 1945; Kirk, 1996).[2] During the following decades, many researchers focused their attention on the issue and worked at describing and explaining this social phenomenon.

2.1. Theoretical background

Theorists have examined a set of plausible explanations being at the root of the demographic transition. The earliest theories on the subject were established by Leibenstein (1957), Becker (1960), but also Easterlin (1969), and Caldwell (1976), among others (Lee, 2015). Standard economic theory was then developed by economists – at about the same period by Schultz (1974) and Becker (1981). In their earliest works, the fertility behaviors of each phase, i.e. before and during the demographic transition, were supposed to be rational. Accordingly, the high fertility rates, which characterized most of European (and worldwide) history, were considered as the product of individuals' and communities' rational choices. However, the demand for children – a key force in the process of demographic transition – changed over time. More precisely, the demand for children was affected by changes in factors such as prices, incomes, and preferences. Lower demand for children meant lower fertility rates, which could therefore initiate the process of demographic transition. These contributions constitute the starting point of the theories of the demographic transition. Since these seminal works, various factors and theories have been suggested and developed to explain the fertility dynamics (e.g. role played by income, mortality decline, value of time, contraception, quantity-quality trade-off, among others).

The gradual rise in the demand for human capital along the process of industrialization has been seen as a prime force leading to the onset of the demographic transition (Galor, 2012), specifically during the second phase of the Industrial Revolution. Taking family

[1] The main objectives of the European Fertility Project were: (i) to create a quantitative record of the European fertility transition; and (ii) to determine the social and economic circumstances that prevailed when the modern decline in fertility began in order to highlight the causal mechanisms of the fertility transition.

[2] Although there have been earlier attempts in defining this phenomenon (Kirk, 1996).

as a single decision-maker, Becker et al. (1990) study the relationship between human capital, fertility and economic growth. In this model with altruistic parents, higher productivity leads to higher wages and favors human capital accumulation which in turn raises the opportunity cost of children. Becker et al. (1990) explicitly consider the difference between the quantity and the quality of children in terms of costs (see also Becker and Lewis, 1973; Willis, 1973).[3]

The trade-off between fertility and education is a key ingredient of unified growth theory, which models the transition from Malthusian stagnation to sustained growth. Galor and Weil (1999, 2000) argue that the acceleration in the rate of technological progress gradually increased the demand for human capital and induced parents to invest in the human capital of their offspring. The increase in the rate of technological progress and the associated increase in the demand for human capital contributed to (i) ease households' budget constraints and provided more resources for the quality as well as the quantity of children; and (ii) induced a reallocation of these increased resources toward child quality. According to the authors, the rate of technological progress stimulated by human capital accumulation eventually induced a reduction in fertility rates.

The decline in the gender gap in education is another argument advanced as a reinforcing mechanism impacting fertility rates. Lagerlöf (2003) investigates the relationship between fertility, gender equality in education and economic growth. The author argues that gender equality in education has a positive impact on economic growth because of its effects on fertility and on the human capital of children. In particular, he shows that more a society moves toward gender equality, higher becomes the women's human capital, and consequently, families choose to substitute quantity for quality in children. More recently, and particularly relevant for this paper, Diebolt and Perrin (2013a,b) assessed the importance of the determinative role played by improvements in gender equality in the economic and demographic development of modern societies. In particular, they show that the increase in gender equality and the rise in technological progress created higher opportunities for women to invest in skilled human capital. The negative correlation linking maternal investments in human capital and fertility eventually encouraged families to have fewer children but to invest more in their quality.

2.2. Empirical literature

Divisions have emerged between those who believed that, before the transition, fertility was not subject to rational choice (in the line of Becker's models) and those who believe that it was the case (Lee, 2015).[4]

2.2.1. The European fertility project

The Princeton fertility project, or European fertility project (EFP), led by Ansley Coale (see Coale and Watkins, 1986), investigates the social and economic circumstances that prevailed when the modern decline in fertility began, with the ambition to disentangle the causal mechanisms of the fertility transition.

It is the most important study of historical fertility patterns in Europe. The project resulted in the publication of several 'high quality' articles and monographs and "was regarded as the definitive study of the subject" (Kirk, 1996, p. 366). One important finding of the EFP was to highlight the existence of varied socio-economic conditions under which the demographic transition took place in the European countries. The historical investigations conducted under the umbrella of the project uncovered additional key features of European fertility patterns. In particular, it demonstrated that large parts of the populations did not voluntarily limit family size before the transition (Knodel and van de Walle, 1979). It also showed that once the fertility decline was engaged, it was an irreversible process. Finally, the EFP concluded that cultural factors played a crucial role on fertility behavior independently of other socio-economic conditions. In contrast, economic factors and social variables were argued to be subordinate to understand fertility dynamics.

2.2.2. The role of nuptiality

A number of studies on the historical development of population have also recognized the direct and indirect influence of nuptiality on fertility patterns. Hajnal (1965) demonstrated that a specific marriage pattern took place in Western Europe from the 16th century until World War I, i.e. the so-called Western European marriage pattern (WEMP; Hajnal, 1965). Yet, Hajnal was not the first one to observe important differences in fertility behaviors within Europe. Malthus already pointed out the existence of such differences since the turn of the 18th century (Malthus, 1798/1960, Engelen and Puschmann, 2011). Nevertheless, Hajnal became famous for highlighting the existence of an imaginary line dividing Western and Eastern Europe, based on their differences in marriage customs and habits. The 'Hajnal line' extends from St. Petersburg in Russia down to Trieste in Italy. On the Western side of this line, marriages are characterized by three main features: (i) a late age at marriage for women – around the age of 25; (ii) an important share of females who never married – above 10%; and (iii) a low share of illegitimate births – below 2% (Hajnal, 1965). This pattern was unique[5]: Western Europe was the only region in the world where couples married at such high ages.[6] Nuptiality was a way to regulate fertility. Late age at marriage was often justified by economic reasons, notably because of the high costs involved in setting up a separate household after marriage. Couples needed time to acquire these resources, which necessarily delayed marriage (Foreman-Peck, 2011). Three factors were crucial to the apparition of the WEMP: (i) the preponderance of consensual marriage; (ii) the possibilities of inheritance for women; and (iii) the easy access to the labor market (De Moor and van Zanden, 2010).

In contrast, early marriages were quasi-universal in Eastern Europe, so that almost every man and woman got married at a point in time. Eastern European societies were characterized by 'high pressure', i.e. featured both high birth and death rates, while Western European societies were described as 'low pressure' societies due to their low birth and death rates (Wrigley and Schofield, 1989). The Western marriage pattern, therefore, contributed to reduce fertility, as the fecund period shrank with the retreat of the age at marriage.

At the same time as it reduced fertility, late marriage enabled women to increase their human capital. More specifically, Foreman-Peck states that the Western European marriage pattern "allowed women to increase their human capital both formally and informally in the years before child bearing so that more informed mothers brought up better educated offspring. The

[3] See Doepke (2015) for a review of Becker's contribution to the analysis of fertility and the child quantity-quality concept.

[4] See De la Croix and Perrin (2016) for an empirical investigation of how much a rational-choice model can explain of the temporal and spatial variation in fertility and education.

[5] Perrin (2013, 2015) has, however, recently shown that the French marriage patterns substantially differed from the rest of Western Europe as it completely reversed its course at the turn of the 19th century, accounting for a decline in the age at marriage, a decrease in the share of definitive celibacy, and a rise in the share of illegitimate births.

[6] However, marrying at a high age is nowadays also common in other parts of the world, including the Arab world (Engelen and Puschmann, 2011).

demographic pattern influenced the stock of human capital and directly contributed to Western Europe's development advantage." (Foreman-Peck, 2011, p. 292). More specifically, "lower mortality requires fewer births, higher age at marriage and permits greater child quality, which in due course raises productivity and innovation" (Foreman-Peck, 2011, p. 297).

2.2.3. Women's education and fertility

Thus, education has long been recognized as a crucial factor influencing fertility patterns (see, for instance, Cochrane, 1979; Cleland and Rodriguez, 1988, among many others). But surprisingly, we still know very little about the relationship linking women's education and fertility limitation in the past. This is all the more surprising since the massive expansion of education in the European regions occurred simultaneously to the onset of fertility declines.

Very few studies have investigated the effect of women's education on fertility in the past. Becker et al. (2013) is one of the rare works doing so at the regional level. Using county-level data for 19th century Prussia, the authors find a strong and unconditional effect of women's primary education on fertility. In particular, they show that a 10 percentage point increase in the female enrollment rate in 1816 is associated with a decrease of about 6 children per 100 women in 1867. Conducting an analysis on the second half of the 19th century in France, Diebolt et al. (2015) find a negative association between the variations in literacy rates and the variations in fertility rates. This result is strongly significant for women only. Counties with higher improvements in female literacy display stronger fertility decline in France at the turn of the 19th century. In a further analysis investigating how much a rational-choice model can explain the temporal and spatial variation in fertility and school enrollment in France over the 19th century, De la Croix and Perrin (2016) stress the importance of mother's education as a key factor shaping the incentives to have children. In particular, they show how quickly fertility drops when mothers' education increases: a rise in enrollment rates by 0.5 (i.e. about three years of education) implies a drop in fertility by one third of a child.

Finally, Diebolt et al. (2016) directly explore the role played by gender equality, measured by the gender gap index capturing the size of the gap between men and women in economic participation and opportunities, educational attainment, and health and survival (see Perrin, 2014), on the education–fertility relationship in the mid-19th century. The authors find that gender equality plays a relevant part in understanding the fertility transition.

In sum, despite the existence of a few works exploring women's education, gender relations, and fertility at the regional level in the past, our knowledge of these indicators and the relationships between them remain significantly limited. In addition, we still know very little about the distribution of these indicators at the European regional level.

3. Data and methodology

The ambition of this paper is to provide a first comprehensive description of the geographical distribution of gender differences in human capital and fertility rates in the past. The data used in this study includes human capital and fertility indicators in the European regions, with a particular focus on the years 1900 and 1960. We extend the newly created database by Diebolt and Hippe (2016a) by distinguishing male and female endowments in human capital, proxied by literacy rates.

3.1. Human capital variables

A range of countries does not provide literacy measures in their 1900 census nor in their 1960 census. This is notably the case of Scandinavian countries, such as Denmark and Sweden, but also of Germany, Switzerland, and the Netherlands (see also Diebolt and Hippe, 2016a). The countries which did not report literacy abilities in their censuses were encouraged to do so for various reasons. The Swiss administration, for instance, considered that a sufficient level of literacy was already attained in 1860 (as documented by the 1860 census, Statistisches Bureau, 1862) and, therefore, did not require further statistics. It based its reasoning on military data showing that 93% of recruits were able to read and write in the region of Bern, and the ratio was even 100% in the region of Solothurn, already in the middle of the 19th century. Similarly, the Netherlands displayed very high literacy levels as only 15% of recruits were illiterate (not or only unsatisfactorily able to read and write) (Statistisches Bureau, 1862) in 1857–58. These few examples highlight the already high levels of literacy already reached in these countries.

In other European countries, in particular in Southern and Eastern European countries (Kaelble, 2013), literacy remained the standard indicator for education in many European countries until the middle of the 20th century and beyond. Illiteracy was unequally distributed not only between but also *within* these countries. Accordingly, we choose to use literacy as our indicator of human capital for the period under consideration.

We define literacy as:

$$Literacy_{it}^{j} = \frac{\sum_{i=10}^{N} rw_{ikt}^{j}}{\sum_{i=10}^{N} n_{ikt}^{j}}, \tag{1}$$

where rw is the ability to read and write, i is the age in years, N is the total number of years, k is a region, the gender is denoted by $j = m, f$ (i.e. male or female), n is the population of individuals of age 10 and plus, and t is a point in time. The age definition (10+ years) is the standard contemporary definition for literacy (see Diebolt and Hippe, 2016a,b; Hippe, 2012, for more information). We measure literacy for both men and women separately.

Over the course of the second half of the 20th century, illiteracy was vastly eliminated throughout Europe. Another indicator of human capital is therefore required for more recent times. Hence, we use the standardly used proxy in the existing literature, which also allows us to account for gender differences: the rate of educational attainment.[7]

Educational attainment is measured as follows:

$$Educational\ attainment_{it}^{j} = \left[\frac{\sum_{i=25}^{64} st_{ikt}^{j}}{\sum_{i=25}^{64} n_{ikt}^{j}}\right], \tag{2}$$

where st is the number of individuals achieving at least upper secondary or tertiary education as their highest education level, i is the age in years, N is the total number of years, k denotes the region, j is the gender, and n is the population of individuals of age 25 and 64.

This indicator provides the attainment rates of individuals who are not early leavers (i.e. who pursued their education after lower secondary school). This corresponds to the reverse 'early leavers' definition currently used by the European Commission (see European Commission, 2015). We decided to adopt this definition because literacy – our proxy for 1900 and 1960 – only accounts for very basic levels of human capital.[8] For example, using tertiary

[7] Ideally, we would prefer to use variables such as PISA scores. Unfortunately, PISA scores are not available at the regional level for most European countries.

[8] Being illiterate in 1900 and 1960 was certainly a disadvantage to labor market prospects in a range of professions. Similarly, today, "[t]he EU regards upper secondary education attainment as a prerequisite for better labour market integration, lowering chances of poverty and social exclusion, and at the same time setting a minimum guarantee for continued personal development and active citizenship" (European Commission, 2015, p. 32).

education rates would correspond to advanced human capital levels, which would not be appropriate for measuring basic levels of education. In addition, the realized scale (with observed rates between 0 and 1) is the same for literacy and this indicator of basic human capital.

To measure gender equality in human capital, we use the Gender Parity Index (GPI).[9] The GPI is commonly used by international organizations, such as UNESCO (e.g., UNESCO, 2011), and is also used by the scarce literature accounting for gender differences in the past (e.g., Becker and Woessmann, 2008).

The GPI is defined as follows:

$$GPI_{kt} = \frac{H_{ktf}}{H_{ktm}}, \quad (3)$$

where H is the human capital indicator (i.e. literacy or educational attainment rates), and f and m designate males and females, respectively.

3.2. Fertility variables

We use information on fertility from the large database provided by the European Fertility Project (Coale and Watkins, 1986), in particular for the years 1870, 1900, 1930 and 1960. The definition of fertility is unique to the project, as it takes as reference data the fertility of Hutterites, an ethno-religious group from North America. More specifically, the EFP considers that the maximum possible fertility one individual can reach during its life cycle should be similar to the level reached by the Hutterites. The Hutterites are indeed known for not controlling their fertility (and even for trying to have as many kids as possible as a religious duty). They had at that time the highest fertility known in the developed world. All European data were therefore constructed in correspondence to this maximum fertility level. The definition of fertility used by the project is relatively elaborated and this indicator is more accurate than some other historical proxies.[10] This enables us to avoid biases which would be involved when using some of the alternative and less stringent measures on the European level.

Accordingly, fertility, as defined by the EFP, can be written as follows:

$$Fertility_{it} = \frac{breal_{it}}{bmax_{it}}, \quad (4)$$

where $breal$ is the actual number of birth and $bmax$ is the maximum attainable number of births in a population. More specifically, we use the total fertility rate.

An additional aim of this paper is to confront current fertility rates to historical data in order to provide a long-run view on the evolution of regional fertility patterns. For this reason, we include recent data for 1992 and 2005 on regional fertility rates provided by Eurostat (2010).

3.3. Methodology

Our methodology follows a descriptive empirical approach. Thus, after providing a brief introductory overview on the European and worldwide fertility decline, we show the regional distribution of our relevant human capital and fertility variables, separately, using maps. Subsequently, we relate these variables to each other by using correlation matrices and scatter plots. Finally, we conduct a simple empirical analysis of the association between gender equality, human capital, and fertility.

The analysis is conducted at the regional level. A region is defined by the NUTS 2 (Nomenclature of Territorial Units for Statistics) classification elaborated by the European Union.[11] This classification enables us to account for constant time-invariant regional administrative units, which gives additional robustness to our analysis. To do so, historical and current administrative regions must be matched in the most accurate way possible. For many countries, this is not a major difficulty as the data are available at more disaggregated levels such as NUTS 3. Countries such as Spain, France, and Cisleithania (the Austrian part of Austria-Hungary) are particular examples for which data are more disaggregated and historical regions are identical to today's regions. Furthermore, the use of the NUTS classification enables us to easily match the data to various time periods (earlier or later) and to other data sources relying on this classification.

Table 1
Date of fertility decline in European countries.

Country	Date of decline in marital fertility by 10%
France	ca. 1800
Belgium	1882
Germany	1885
Hungary	ca. 1890
England and Wales	1892
Sweden	1892
Scotland	1894
Netherlands	1897
Denmark	1900
Norway	1904
Austria	1908
Finland	1910
Italy	1911
Bulgaria	1912
Spain	1918
Ireland	1929

Source: Knodel and van de Walle (1979).

4. The fertility transition

In Europe (and the European offshoots), the process of modernization started earlier than in any other region of the world. Europe was a predecessor of a process that took place later in the other parts of the world. Interestingly, the onset of the fertility decline occurred in a rather small period of time in most European countries, i.e. between 1880 and 1910 in two out of three regions which are part of the table (Knodel and van de Walle, 1979, see Table 1).

Let us now come to the main aim of this paper and consider the regional level in Europe. To document the historical evolution of fertility rates in the European regions, we correlate historical fertility indices (in 1870, 1900, 1930 and 1960) with more recent fertility data (1992 and 2005).[12] Table 2 presents the correlation matrix of fertility indices over the 1870–2005 period. Fertility rates appear to be highly correlated throughout time. This is not a very surprising result as fertility rates are often reported as being 'hereditary', i.e. daughters often take similar fertility decisions as their mothers and

[9] For an overview on gender (in)-equality indices, see Subrahmanian (2005) and Perrin (2014).

[10] For example, Becker et al. (2010) use the child–woman-ratio as their indicator of fertility. Clearly, this indicator is arguably rough. The fertility indicators used by the EFP have nonetheless also been criticized for presenting drawbacks, see e.g., Brown and Guinnane (2007).

[11] We use a similar regional level for countries which are not part of the NUTS classification, such as Russia.

[12] We chose to use the year 2005, and not a more recent years, because this is the year for which data are available for a larger number of regions. The results of the year 1992 are less conclusive than the other years. This may be explained by the fact that fewer data were available for that specific year. The results for this year must, therefore, be considered with caution.

Table 2
Correlation matrix of fertility indices (NUTS 2), 1870–2005.

	2005	1992	1960	1930	1900	1870
2005	1.000					
1992	0.459 (0.000)	1.000				
1960	0.237 (0.000)	0.262 (0.001)	1.000			
1930	−0.280 (0.000)	−0.155 (0.055)	0.209 (0.000)	1.000		
1900	−0.577 (0.000)	−0.024 (0.772)	−0.027 (0.659)	0.818 (0.000)	1.000	
1870	−0.467 (0.000)	−0.014 (0.872)	−0.078 (0.209)	0.762 (0.000)	0.935 (0.000)	1.000

Source: own calculations, data provided by Coale and Watkins (1986) and Eurostat (2010).
Note: *p*-values in parenthesis.

can also be influenced by the socio-economic characteristics of the region in which they leave.

However, what is more interesting is the sign of the correlation. The correlations between the most recent years are positive (i.e., 1960, 1992 and 2005); suggesting the existence of positive links between them (see the upper part of column 1 and 2). In contrast, past and recent years are strongly and significantly negatively correlated (see lower part of column 1 and 2). The correlations suggest the tendency of a complete reversal of fertility rates in the European regions by the middle of the 20th century. The regions which had on average the highest fertility rates around the turn of the 19th century have the tendency to exhibit the lowest rates today, and conversely. The highest negative correlation is found between the years 2005 and 1900. This indicates that an important change occurred at a certain point in time during the 20th century; the century which marked the start of the demographic transition in many European countries. This period also coincides with the end of the 'unique' European marriage pattern. As Foreman-Peck (2011) notes, Hajnal was already aware that the "First World War reinforced a shortage of marriageable men already established by emigration, thereby temporarily increasing female celibacy and age at marriage. Unsurprisingly the turbulent half century after 1914 that saw such enormous social, economic and ideological changes in Europe, also eventually eliminated the 'European Marriage Pattern"' (Foreman-Peck, 2011, p. 293).

Thus, the initially negative correlation switches to a positive one between 1930 and 1960 (see column 1), suggesting the entry into a new regional fertility pattern in Europe. What happened over this period? For example, the proportion of women, especially married women, engaged in paid work increased drastically during WWII. In the US, for instance, one out of two women working in 1950 has been employed in 1940, as shown by Goldin (1991). Women performed a variety of work and activities. The war certainly contributed to modifying deeply the role of women in society, as well as gender relations. Additional factors explain the strong rise in the female labor force participation observed after WWII. Increased education, contraception, Cultural Revolution, among others, are some of the main determinants, which have directly or indirectly affected women's economic role during the second half of the 20th century.

We use scatter plots to provide a deeper insight into the relationships between past and present fertility (see Fig. 1a–e). Fig. 1a presents a clear negative relationship, although a few regions appear as outliers. The most obvious outlier is the South-East region (RO22) of Romania, located on the Black Sea and at the border with Ukraine, Moldova, and Bulgaria.[13] In contrast, in 2005, the highest fertility is found in Pohjois-Suomi (FI1A), the region located in the most Northern area of Finland.

In 1900 (Fig. 1b), apart from the already mentioned regions, Berlin (DE30) had one of the lowest fertility, which remained among the lowest in 2005. The large urban population characterizing this region may contribute to explain this particularly low fertility rate.[14] The former Yugoslav Republic of Macedonia (FYROM) (MK00) appears as a clear outlier in 1930 by reaching more than 5 children per woman, as highlighted by Fig. 1c. Similarly to Berlin, the capital of Austria, Vienna (AT13), presents the lowest fertility rate in 1930. The negative relationship linking current and historical fertility rates appears clearly in Fig. 1a–c.

The negative relationship, systematically observed so far, between past and present fertility suddenly reverses to become positive around the 1960s, as observed in Fig. 1d. Looking at each figure separately, Iceland (IS00) appears to experience a gradual and fast increase in fertility over the 1870–1960 period, until it reaches similar fertility rates as FYROM (at the top of fertility levels in Europe) in 1960. In contrast, the capital of Czechoslovakia, Prague (CZ01), overcomes Vienna's leadership by displaying the lowest fertility rate in 1960.

Finally, Fig. 1e displays a strong positive relationship between the two most recent years, 1992 and 2005. As described in Section 4, the fertility indices used for the years 1992 and 2005 are different from those used for the earlier years (i.e. 1870, 1900, 1930, and 1960). This difference explains the specific shape observed in Fig. 1e. Although the sample is reduced on that period, Fig. 1e highlights Cyprus (CY00), a new country entering the sample, as the most fertile European region (in 1992), while the Northern Spanish regions of Asturias (ES12) and País Vasco (ES21) appear at the bottom of the list, by exhibiting the lowest fertility rates.

In consequence, fertility rates evolved significantly over the 1870–2005 period. Interestingly, we observe a change in the relative fertility rates of some countries (or regions) but also a reversal of the overall European fertility pattern, i.e. many regions with historically low fertility exhibit nowadays high fertility rates, and vice versa.[15]

5. Gender equality in human capital in the past

5.1. Regional distribution of the GPI in the past

After having considered the evolution of national and regional fertility changes in the previous sections, let us now turn to the issue of gender inequality in human capital. More specifically, this section highlights the general evolution of gender equality in human capital in the European regions between 1900 and 1960. According to UNESCO (2011), gender parity in human capital is attained when the Gender Parity Index (GPI) is within the limits

[13] This region exhibits, by far, the highest fertility rates of our sample in 1870. The region was characterized by a large population of Turkish people, known for having on average a greater number of children.

[14] Urban areas are usually more advanced and exhibit higher education levels and lower fertility rates. The cost of living being higher in urban areas, the cost of having children is also consequently higher.

[15] Always considering the sample, without Eastern European countries, such as Russia.

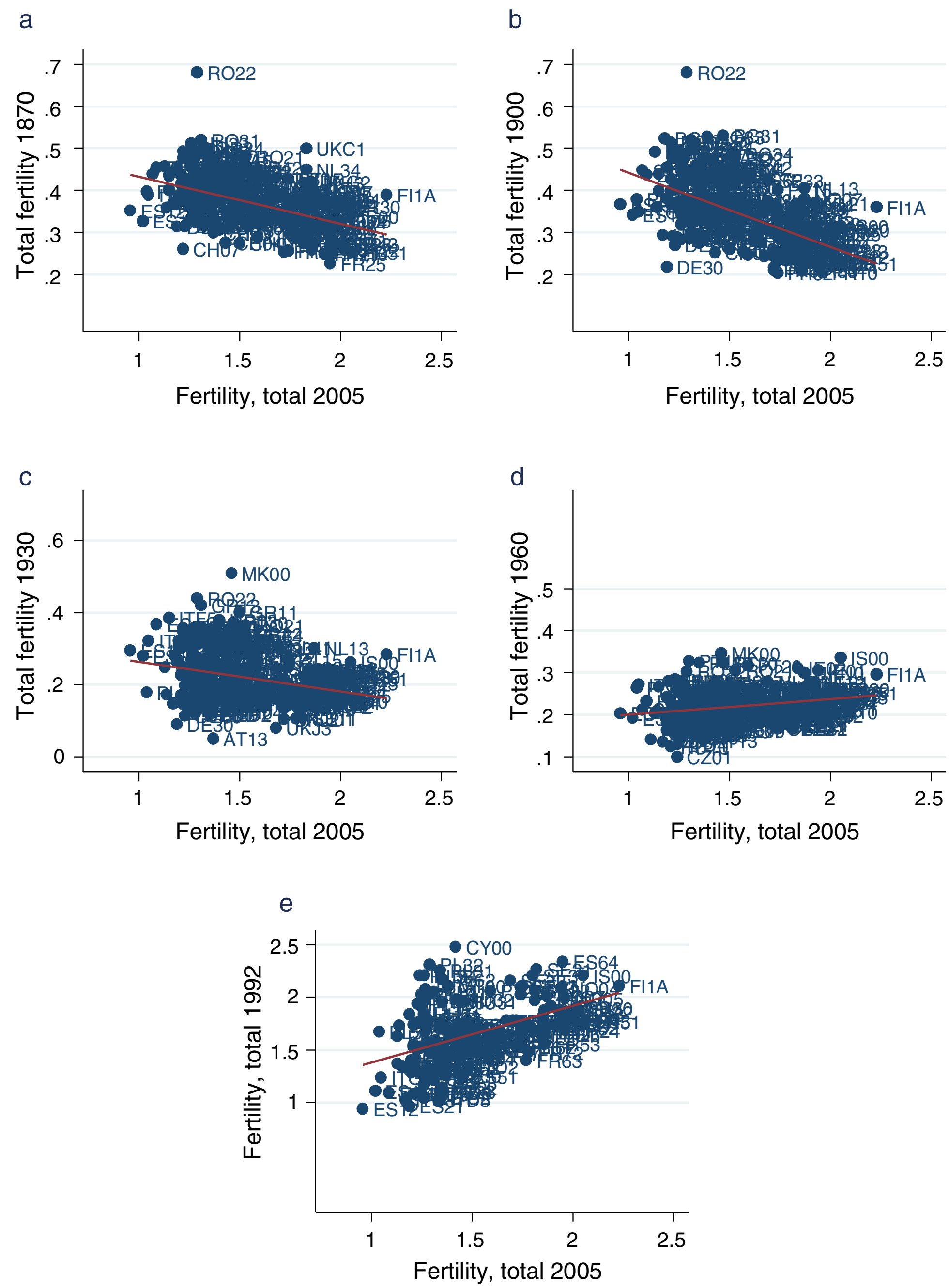

Fig. 1. Comparison of current versus historical fertility. *Note*: Data refer to EU27, EFTA members, and candidate countries to the EU. Fertility data refer to the number of children per woman for all women aged 10 and more. No data available for Belgium and Denmark in 2005, neither for Denmark, Germany, Ireland, Latvia, Netherlands, Romania, Slovakia, Liechtenstein, United Kingdom, Croatia, and FYROM in 1992.
Source: Coale and Watkins (1986), Eurostat (2010).

of 0.97 and 1.03. Accordingly, Figs. 2 and 3 – which present the GPI in human capital in 1900 and 1960, respectively – account for 7 intervals, one of which relies on the limits of gender parity as defined by the UNESCO (in red on the map). As specified previously, we decide to focus more specifically on the years 1900 and 1960 to be in accordance with the data on regional fertility. We would prefer to have even earlier data for 1870, however data availability limits our analysis to 1900. In particular, a large part of the European regions is constituted by the Russian Empire which had its first complete census taken only in 1897. Still, this allows us to see the changes that took place over a 60 year period.

For the year 1900, we observe a clear core-periphery pattern (Fig. 2).[16] According to these data, gender parity in literacy was already attained in Ireland, Northern France, in most parts of today's Austria, but also in Czech Republic, Slovenia, and Latvia. One could draw an imaginary circular arc line going from the South of Ireland to Slovenia (through France), and then moving progressively northwards toward Latvia. This impression is likely to be confirmed and even reinforced by the introduction of the countries for which

[16] Although some countries are absent from our sample, as explained in the data section.

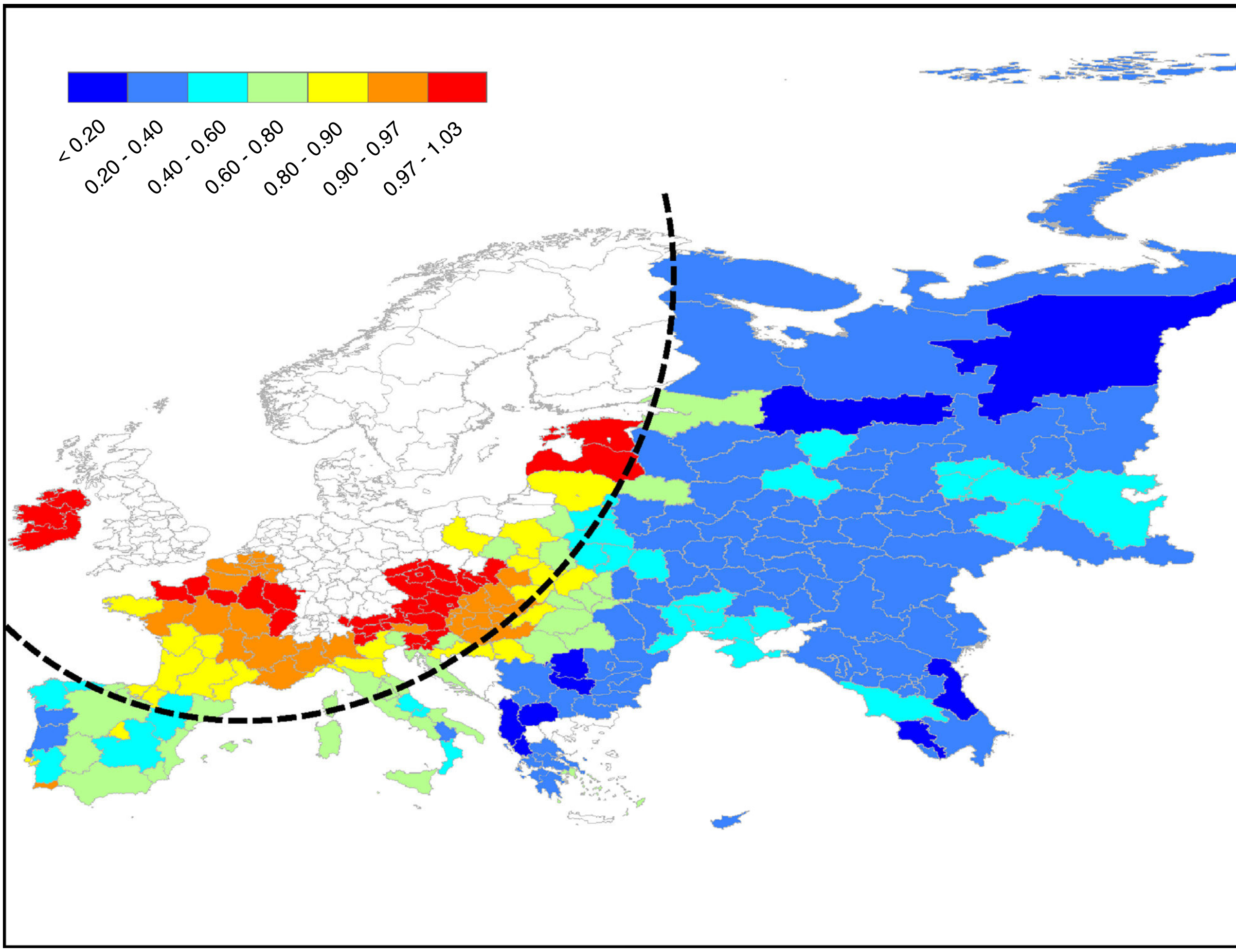

Fig. 2. Gender Parity Index for literacy, 1900.

Source: See Appendix.

literacy data at the regional level are not available, such as the UK (to the exception of Northern Ireland), the Netherlands, Germany, Switzerland, and the Scandinavian countries. This imaginary line would therefore cut Europe in two geographic areas – characterized by high gender parity on the inner side and significant gender differences on the outer side.

We observe that the further one moves away from the regions characterized by high gender parity (i.e. from the imaginary line), the greater the downward trend is. This pattern of gradual spread also appears clearly within France, where the Northeastern regions, highly industrialized, exhibit gender parity. Gender parity is reduced as we move away toward the Southwestern part of the country. The same tendency of gradual spread can be found within Italy, from regions displaying higher gender parity indices in the North to large gender differences in the South. Larger GPI values are sometimes observed in the capital regions, such as Madrid or Lisbon.

The regions presenting the lowest GPI values are located in Eastern Europe. A closer look at the Balkans reveals that the lowest GPIs are located in two clusters: one located in the Southwestern part of Romanian and its neighboring North-Western regions of Bulgaria and another one in the greater Albanian region comprising Albania, Macedonia and the North-Western regions of Greece. In the Russian Empire, a few regions located in the South and in the North also display very low GPIs. The highest values appear in the Baltic countries and in Poland. The regions surrounding the Black Sea, but also St. Petersburg and Moscow, present higher GPIs than the rest of Eastern Europe.

A similar pattern emerges on the geographical distribution of the GPI in literacy in 1960 (Fig. 3). The circular line of gender parity seems to have slightly moved toward the Southeastern part of Europe. The map, however, presents a few outliers. This is notably the case of the Italian Sicily, which has moved toward gender parity (much faster than other regions of the Southeastern side of the arc line). This is also true for the central region of Romania (characterized by an important Hungarian speaking population) and, as in 1900, for Portuguese Algarve. But the most striking change occurred in Armenia, which moved from the bottom of the list, i.e. the region/country with the lowest GPI, to attain gender parity in 1960. No other region has made such a transformation during this time span. The regions now lacking behind are located in the Balkans (in particular Kosovo, which exhibits the lowest GPI in our sample) and in Portugal (in the center and north).

Section 5 has explored the geographical distribution of gender equality in human in the past. It shows the existence of a clear core-periphery pattern with an imaginary line cutting Europe in two geographic areas – characterized by high gender parity on the inner side and significant gender differences on the outer side.

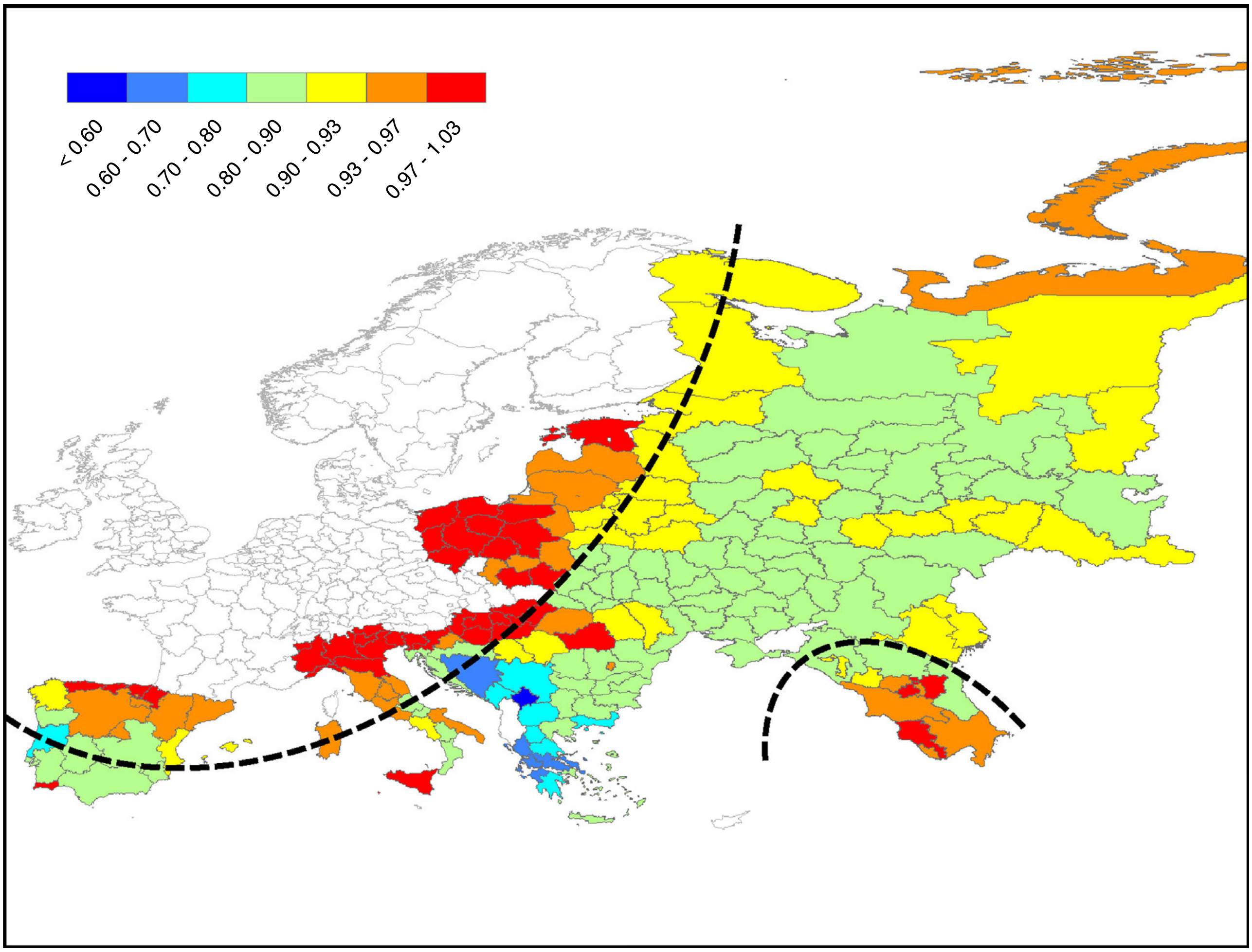

Fig. 3. Gender Parity Index for literacy, 1960. *Note*: Only national averages are available for Belarus and Ukraine. Data are not available for the Opolskie region in Poland. *Source: See Appendix.*

6. Relationship between gender equality, human capital and fertility

6.1. GPI and human capital

As a further exercise, we investigate how the GPI is related to the overall human capital value. One may expect a positive relationship linking human capital to the gender parity index. Indeed, men and women's investments in human capital are much more likely to be similar in regions presenting high average literacy rates, i.e. rates close to 1, than in regions displaying low average literacy rates. Investments in human capital appear to be a gradual process. Historical data show that large efforts were first placed in men's education, before any effort could be placed in women's education (see, for instance, the case of France during the 19th century in Perrin (2013), Chapter 3).

Evidence provided by Fig. 4 seems to validate the existence of a positive relationship between human capital and GPI across European regions in 1900. The scatter plot presents a strong positive relationship. Three regions located in the upper-left corner of the scatter plot, however, appear as outliers: the Portuguese islands in the Atlantic Ocean (i.e. Acores and Madeira), and the Algarve (the most Southern region in Portugal). All three of them had high levels of gender equality, but at quite low literacy levels. These regions were not governed by the central Portuguese authorities during that period (but had their own autonomy), which might be at least one factor influencing this special result. At the same time, they were rather rural and remote regions with low economic performance. Furthermore, Fig. 4 categorizes regions according to their adherence to the Western or to the Eastern fertility pattern,

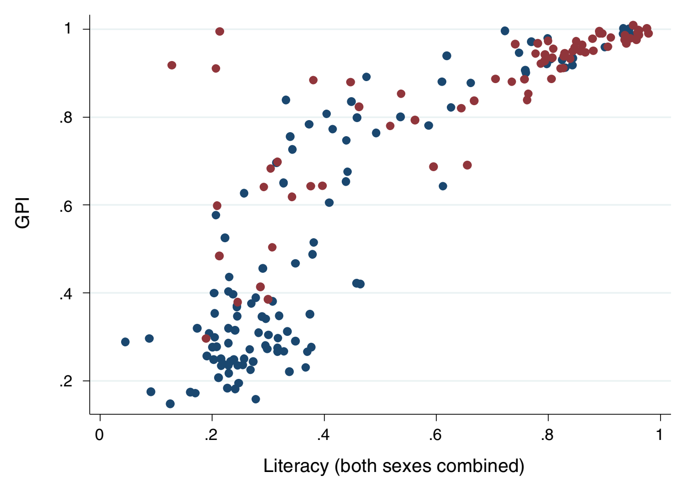

Fig. 4. Relationship between human capital and GPI, 1900. *Note*: Regions in the East of the Hajnal line are depicted in blue; regions from the West side appear in red. *Sources: See Appendix.*

as designed by Hajnal (1965).[17] Overall, the regions located on the West side of the Hajnal line present higher levels of human capital as well as higher levels of gender equality.[18] However, there are important regional variations and there are some outliers to this pattern.

During the first half of the 20th century, literacy rates progressed (almost) everywhere in Europe (where reading and writing ability was not yet universal). Thus, while the literacy rates were fluctuating between 0 and 1 in 1900, they ranged only between 0.5 and 1 in 1960. The countries in which the population was still not completely literate in 1960 (i.e. belonging to the interval [0.55; 1]) present a less clear literacy-GPI relationship, as evidenced by the point cloud in Fig. 5. Dividing our sample so as to isolate the point cloud from the rest of the regions, it appears that the bulk of regions forming the round-shaped heap is composed of the Soviet Union and Portugal (Fig. 5b). In the Soviet Union, this surprising pattern may perhaps be explained, for example, by the stress laid on gender equality, particularly in education, under the Communist regime.[19] Enhancing educational opportunities for women and the overall progress in gender equality during the 20th century may contribute to explain this specificity of the Soviet Union regions.[20] Another explanation must be found for the case of Portugal. Portuguese exceptionalism is again, as in 1900, mainly driven by the unique pattern found in its Island regions (e.g. Madeira) and the Algarve. Another alternative possibility is the availability and enforcement of compulsory schooling as the driving force behind this pattern. However, given the data limits of this paper, we cannot clarify the reasons; it is up for future research to investigate the reasons for this peculiarity more closely.

Finally, the remaining regions form a similar pattern to that observed in 1900 (Fig. 5c).

6.2. GPI and fertility

Finally, we explore the relationship between the fertility and the GPI at the regional level in 1900 and 1960. Fig. 6a shows the existence of a strong negative relationship between fertility and the GPI in 1900 (with a correlation coefficient of −0.74). However, the relationship is not totally clear cut. A number of regions, mainly located on the West side of the Hajnal line, present relatively high fertility and display high GPI. By contrast, many regions located on the East side of the Hajnal line exhibit both low fertility and GPI rates.

This relationship remains negative in 1960, but to a much weaker extent, as can be seen in Fig. 6b. The decline of the coefficient of correlation and its significant is likely to be explained by the GPI indicator used (i.e. female-to-male literacy), which does not allow for sufficient variations (as men and women's literacy rates were mostly already both very high in 1960) and by the timing, as the demographic transition was already well-advanced in Europe at that time. As was also pointed out by Hajnal, the Hajnal line is of no real use anymore at this point in time, as the distinction and the WEMP has vanished by 1960. One of the most obvious outliers in the figure is the Kosovo in the very lower right corner of the figure, characterized by very high fertility (the highest fertility rate observed in 1960) and a very low GPI. Excluding this outlier would give us an even weaker negative relationship (not shown). A similar exercise conducted between fertility and female human capital presents comparable negative associations (see Fig. B.1. in Appendix).

The confrontation of the gender equality data, first, to literacy rates suggests the existence of an intricate relationship between these variables in 1900. The regions characterized by high gender equality present high endowment in human capital and exhibit low fertility rates. Inversely, regions characterized by low gender equality display low endowment in human capital and high fertility rates. While the positive relationship between human capital and GPI remains strong in 1960, the association between the GPI and fertility becomes significantly weaker.

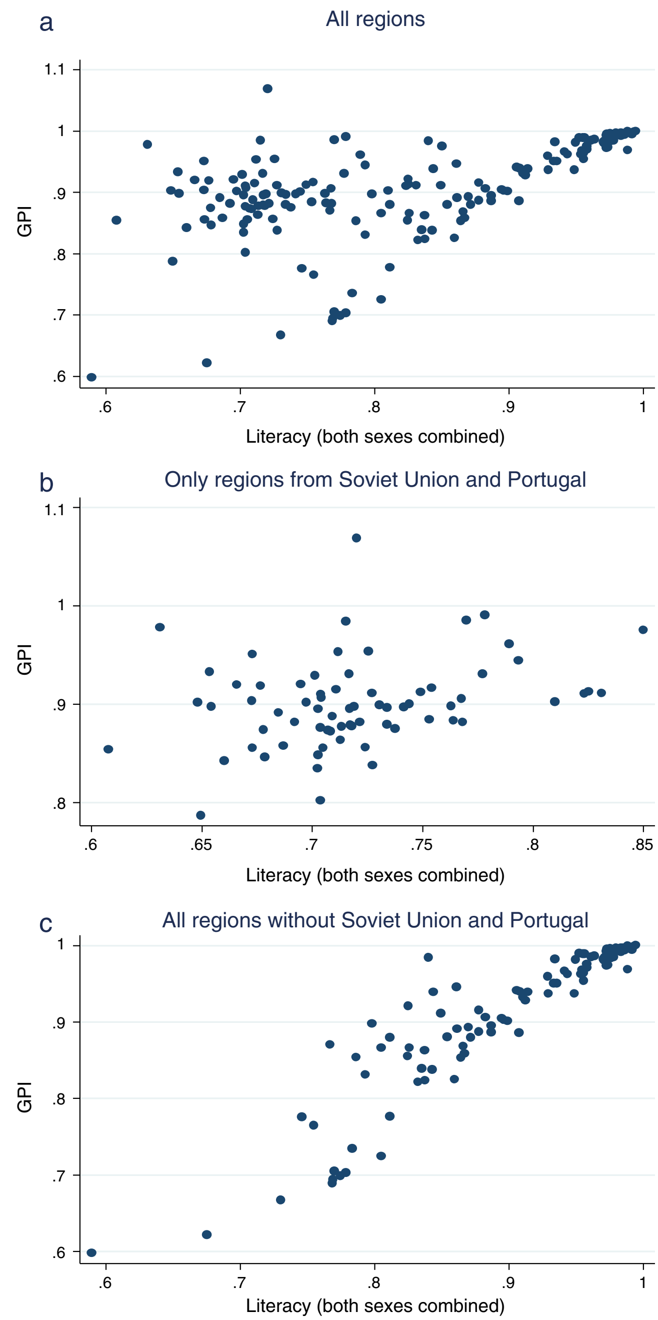

Fig. 5. Relationship between human capital and GPI, 1960.
Sources: See Appendix.

[17] See the appendix for more information on the geographical definition of the Hajnal line.
[18] Of course, there may be other factors which are the underlying causes of the observed relationship. We cannot control for those here.
[19] In the Soviet period, education was highly centralized and featured total access to primary and middle education for all citizens. The Soviet leaders put effort to achieve the optimum educational investment in each youth (Ross, 1959), boys and girls alike.
[20] Other Eastern European Communist countries, such as Romania, however, present a similar positive relationship (linking GPI and human capital) to the classical pattern observed in the rest of European regions.

Table 3
Ordinary least-squares estimates – 1900.

Dependent variable	Gender parity index				
	(1)	(2)	(3)	(4)	(5)
Human capital	0.89*** (0.000)	0.69*** (0.000)	0.68*** (0.000)	0.71*** (0.000)	0.71*** (0.000)
Fertility		−0.69*** (0.000)	−0.42*** (0.001)	−0.39*** (0.001)	−0.36** (0.015)
Marital status			−0.37*** (0.006)	−0.31** (0.012)	−0.28** (0.044)
Population density				0.00 (0.754)	−0.00 (0.754)
Urbanization					0.07 (0.490)
Constant	0.21*** (0.000)	0.59*** (0.000)	0.70*** (0.000)	0.63*** (0.000)	0.59*** (0.000)
Observations	199	166	166	165	165
R^2	0.71	0.74	0.75	0.77	0.77

Note: This table reports OLS estimates of our variables of interest. Robust *p*-values in brackets.
*10% of significant level.
** 5% of significant level.
*** 1% of significant level.

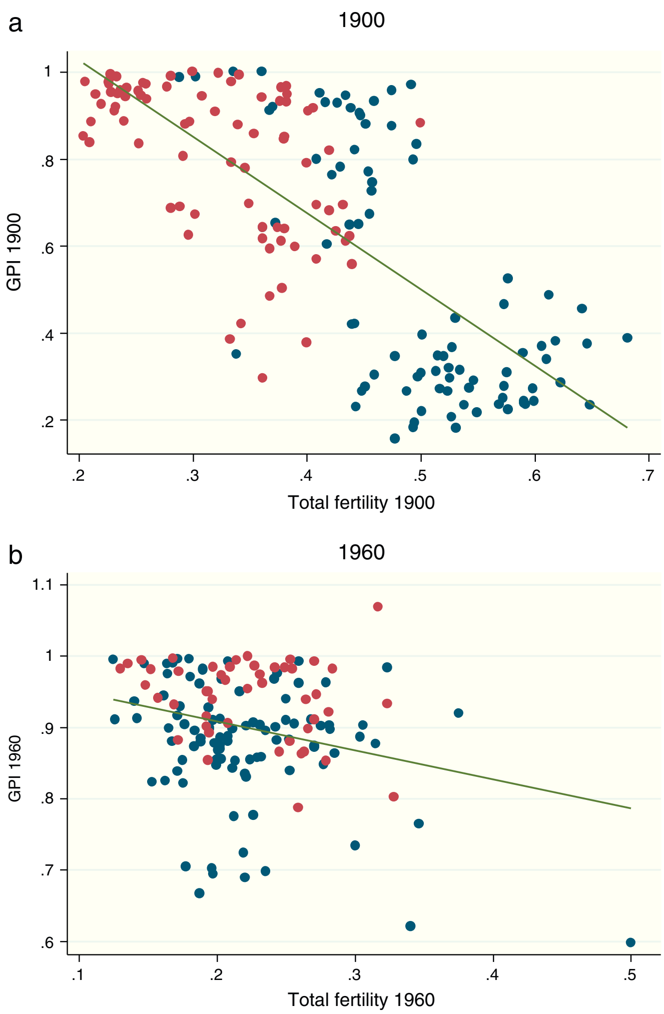

Fig. 6. Relationship between total fertility and GPI, 1900 and 1960. *Note:* The red and blue dots correspond to the regions located on the West and on the East side of the Hajnal line, respectively.
Source: See Appendix.

6.3. *Empirical investigation of the GPI, human capital and fertility relationship*

In order to better understand the correlations highlighted in Section 6, we estimate how gender equality in human capital is affected by human capital and fertility in 1900 and 1960.

We establish the descriptive association between our variables by using the following OLS regression model:

$$GPI_i = \beta_0 + \beta_1 human\,capital_i + \beta_2 fertility_i + \beta_3 marital\,status_i + \beta_4 density_i + \beta_5 urbanization_i + \varepsilon_i \quad (5)$$

where i stands for the region and ε_i for the error term. GPI_i is the gender parity index in human capital. $human\,capital_i$ refers to the ability to read and to write of individuals living in region *i*. $fertility_i$ is the fertility rate of region *i* (as measured by the EFP), $marital\,status_i$ is measured by the age at marriage.[21] $density_i$ is the number of individuals in region *i* per square kilometer. Finally, $urbanization_i$ captures the degree of urbanization (i.e. the share of people living in urban areas).

Table 3 presents OLS estimates of Eq. (5). The dependent variable is the gender parity index in human capital in 1900. The bivariate regression (column 1) shows that the GPI and human capital are significantly positively correlated. Column 2 introduces fertility. Fertility is strongly negatively associated with the GPI. Regions associated with high fertility rates display higher gender inequalities. In the subsequent columns, we progressively introduce control variables. The marital status is negatively associated with the GPI. The younger individuals get married; the lower is the gender equality. Controlling for more densely populated and urban environments increases the explanatory power of the model but is not significantly associated with the GPI.

The human capital coefficient remains statistically significant throughout the different specifications, so do the fertility coefficients. Considering the richest specification (column 5), an increase in human capital by 1% points is associated with an increase in gender equality by 0.71 and by a decrease in fertility by 0.36.

We run a similar exercise using panel data for the year 1960. Table 4 reports the estimates obtained from Eq. (5) using various specifications. Similarly to Table 3, we successively introduce control variables: marital status (column 3), population density (column 4), and Urbanization (column 5).

Although the magnitude of the coefficients decreased substantially in comparison to the situation in 1900, the GPI and human capital remain significantly positively correlated throughout the different specifications. Fertility, however, becomes non-statistically significant. Fertility does not seem to explain anymore the variations in the gender parity index observed across regions.

Regions characterized by young age at marriage exhibit lower gender equality, confirming our previous results. Nonetheless, controlling for the level of urbanization (column 5), the marital status

[21] The younger the age at marriage, the higher the marital status.

Table 4
Ordinary least-squares estimates – 1960.

Dependent variable	Gender parity index				
	(1)	(2)	(3)	(4)	(5)
Human capital	0.39*** (0.000)	0.42*** (0.000)	0.46*** (0.000)	0.49*** (0.000)	0.41*** (0.000)
Fertility		−0.15 (0.335)	−0.12 (0.397)	−0.18 (0.156)	−0.08 (0.529)
Marital status			−0.15** (0.021)	−0.17*** (0.009)	−0.03 (0.806)
Population density				−0.00 (0.181)	−0.00 (0.829)
Urbanization					0.25** (0.030)
Constant	0.59*** (0.000)	0.59*** (0.000)	0.64*** (0.000)	0.64*** (0.000)	0.48*** (0.000)
Observations	169	145	145	144	144
R^2	0.30	0.36	0.38	0.42	0.49

Note: This table reports OLS estimates of our variables of interest. Robust *p*-values in brackets.
*10% of significant level.
** 5% of significant level.
*** 1% of significant level.

coefficient lose its significance. Henceforth, in 1960, more urbanized regions displayed greater gender equality.

This empirical analysis supports the existence of a positive relationship of the GPI and the level of human capital. However, although the analysis confirms the existence of a negative relationship with fertility in 1900, this finding is not found when using the regional-level data in 1960. In this way, the regressions confirm overall the tendencies in the visual graphs obtained in the previous section.

7. Concluding remarks

This paper provides a descriptive analysis of the relationship between fertility, human capital, and gender equality in human capital in the European regions in 1900 and 1960. According to the Unified Growth Theory (Galor and Weil, 1996; Galor, 2005, 2012), human capital was a crucial factor for the demographic transition. Gender equality in human capital, in particular, has been shown to have played a crucial role in the transition to Modern Growth, characterized by sustained economic growth and fertility transition (Diebolt and Perrin, 2013a,b).

To bring together gender equality and human capital indicators, we extend the evidence presented by Diebolt and Hippe's (2016a) new and large database on human capital by providing separate regional literacy values for men and women in 1900 and 1960. These data have been combined with the European Fertility Project database (EFP; Coale and Watkins, 1986). Although a large literature discusses the results of the EFP, no study has specifically investigated a similar scope on gender equality in human capital in a European regional perspective. Besides providing evidence on the geographical distribution of gender equality in human capital in the past, we explore the underlying relationships linking gender equality with human capital and fertility.

We find a reversal of both educational fortunes and regional fertility rates in the 20th century. While men had higher basic human capital than women in 1900 and 1960, women have nowadays mostly overtaken and surpassed men (when considering educational attainment; see e.g. Meschi and Scervini, 2014). In addition, we observe that the regions displaying the highest fertility in 1900 rates have the tendency to present, on average, the lowest rates today. In line with theoretical expectations, we also emphasize the existence of a negative relationship between gender equality in human capital and fertility in the past. Regression analyses for 1900 confirm these predictions, while the association is not significant anymore in 1960.

The gendered approach in human capital appears as a cornerstone of our understanding of fertility dynamics. The geographic distribution and evolution of gender discrepancies, in the past and the present, strengthens the view that gender matters and that gender equality in a European regional perspective deserves further attention and investigation. The descriptive analysis and findings presented in this paper call for more research on the long-run relationships linking human capital, gender equality, and fertility. It further requires integrating additional socio-demographic variables in a European regional perspective to deepen the analysis and enhance our understanding of historical and social processes.

Appendix A. Data sources and definitions

Data sources for regional literacy in 1900 and 1960: census data; original sources used by Diebolt and Hippe (2016a).

- Data for literacy in 1900

Country	Census year	Source
Albania	1918	Preliminary dataset "Albanische Volkszählung von 1918", entstanden an der Karl-Franzens-Universität Graz unter Mitarbeit von Helmut Eberhart, Karl Kaser, Siegfried Gruber, Gentiana Kera, Enriketa Papa-Pandelejmoni und finanziert durch Mittel des Österreichischen Fonds zur Förderung der wissenschaftlichen Forschung (FWF). Special thanks to Siegfried Gruber for providing the data.
Austria (Cisleithania)	1900	K. K. Statistische Central-Commission (1903). Oesterreichische Statistik. Ergebnisse der Volkszählung vom 31. December 1900, 2. Band, 2. Heft, Wien, Kaiserlich-Königliche Hof- und Staatsdruckerei.
Belgium	1900	Statistique de la Belgique (1903). Population. Recensement général. 31 décembre 1900, Bruxelles, Typographie-Lithographie A. Lesigne.
Bulgaria	1900	Principauté de la Bulgarie (1906). Résultats généraux du recensement de la population dans la principauté de Bulgarie au 31 décembre 1900, 1-ère livraison, Sophia: Imprimerie "Gabrovo".
Cyprus	1911	Mavrogordato, A. (1912). Cyprus. Report and general abstracts of the census of 1911, London: Waterlow & Sons Ltd.
France	1906	Statistique Générale de la France (1908). Résultats statistiques du recensement général de la population effectué le 6 mars 1906, Paris.
Greece	1907	Royaume de Grèce (1909). Résultats statistiques du recensement général de la population effectué le 27 octobre 1907, Tome I, Athènes: Imprimerie nationale.
Hungary (Transleithania)	1900	Magyar Statisztikai Közlemények (1907). A magyar szent korona országainak 1900. Evi. Népszámlálása. Harmadik rész. A népesség részletes leirása. Budapest: Pesti Könyvnyomda-Részvénytársaság.
Ireland	1901	Census of Ireland, 1901 (1902). Part II. General Report, Dublin: Brown & Nolan Ltd.
Italy	1901	Ministero di agricultura, industria e commercio (1907). Anuario statistico italiano 1905–1907, Fascicolo Primo, Roma: G. Bertero e C.
Portugal	1900	Ministerio dos negocios da fazenda (1906). Censo Da Populacao Do Reino De Portugal No 1. De Dezembro De 1900, Vol. II, Lisboa: A Editora.
Romania	1899	Royaume de Roumanie (1905). Résultats définitifs du dénombrement de la population (décembre 1899), Bucarest, Eminesco.

Country	Census year	Source
Russian Empire	1897	издание центрального статистичесКого Комитета министерства внутренних (1899–1905). первая всеобщая. перепись населения, российсКой империи, 1897 г., с.-петербург. Various tombs.
Serbia	1906	Direction de la Statistique d'Etat du Royaume de Serbie (1908). Annuaire statistique du Royaume de Serbie, Onzième Tome, Belgrade, Imprimerie de l'Etat du Royaume de Serbie.
Spain	1900	Dirección general del Instituto geográfico y estadístico (1903). Censo de la Poblacion de Espana, según el Empadronamiento hecho en la Península é Islas adyacentes en 31 de diciembre de 1900, Tomo II and Tomo III, Madrid: Imprenta de la Dirección general del Instituto geográfico y estadístico.

Age definitions are as follows: Italy = 6+; Austria; Spain = 8+; Albania, Belgium, Bulgaria, France, Greece, Ireland, Portugal, Russian Empire = 10+; Hungary, Romania, Serbia = 11+; Cyprus = 15+. Age definitions are either directly given in the publication or have been linearly estimated from available age definitions in order to be as close as possible to the standard definition of ages above 10 years. The various age definitions as calculated here may possibly not significantly affect the final results. For example, the percentage change of using a 5+ instead of a 10+ definition is below 1% in Ireland.

- Data for literacy in 1960

Country	Census year	Source
Bulgaria	1956	централно статистичесКо управление при министерсКия свет (1960). преброяване на населението в народна републиКа българия на 1. XII. 1956 година, общи резултати, Книга II, софия: дърЖавно издателство “науКа и изКуство”.
Greece	1961	Royaume de Grèce (1968). Résultats du recensement de la population et des habitations effectué le 19 mars 1961, Vol. III, Athènes: Office nationale de Statistique.
Hungary	1960	Központi Statisztikai Hivatal (1962). 1960. Évi népszámlálás, Budapest, Allami Nyomda.
Italy	1961	Istat (2012). Serie Storiche, Tavola 7.1.1, online, last accessed 3 August 2012, http://seriestoriche.istat.it/fileadmin/allegati/Istruzione/tavole/Tavola_7.1.1.xls.
Poland	1960	Glówny Urzad Statystyczny (1960). Biuletyn statystyczny. Spis powszechny z dnia 6 grudnia 1960 r., Ludnosc. Gospodarstwa domowe, Wyniki ostateczne, Seria "L", various issues, Warszawa.
Portugal	1960	Instituto nacional de Estatística (1960). X Recenseamento Geral Da Populacao, Tomo III, Lisboa: Sociedade Tipográfica.
Romania	1956	Republica Populara Romîna (1961). Recensamîntul Populatiei din 21 Februarie 1956, Rezultate Generale, Bucuresti, Directia Centrala de Statistica.
Spain	1960	Instituto nacional de Estadística (1969). Censo de la Poblacion y de la Viviendas de Espana, según la Inscripción realizada el 31 de diciembre de 1960, Tomo III, Madrid: I.N.E. Artes graficas.
USSR	1959	Demoscope (2012). Всесоюзная перепись населения 1959 года. Таблица 7. Распределение населения по возрасту и уровню образования. РГАЭ. Ф.1562 Оп. 336 Д. 1591–1594, online, last accessed 8 August 2012, http://demoscope.ru/weekly/ssp/rus_edu_59.php. Demoscope (2012). Всесоюзная перепись населения 1959 года. Таблица 2,5. Распределение всего населения и состоящих в браКе по полу и возрасту. РГАЭ. Ф.1562 Оп. 336 Д. 1535–1548. РоссийсКая Государственная библиотеКа, отдел “литературы ограниченного пользования”, online, last accessed 8 August 2012, http://demoscope.ru/weekly/ssp/sng_mar_59_r.php.
Yugoslavia	1961	Statisticni urad Republike Slovenije (2012). Popis prebivalstva 1961, Prebivalstvo, staro 10 let ali več, po spolu, starosti in pismenosti, online, last accessed 8 August 2012, http://www.stat.si/publikacije/popisi/1961/1961_2_40.pdf.

Note: Age definitions are as follows: Italy = 6+; Hungary, Poland, Portugal, Spain = 7+; Bulgaria, Romania = 8+; Greece, USSR, Yugoslavia = 10+. Age definitions are either directly given in the publication or have been linearly estimated from surrounding available age definitions to be as close as possible to the standard definition of ages above 10 years.

Geographical definition of the Hajnal line: no precise distinction of European regions located on the West and East side of the Hajnal line was provided in the literature. Thus, this line was approximated in the following way: countries belonging to the West side of the Hajnal line (following the WEMP) are Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK and all regions of Poland, Czech Republic and Austria except the following NUTS regions (belonging east of Hajnal line): AT11, AT12, AT13, AT31, AT32, CZ06, CZ07, CZ08, PL11, PL12, PL21, PL22, PL31, PL32, PL33, PL34, PL52. All NUTS codes in this article refer to NUTS 2006.

Data on urbanization and population density have been computed from the following sources:

Klein Goldewijk, K., A. Beusen, de Vos, M. and G. van Drecht (2011). The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years, Global Ecology and Biogeography, 20 (1): 73–86.

Klein Goldewijk, K., Beusen, A. and P. Janssen (2010). Long term dynamic modeling of global population and built-up area in a spatially explicit way, HYDE 3.1, The Holocene, 20 (4): 565–573.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ihe.2017.02.001.

References

Becker, G.S., 1960. An economic analysis of fertility. In: National Bureau of Economic Research. Demographic and Economic Change in Developed Countries. Columbia University Press, New York/London, pp. 209–240.

Becker, G.S., 1981. A Treatise on the Family. Harvard University Press, Cambridge.

Becker, G.S., Tomes, N., 1986. Human capital and the rise and fall of families. Labor Economics 4 (3–2), S1–S39.

Becker, G.S., Murphy, K.M., Tamura, R., 1990. Human capital, fertility and economic growth. Journal of Political Economy 98 (3–2), S12–S37.

Becker, G., Lewis, H.G., 1973. On the interaction between the quantity and quality of children. Journal of Political Economy 84 (2), S279–S288.

Becker, S.O., Woessmann, L., 2008. Luther and the girls: religious denomination and the female education gap in nineteenth-century Prussia. Scandinavian Journal of Economics 110 (4), 777–805.

Becker, S.O., Cinnirella, F., Woessmann, L., 2010. The trade-off between fertility and education, evidence from before the demographic transition. Journal of Economic Growth 15, 177–204.

Becker, S.O., Cinnirella, F., Woessmann, L., 2012. The effect of investment in children's education on fertility in 1816 Prussia. Cliometrica 6, 29–44.

Becker, S.O., Cinnirella, F., Woessmann, L., 2013. Does women's education affect fertility? Evidence from pre-demographic transition Prussia. European Review of Economic History 17, 24–44.

Bignon, V., García-Peñalosa, C., 2016. Protectionism and the Education-Fertility Trade-off in Late 19th Century France, Working Papers halshs-01264614, HAL.

Bleakley, H., Lange, F., 2009. Chronic disease burden and the interaction of education, fertility, and growth. Review of Economics and Statistics 91 (1), 52–65.

Brown, J.C., Guinnane, T.W., 2007. Regions and time in the European fertility transition: problems in the Princeton Project's statistical methodology. Economic History Review 60 (3), 574–595.

Caldwell, J.C., 1976. Toward a restatement of demographic transition theory. Population and Development Review 2 (3/4), 321–366.

Cleland, J., Rodriguez, G., 1988. The effect of parental education on marital fertility in developing countries. Population Studies 42 (3), 419–442.

Coale, A.J., Watkins, S.C. (Eds.), 1986. The Decline of Fertility in Europe: The Revised Proceedings of a Conference on the Princeton European Fertility Project. Princeton University Press, Princeton.

Cochrane, S.H., 1979. Fertility and Education: What Do We Really Know? Johns Hopkins University Press, Baltimore.

De la Croix, D., Perrin, F., 2016. French Fertility and Education Transition: Rational Choice vs. Cultural Diffusion, IRES Discussion Papers, No. 2016-7, 36 pp.

De Moor, T., van Zanden, J.L., 2010. Girl power: the European marriage pattern and labour markets in the North Sea in the late medieval and early modern period. Economic History Review 63 (1), 1–33.

Diebolt, C., Perrin, F., 2013a. From stagnation to sustained growth: the role of female empowerment. American Economic Review 103, 545–549.

Diebolt, C., Perrin, F., 2013b. From Stagnation to Sustained Growth: The Role of Female Empowerment, AFC Working Paper, WP2013-4.

Diebolt, C., Hippe, R., 2016a. Regional human capital inequality in Europe in the long run, 1850–2010, AFC Working Paper 2016-9.

Diebolt, C., Hippe, R., 2016b. Remoteness equals backwardness? Human capital and market access in the European regions: insights from the long run, AFC Working Paper 2016-8.

Diebolt, C., Menard, A.-R., Perrin, F., 2016. Behind the Fertility-Education Nexus: What Triggered the French Development Process?, BETA Working Paper, No. 2016-10.

Diebolt, C., Mishra, T., Perrin, F., 2015. Did Gender-Bias Matter in the Quantity-Quality Trade-off in 19th Century France, Lund Papers in Economic History No. 141.

Doepke, M., 2015. Gary Becker on the quantity and quality of children. Journal of Demographic Economics 81 (1), 59–66.

Easterlin, R.A., 1969. Towards a socio-economic theory of fertility: a survey of recent research on economic factors in American fertility. In: Behrman, S.J., et al. (Eds.), Fertility and Family Planning: A World View. University of Michigan Press, Ann Arbor.

Engelen, T., Puschmann, P., 2011. How unique is the Western European marriage pattern? A comparison of nuptiality in historical Europe and the contemporary Arab world. The History of the Family 16 (4), 387–400.

European Commission, 2015. Education and Training Monitor 2015, http://ec.europa.eu/education/library/publications/monitor15_en.pdf (accessed 14.07.16).

Eurostat, 2010. Fertility rates by age–NUTS level 2 regions, data file demo_r_frate2 [older version] (accessed 2012).

Foreman-Peck, J., 2011. The Western European marriage pattern and economic development. Explorations in Economic History 48, 292–309.

Galor, O., Weil, D.N., 1999. From Malthusian stagnation to modern growth. American Economic Review 89, 150–154.

Galor, O., Weil, D.N., 2000. Population, technology, and growth: from Malthusian stagnation to the demographic transition and beyond. American Economic Review 90, 806–828.

Galor, O., 2005. From Stagnation to growth: unified growth theory. In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. North Holland, Amsterdam, pp. 171–293.

Galor, O., 2012. The demographic transition: causes and consequences. Cliometrica 6, 1–28.

Galor, O., Weil, D.N., 1996. The gender gap, fertility, and growth. American Economic Review 86 (3), 374–387.

Goldin, C., 1991. The role of World War II in the rise of women's employment. American Economic Review 81 (4), 741–756.

Guinnane, T.W., 2008. The Historical Fertility Transition and Theories of Long-run Growth: A Guide for Economists. Yale University, Mimeo.

Hajnal, J., 1965. European marriage pattern in historical perspective. In: Glass, D.V., Eversley, D.E.C. (Eds.), Population in History. Aldine Publishing Company, Chicago.

Hanushek, E.A., 1992. The trade-off between child quantity and quality. Journal of Political Economy 100 (1), 84–117.

Hippe, R., 2012. How to measure human capital? The relationship between numeracy and literacy. Economies et sociétés 45 (8), 1527–1554.

Hippe, R., (PhD thesis) 2013. Human capital formation in Europe at the regional level – implications for economic growth. University of Tübingen & University of Strasbourg.

Hippe, R., Baten, J., 2012. Regional inequality in human capital formation in Europe, 1790–1880. Scandinavian Economic History Review 60 (3), 254–289.

Kaelble, H., 2013. A Social History of Europe, 1945–2000: Recovery and Transformation After Two World Wars. Berghahn Books.

Kirk, D., 1996. Demographic transition theory. Population Studies 50, 361–387.

Knodel, J., van de Walle, E., 1979. Lessons from the past. Policy implications of historical fertility studies. Population and Development Review 2 (2), 217–245.

Lagerlöf, N.P., 2003. Gender equality and long-run growth. Journal of Economic Growth 8 (4), 403–426.

Lee, R., 2015. Becker and the demographic transition. Journal of Demographic Economics 81 (1), 67–74.

Leibenstein, H., 1957. Economic Backwardness and Economic Growth. Wiley and Sons, New York.

Malthus, T., 1798/1960. An Essay on the Principle of Population, as it Affects the Future Improvement of Society. Random House, NewYork.

McKinsey Global Institute, 2016. The power of parity: how advancing women's equality can add $12 trillion to global growth, http://www.mckinsey.com/global-themes/employment-and-growth/how-advancing-womens-equality-can-add-12-trillion-to-global-growth (accessed 15.07.16).

Meschi, E., Scervini, F., 2014. A new dataset on educational inequality. Empirical Economics 47 (2), 695–716.

Murphy, T., 2015. Old habits die hard (sometimes). Can département heterogeneity tell us something about the French fertility decline? Journal of Economic Growth 20, 177–222.

Notestein, F., 1945. Population: the long view. In: Schultz, T. (Ed.), Food for the Worm. Chicago University Press, Chicago, pp. 36–57.

Perrin, F., 2014. On the Construction of a Historical Gender Gap Index. An Implementation on French Data, AFC Working Paper, WP2014-5, 28 pp.

Perrin, F., (PhD dissertation) 2013. Gender Equality and Economic Growth in the Long-Run. A Cliometric Analysis. University of Strasbourg and Sant'Anna School of Advanced Studies www.theses.fr/2013STRAB006.pdf.

Perrin, F., 2015. An exploration of the cliometric relationship between gender equality and economic growth. Economics and Business Review 1, 34–54.

Ross, L.W., 1959. Some aspects of soviet education. Journal of Teacher Education 11 (4), 539–552.

Schultz, T.W. (Ed.), 1974. Economics of the Family: Marriage, Children and Human Capital. Chicago University Press, Chicago.

Statistisches Bureau, 1862. Schweizerische Statistik, Bevölkerung. Eidgenössische Volkszählung vom 10. December 1860. I. Lieferung, Statistisches Bureau des eidg. Departement des Innern. Walder & Schiller, Bern.

Subrahmanian, R., 2005. Gender equality in education: Definitions and measurements. International Journal of Educational Development 25 (4), 395–407.

UNESCO, 2011. EFA Global Monitoring Report 2011, Paris.

Van de Walle, F., 1980. Education and the demographic transition in Switzerland. Population and Development Review 6 (3), 463–472.

Willis, R., 1973. A new approach to the economic theory of fertility behavior. Journal of Political Economy 81 (2), S14–S64.

Wrigley, W.A., Schofield, R.S., 1989. The Population History of England 1541–1871: A Reconstruction. Cambridge University Press, Cambridge.