The International Epidemiological Transition and the Education Gender Gap

Mariko J. Klasing   Petros Milionis

University of Groningen*

March 2018

Abstract

We explore the impact of the international epidemiological transition on educational attainment of males and females over the second half of the 20th century using an instrumental variables strategy that exploits pre-existing variation in mortality rates across infectious diseases and gender differences in the responsiveness of the immune system to vaccination. We document that health improvements associated with the transition led to larger gains in female life expectancy. These relative gains were associated with relative increases in female educational attainment and account for a large share of the reduction in the education gender gap that took place over this period.

Keywords: International Epidemiological Transition, Vaccination, Life Expectancy, Education, Gender Differences, Economic Development.

JEL Classification: I15, I24, J16, O11.

---

*Faculty of Economics & Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail Addresses: m.j.klasing@rug.nl, p.milionis@rug.nl.

1 Introduction

One of the greatest achievements of the world during the 20th century has been the progress made towards gender equality. Comparing the role of women in economic and social life in 2000 relative to that in 1900 reveals the remarkable changes that have happened in terms of women's legal status, political rights, access to the labor market, and other areas. Nowhere is this change more visible, though, than when looking at female educational attainment. In many countries of the world women nowadays outperform men in all levels of education, to the point that the relative under-performance of men is being considered an emerging problem (OECD, 2015).

The progress achieved toward the closing of the gender gap in education can be seen in Figure 1. The figure depicts for a broad sample of 146 countries the evolution of the female-to-male ratio of average years of schooling over the 20th century based on data from Barro and Lee (2013). In the beginning of the century this ratio fluctuated around 0.75, implying that females had on average only 3/4 of the years of schooling of males. Starting from around the 1950s, though, we see a clear upward trend in this ratio, which eventually resulted in the complete elimination of the female gap in education in most countries of the world.

[Insert Figure 1 around here]

Understanding the driving forces behind this remarkable transition has been the focus of a growing literature in economics. Early contributions by Goldin (1995), Galor and Weil (1996), as well as Goldin et al. (2006) have stressed the role of improved labor market opportunities for women. More recent work by Chiappori et al (2009), Fernandez and Wong (2011), and Reijnders (2018) have highlighted changes in marriages patterns as an important factor. Fernández, Fogli and Olivetti (2004), Beaman et al. (2012), Fernández (2013), and Hazan and Zoabi (2015) have emphasized the importance of changing norms and biases regarding the role of women in society. Greenwood et al. (2016) have stressed the role of the decline in the price of household durables, which freed women from household work. The literature has also explored the role of medical innovations such as the introduction of the birth control pill (Goldin and Katz, 2002) and improvements in maternal health (Albanesi and Olivetti, 2016) in raising female schooling. Overall, however, most of the existing contributions have focused on the experiences of developed countries and particularly on the case of the United States.

Yet, developed countries may not provide the most striking examples in terms of the evolution of the education gender gap. As Figure 2 demonstrates, the relative rise in female educational attainment over the second half of the 20th century has been much more rapid in developing countries, where many of the aforementioned factors have played a less important role. Countries such as Bolivia, China, Congo, Egypt, India, Indonesia

and Turkey, despite their big differences in terms of economic development, all witnessed a closing of the gender gap. With that in mind, in this paper we study the evolution of the education gender gap in a broad sample of countries and consider the role of general improvements in health as a driving force behind the observed transition. In this context, we demonstrate that the closing of the education gender gap can largely be attributed to differential health improvements between men and women.

[Insert Figure 2 around here]

Exploring the link between health improvements and educational attainment, although not central in the gender gap literature, is natural as health improvements have commonly been argued to promote educational attainment. Healthier individuals, who expect to live longer and more productive lives, are bound to have stronger incentives to invest in their own education. Similarly, parents who have healthier children also have stronger incentives to invest in their children's education. The nature of the link between health improvements and educational attainment has been highlighted in a series of theoretical models starting with Ben-Porath (1967) and more recent contributions by Boucekkine et al. (2002), Kalemli-Ozcan (2002), Cervellati and Sunde (2005, 2015), Soares (2005), Hazan and Zoabi (2006) and de la Croix and Licandro (2012). At the same time Soares (2006), Bleakley (2007), Jayachandran and Lleras-Muney (2009), Lucas (2010), Oster et al. (2013), and Hansen and Strulik (2017) have provided empirical evidence in support of this association.

Following the conclusions from this literature, we would expect that a reduction in the education gender gap would result from relatively faster increases in female education triggered by relatively stronger improvements in female health. As shown in Table 1, a simple comparison of the evolution of educational attainment of males and females and life expectancy at birth over the 20th century indeed suggests this pattern. While up until the 1950s the differences between males and females in terms of life expectancy and average years of schooling were fairly constant, in the years that followed we see relative female life expectancy rising sharply and the education gender gap declining.

[Table 1 around here]

As the similarity in the time trends of relative female life expectancy and education is only suggestive, in our analysis we explore more carefully the link between the two by exploiting exogenous variation in life expectancy gains triggered by the so-called International Epidemiological Transition (IET). This term refers to a period of rapid decline in mortality from previously highly fatal infectious diseases which started after the end of World War II and resulted in unprecedented improvements in life expectancy around the globe (Becker et al., 2005; Cutler et al.,2006). These improvements were brought

3

forward by a sequence of important medical innovations related to the development of antibiotics, vaccines and other treatment methods. Largely products of medical research in developed countries, these innovations diffused quickly around the globe as a result of the coordinated efforts of the United Nations and the World Health Organization. The consequence of this diffusion process was the almost complete eradication of many infectious diseases, which previously had affected large shares of the world's population, within a few decades.

In this historical context and following a line of research pioneered by Acemoglu and Johnson (2007), we estimate the impact of life expectancy improvements, triggered by the global spread of the new medical innovations, on educational attainment of males and females in an instrumental variables regression setup. For this purpose we use variation in the disease environment across countries prior to onset of the IET, as the introduction of the new treatment methods naturally had a large impact in places where mortality from infectious diseases was initially higher. In contrast to existing work that has studied the aggregate effects of the IET for the overall population (Acemoglu and Johnson, 2007; Cervellati and Sunde, 2011; Hansen, 2013), however, we consider the differential impact that the new medical technologies had on males and females, an important dimension that the literature has until now largely overlooked. In particular, we demonstrate that women benefited more in terms of life expectancy from the medical innovations associated with the IET than men. This in turn resulted in differential increases in female and male educational attainment and contributed to a sizeable reduction in the education gender gap.

Exploring the nature of these gender-specific effects of the IET on life expectancy, we show that they are rooted in the differential biological responses of males and females to vaccination. This finding relates to a recent medical literature which has documented that females exhibit stronger immune responses to vaccines than men (Cook, 2008; Klein et al., 2010; Furman et al., 2013). As a consequence of that, vaccines are more effective in providing disease protection for women compared to men. On the other hand, no such differences have so far been documented for antiviral and antibacterial drugs. In light of this evidence, we explore the extent to which the mortality reductions were due to vaccine- or non-vaccine-related medical innovations. Separating mortality improvements along these lines, we document that women experienced larger increases in life expectancy and educational attainment than men specifically in countries were the IET brought forth mortality reductions in infectious diseases that became preventable with vaccines. We also highlight the absence of such differential effects in countries where mortality reductions were primarily achieved for diseases for which vaccines did not constitute an effective method of control.

Taking into account both the gender- and the disease-specific characteristics of the IET allows us to explain a sizeable share of the reduction in the education gender gap that

4

occurred across countries between 1940 and 1980. Based on the estimated magnitudes for the above described effects we are able to explain 39% of the actual life expectancy increases of women and 33% of those of men. In terms of educational attainment, the estimated effects can account for 26% and 21% of the observed increases in female and male education. These differential increases in male and female education in turn imply that the differential impacts of the medical innovations associated with the IET on male and female health can account for 80% of the observed global reductions in the education gender gap between 1940 and 1980.

Going one step further, we also study the broader macroeconomic implications of these differential improvements in male and female health. In this context we provide evidence that the negative association between GDP per capita and life expectancy reported by Acemoglu and Johnson (2007) occurred largely in countries where the relative increases in female life expectancy over this period were small. Furthermore, we show that the positive effects of the post-1940 life expectancy gains on income per capita documented by Cervellati and Sunde (2011) for countries that had already undergone the demographic transition were concentrated among countries where female life expectancy rose more than male life expectancy. Overall these results suggest that female health improvements can have an important positive contribution to economic development.

To establish these results we proceed as follows. Section 2 reviews the evidence from the medical literature on gender differences in the context of infectious diseases. Section 3 outlines the empirical strategy that we follow in the paper. Section 4 describes the data that we use. Section 5 presents our baseline results regarding the effect of life expectancy on educational attainment, while their robustness is assessed in Section 6. Section 7 presents evidence regarding the effects of improvements in male and female life expectancy on GDP per capita. Section 7 offers some concluding remarks.

2   Gender Differences related to Infectious Diseases

Since Grossman (1985), the medical literature has recognized that immune responses to pathogens differ for men and women. Men generally show weaker immune responses than women, making them more susceptible to contract infectious diseases and having more severe disease outcomes compared to women. In this section we summarize the key evidence regarding the different ways in which men and women are affected by infectious diseases and how they respond to vaccination. These gender differences are later on explored in our empirical analysis.[1]

In their review of recent epidemiological studies, Bernin and Lotter (2014) document that parasitic diseases, such as malaria, affect men more frequently and more severely than

---

[1]As this paper relates biological differences between men and women to educational outcomes in a given social context, we use the terms 'sex' and 'gender' interchangeably to refer to these differences.

women. Neyrolles and Quintana-Murci (2009) and Nhamoyebonde and Leslie (2014) show that males are also more frequently affected by tuberculosis compared to females, with the male-to-female ratio being 1.9 based on case notifications. Muenchhoff and Gouldner (2014) discuss gender differences in pediatric infectious diseases and provide evidence that males have typically higher disease susceptibility as well as higher morbidity and mortality than females. Giefing-Kroell et al. (2015) review the evidence on the effect of 24 infectious diseases on adults and report a male bias for 54% of the diseases in terms of incidences and of 57% of the diseases in terms of mortality.

Looking at broader sets of data on mortality rates from infectious diseases also indicates the presence of a male bias. Owens (2002) provides some evidence for this using data from the United States, while Lozano et al. (2012) demonstrate it based on the Global Burden of Disease Study. In particular, Lozano et al. estimate sex-specific mortality rates for 235 causes of deaths based on vital statistics from 187 countries between 1990 and 2010. These data show that across the 27 most important infectious diseases male mortality rates are on average 13% higher than the corresponding rates for females. Only five out of the 27 major diseases show a slight female bias, while in four cases the mortality rates appear to be the same for males and females.

Looking at the age-specific nature of the reported male bias, the existing evidence suggests that the bias is partially driven by physiological differences related to sex hormones. In the case of parasitic diseases, Bernin and Lotter (2014) provide evidence that the higher disease susceptibility of males relative to females tends to peak during middle ages, following the increase in testosterone levels. Similarly, Guerra-Silveira and Abad-Franch (2013) document that the male bias in disease incidence for most pathogens typically first emerges during the first year of life. This coincides with the surge of sex hormones during the so-called mini-puberty. It then falls during early childhood when hormonal differences are minimal and peaks again during puberty and reproductive ages when the variation in sex hormone levels between males and females is highest.[2] In addition to variation in sex hormones, Baghdadi et al. (2006) and Davila et al. (2008) show that genetic factors related to the X-chromose may play a role in explaining the sex bias. This evidence, together with the fact that exposure to pathogens is typically sex-unbiased (Guerra-Silveira & Abad-Franch, 2013) imply that the observed male bias in disease susceptibility and mortality is largely of biological nature and unrelated to behavioral or environmental factors.

The conclusion that the male bias in disease susceptibility and mortality from infectious diseases is related to variation in sex hormones has also been strongly supported

---

[2]The role of hormones for the male bias in infectious diseases is also supported by specific case studies among humans. For example, Hamilton and Mestler (1969) show that castrated males have a lower mortality rate from tuberculosis than males with intact genitalia. Similarly, Svanberg (1981) finds that women who underwent surgical removal of the ovaries have higher rates of tuberculosis mortality than the overall female population.

by laboratory experiments with animals. Such experiments have shown, for instance, that treatment of female mice with testosterone reduces their resistance to parasites and increases their disease burden (Cernetitch et al., 2006), whereas castration renders male mice more resistant (Wunderlich et al., 2002; Krucken et al., 2005). For the case of tuberculosis, Yamamoto et al. (1991) have shown that male mice are more susceptible to infections and develop more severe disease outcomes. Just like in the above case, the authors provide evidence that castration increases resistance to the disease, while treatment with testosterone reduces it. In a similar vein, Tsuyuguchi et al. (2001) have shown that female mice whose ovaries had been removed exhibit lower resistance to the tuberculosis bacterium than control mice that underwent placebo surgery. Moreover, treatment with estrogen of the female mice without ovaries was shown to restore disease resistance to the same level as in the control group of mice.

More recent studies have documented that such differences between males and females characterize also their responses to vaccines against infectious diseases. This was first documented by Stanberry et al. (2002) and Cook (2008). These studies demonstrate that vaccine efficacy, namely the relative reduction in disease susceptibility of a vaccinated group of individuals compared to an non-vaccinated one, is significantly higher for women. While Stanberry focused on the case of herpes vaccines, Cook demonstrated that this pattern can be observed both for viral as well as for common bacterial diseases. Their findings have also been confirmed in a series of replication studies by Klein et al. (2010), Klein and Pekosz (2014) and Klein et al. (2015). These authors document that disease protection from vaccines is stronger among women for several common infectious diseases.[3] These differences in vaccine efficacy have also been shown to be substantial. For example, Klein et al. (2010) show that in the case of influenza women exhibit the same antibody response to half a dose of the vaccine compared to a full dose in men.

It is important to note here that this pattern is not due to differences in the vaccine doses administered to men and women. In fact, the standard medical practice is to universally administer the same dose (Poland et al., 2008a), irrespective of sex. Instead, as in the case of the male bias in disease susceptibility, the evidence suggests that these differences should be attributed to sex hormones, especially estrogens (Klein et al., 2010) and testosterone (Furman et al., 2013). Estrogens have been shown to stimulate the activity of immune cells more generally, while testosterone has been shown to suppress it (Fish, 2008; Giefing-Kroell et al., 2015). Fish (2008) and Libert et al. (2010) also link differences in vaccine efficacy between males and females to sex chromosomes. The fact that females have two X-chromosomes, a maternal and a paternal one with a random pattern of expression, contributes to a genetic advantage due to a more extensive repertoire of proteins and greater diversity in physiological processes. As such, the stronger

---

[3]These diseases include herpes, hepatitis, influenza, measles, mumps, rubella, smallpox, tuberculosis and yellow fever.

immune response to vaccines in women is in line with women being less susceptible to contract infectious diseases and having less severe disease outcomes, as discussed above. This is because sex hormones and genetic differences not only influence the functioning of immune cells in response to vaccines, but also affect the response to naturally occurring pathogens (Cook, 2008; Fish, 2008).

These stark gender differences in the effectiveness of vaccination against infectious diseases contrasts with the case of other treatments methods. In fact, for antivirals, antibiotics and other drugs acting on the immune system there is no similar evidence of clear differences in the responses of males and females. Only for drugs acting on the central nervous system, such as anti-psychotic drugs and antidepressant, as well as beta-blockers reducing the heart rate and systolic blood pressure there is weak evidence for more beneficial effects on women compared to men (Franconi et al., 2007).

Overall these medical studies underscore an important pattern: Immune responses of men and women to infectious diseases are different and this is true for both exposure to the naturally occurring pathogen as well as to the associated vaccine. These differences do not appear to be driven by behavior, but are rooted in biological differences related to sex hormones. This suggests that vaccination campaigns will give rise to differential effectiveness across genders.[4] In particular, vaccination campaigns against infectious diseases are expected to trigger larger mortality reductions for females than for males, while this is not expected to be the case for other public health interventions aimed at controlling infectious diseases which do not rely on vaccination. In the following section we describe how we are going to exploit this pattern in the context of our empirical strategy.

3 Empirical Strategy

Our empirical strategy builds on the approach of Becker et al. (2005), and Acemoglu and Johnson (2007) who investigate the macroeconomic effects of the large improvements in life expectancy that took place over the second half of 20th century. In particular, we explore the fact that countries where the mortality burden from infectious diseases was higher before the onset of the IET were the ones that benefited more in terms of mortality reductions from the new treatment methods that became available globally after the 1940s. Following the identification strategy proposed by Acemoglu and Johnson, we exploit the exogenous nature of the initial differences in mortality rates from infectious diseases to identify the impact of the subsequent changes in life expectancy on educational

---

[4]This is provided that these campaigns are equally effective at targeting males and females. As documented in World Bank (2012), though, there is no evidence for gender discrimination in the use of preventative health services, such as vaccination. Even in a country like India with very unequal gender norms vaccination rates do not appear to differ between boys and girls (Banerjee et al., 2010).

attainment in an instrumental variables regression.

Our analysis, however, differs from previous contributions in the literature, as we investigate the gender-specific nature of the improvements in life expectancy and their effect on educational attainment. Specifically, we take into consideration whether the initial mortality environment in a given country was dominated by infectious diseases that were subsequently controlled by vaccination or by other prevention methods, such as the use of antibiotics. This strategy allows us to exploit the fact that different health campaigns aimed at controlling infectious diseases have had different effects on males and females, as discussed in the previous section.

With the global spread of western medical technologies in the post-war era, most infectious diseases became preventable and mortality rates from these diseases fell substantially around the world. Yet, this effect was not uniform for males and females. In countries where the mortality environment prior to 1940 was dominated by infectious diseases that subsequently became preventable thanks to the introduction of new or improved vaccines we would expect to see a stronger increase in life expectancy for females than for males. In contrast, in countries where the mortality environment prior to 1940 was dominated by infectious diseases that subsequently became preventable thanks to other medical innovations we would expect to see similar increases in female and male life expectancy.

Exploiting this pattern, we estimate the impact of changes in life expectancy of males and females on their educational attainment in a long-differences panel between two dates, which for most of our analysis we take to be 1940 and 1980. Specifically, our baseline estimation equation is the following:

$$AYS_{gct} = \alpha \cdot LE_{gct} + \mu_{gc} + \gamma_t + u_{gct}. \tag{1}$$

Here AYS_{gct} denotes the average years of schooling of a given cohort of gender g in country c in year t and LE_{gct} denotes the life expectancy at birth of gender group g in country c at time t. The specification includes gender-country fixed effects, μ_{gc}, and year fixed effects γ_t. Given this fixed effect structure, our panel regression with two dates is equivalent to a specification in first differences. Conditional on these fixed effects, a positive α coefficient would suggest that life expectancy improvements between between 1940 and 1980 were associated with increases in educational attainment over this period.

To account for the potential endogeneity bias in the estimation of α, we use a two-stage least square (2SLS) estimation procedure where we instrument LE_{gct} in equation (1) using the following first-stage specification:

$$
\begin{aligned}
LE_{gct} = {} & \beta^V \cdot \sum_{d \in V} M_{dct} + \beta^{Vf} \cdot I_g^f \cdot \sum_{d \in V} M_{dct} + \dots \\
& \dots + \beta^{NV} \cdot \sum_{d \in NV} M_{dct} + \beta^{NVf} \cdot I_g^f \cdot \sum_{d \in NV} M_{dct} + \eta_{gc} + \delta_t + \varepsilon_{gct}.
\end{aligned}
\tag{2}
$$

The subscripts g, c, and t again denote gender, country, and year, while the index d denotes different infectious diseases. Thus, M_{dct} is the mortality rate from infectious disease d in country c and year t, which, following Acemoglu and Johnson (2007), we take to be equal to the actual country-specific values in 1940 and the mortality rate at the global health frontier in 1980. As M_{dct} takes the same values for all countries in 1980, the variable does not capture actual changes in mortality over time for disease d and the pace at which the disease was eradicated. Instead it reflects predicted changes in mortality from a given infectious disease that followed as a consequence of the medical innovations and their global diffusion after 1940. Thus, changes in M_{dct} are treated as a predictor for the actual changes in mortality from infectious diseases that occurred across countries and genders between 1940 and 1980 and are used as an instrument for the associated changes in life expectancy.

These mortality rates are then aggregated into two groups. A group of vaccine-preventable diseases, which we denote as V and refer to as the *V-disease group*, and a group of non-vaccine-preventable diseases, which we denote as NV and refer to as the *NV-disease group*. The aggregate mortality rates for these two groups are interacted with a dummy variable for all female observations, I_g^f, in order to estimate the differential effects of the predicted mortality reductions across genders. The specification also includes gender-country and year fixed effects in line with equation (1).

As larger changes in M_{dct} over time indicate greater predicted reductions in mortality from infectious diseases of a given group, they should be associated with larger increases in life expectancy. Thus, we would expect both β^V and β^{NV} to be negative. Comparing these effects across males and females, we would expect the interaction coefficient β^{Vf} to be negative. This is because, in light of the medical evidence presented above, reductions in mortality from infectious diseases that became preventable with the introduction of new vaccines are bound to be larger for females. On the other hand, as there is no expectation for larger increases in female life expectancy from reductions in mortality from non-vaccine-preventable infectious diseases, we expect the interaction coefficient β^{NVf} to be zero.

The unbiased estimation of the key coefficients of interest in our empirical setup requires the exogeneity of the employed predicted mortality changes for the different groups of infectious diseases. In that respect, the crucial assumption is that the initial mortality rates, which reflect the mortality environment in each country prior to the IET, are uncorrelated with other time-varying country and gender specific characteristics that influence education through channels other than life expectancy. To ensure the validity of this assumption, as part of our robustness analysis, we control for several time-varying correlates of male and female health and educational attainment, the omission of which could bias our results. Other factors leading to persistent differences in the level of educational attainment of males and females in a given country are not explicitly controlled

for in our regression as they will be filtered out by the gender-country fixed effects.[5]

4 Data

Our estimation is based on a panel data set covering 75 countries in two time periods. As already mentioned in the previous section, for our baseline regressions the two time periods correspond to the years 1940 and 1980.[6]. We focus on changes in education, life expectancy and mortality between these two time periods in order to assess the impact of the medical innovations associated with the IET over the second half of the 20th century before the start of the global HIV/AIDS epidemic, as in Acemoglu and Johnson (2007), Cervellati and Sunde (2011), and Hansen (2013). As we are interested in the differential impact that the new medical technologies had on the outcomes of males and females, our data set includes for each country and year two observations, one for the female population and one for the male population. This leads to a total sample size of $75 \cdot 2 \cdot 2 = 300$ observations.

Section A of the Appendix lists the 75 countries that are included in our main sample. They span all regions of the world with the exception of Sub-Saharan Africa for which reliable data on life expectancy and mortality rates only start in 1950. In this section we briefly describe the key variables of interest, namely years of schooling, life expectancy and mortality rates. Further information on the data sources for these variables as well as for all additional data that we employ in our analysis can be found in Section B of the Appendix.

To measure educational attainment, we use the gender- and cohort-specific average years of schooling data provided by Barro and Lee (2013). Following the approach of Hansen (2013), we focus on the cohort of individuals that was between 5 and 9 years old in the two respective time periods (1940, 1980) and measure their educational attainment as observed in the Barro and Lee (2013) data 10 years later (1950, 1990).[7] These are the cohorts of boys and girls that entered the formal schooling system between 1937 and 1941, and between 1977 and 1981, respectively. By comparing them we can assess how levels of educational attainment around the globe were affected by the life expectancy improvements associated with the IET.

Data on life expectancy at birth in 1940 are drawn mainly from the various editions of the UN Demographic Yearbook, supplemented with additional data for selected countries as explained in Section B of the Appendix. The respective information for 1980 is obtained

---

[5]Such differences could be related to the the relative status of women in society as well as to their relative access to health care and schooling.

[6]In our robustness analyses we also consider alternative setups where the two time periods correspond to different years.

[7]As part of our robustness analysis we explore both the measurement of educational attainment at higher ages and the impact of life expectancy on primary, secondary and tertiary education.

from the electronic version of World Population Prospects of the UN Population Division. We furthermore collected data for life expectancy at higher ages. The sources for these data are the same as for life expectancy at birth, namely UN Demographic Yearbooks for 1940 and World Population Prospects for 1980. In Section B of the Appendix we also explain how we deal with missing data.

The information on mortality rates from infectious diseases used in our baseline analysis is drawn from Acemoglu and Johnson (2007). The authors collect and report disease-specific mortality rates for 13 infectious diseases in 1940. We classify these diseases into two groups depending on whether the disease became preventable with vaccines during the post-war era or not. The V-disease group of vaccine-preventable diseases includes the following seven diseases: diphtheria, influenza, measles, pneumonia, smallpox, tuberculosis, whooping cough. The remaining six diseases that fall into the NV-disease group of non-vaccine-preventable diseases are: cholera, malaria, plague, scarlet fever, typhoid fever, typhus. Section C of the Appendix provides detailed information on the characteristics and key methods of control of each disease based on which we perform this classification.

For the year 1980 we do not collect actual mortality rates. This is because our mortality variable should not reflect the actual country-specific mortality environment in that period, but the conditions at the global health frontier. For our baseline analysis we follow Acemoglu and Johnson (2007) and take the frontier mortality rates in 1980 to be zero for all infectious diseases. As an alternative to this approach, we assume that all countries experienced the same proportionate reductions in mortality between 1940 and 1980, rather than reaching the same level in 1980. In this case, the predicted mortality rate for each disease in 1980 is taken to be equal to the country-specific rate in 1940 scaled down by the average rate at which mortality fell for that disease at the global level. As a second alternative, we assume that the predicted mortality rates for all countries in 1980 were equal the values observed in the United States in that year, which are close to but not equal to zero. By using the predicted mortality changes for each country between 1940 and 1980, rather than the actual changes, our mortality variable can function as an instrument for changes in life expectancy. This is because changes in predicted mortality are correlated with actual changes in mortality, which in turn determine the evolution of life expectancy.

We should note that our mortality rates are not gender-specific, but refer to the total population of each country. Neither Acemoglu and Johnson (2007) nor the original sources of their data provide any gender-specific mortality rates. When estimating our first-stage regression specification (2), therefore, we assign the same values to the male and female population in each respective country and only allow their effect on life expectancy to differ across genders. For a smaller sample of 22 countries, though, we were able to obtain gender-specific mortality rates from infectious diseases in 1940 from Preston et al. (1972).

These countries are listed in Section A of the Appendix. Based on this smaller sample we provide some robustness analyses in Section 6.1 and discuss the direction of the bias caused by the potential measurement error.

Table 2 shows the descriptive statistics for all key variables. For average years of schooling and life expectancy at birth the statistics are reported separately for males and females. As these figures clearly indicate, both schooling and life expectancy increased substantially between 1940 and 1980 and these increases were larger for females than for males. The exact nature of this relationship is what we investigate in the next section.

[Insert Table 2 around here]

5 Baseline Regression Results

Table 3 presents the results from the estimation of our baseline specification with 2SLS. Panel A shows the results of the first-stage estimation based on different variants of equation (2), while Panel B shows the results of the corresponding second-stage estimation of equation (1). Standard errors, reported in brackets, are clustered at the gender-country level in line with the employed fixed-effects.

In column 1 we present a simple variant of the first-stage estimation where we omit the interaction term with the female dummy and we do not separate the mortality rates into different disease groups. Instead we look at the overall effect of the predicted mortality reductions from the 13 infectious diseases on life expectancy at birth. As the estimated coefficient indicates, the improvements in the mortality environment that took place between 1940 and 1980 had a large and statistically significant effect on life expectancy. On average the estimates imply an increase in life expectancy at birth by 7 years, which is similar to the magnitudes reported by Acemoglu and Johnson (2007).

[Insert Table 3 around here]

In column 2 we estimate separately the effect of the predicted mortality reductions for males and females by including in the specification the interaction term between the female dummy and the predicted mortality rates. The estimated coefficient for the interaction term is negative and statistically significant at the 1% level, indicating that female life expectancy, on average, rose faster between 1940 and 1980 than male life expectancy in response to the same predicted improvements in the mortality environment. Specifically, the estimated coefficients of −12.9 and −4.9 imply that male life expectancy increased on average by 5.6 years in response to changes in predicted mortality and female life expectancy increased by 7.8 years. This corresponds to 33% and 39% of the actual life expectancy increases for males and females respectively observed over this period in our sample of countries. The relatively larger effect on female life expectancy does not

13

imply that female life expectancy increases more in response to the same reduction in actual mortality. By construction life expectancy at birth reflects the overall mortality environment at a specific point in time and its calculation is gender-invariant. Instead, we should interpret the absolutely larger coefficient for females as reflecting that the actual mortality reductions experienced by females in response to the post-1940 medical innovations were larger than those for males. These effectively larger mortality reductions in turn gave rise to relatively larger increases in life expectancy.

In column 3 we estimate again the differential effects of the predicted mortality reductions for males and females, but distinguishing this time between mortality from vaccine-preventable diseases (V-disease group) and from non-vaccine-preventable diseases (NV-disease group). As we can see from the estimated coefficients, the predicted reductions in mortality rates from both groups of diseases were associated with significant increases in life expectancy. Moreover, as the interaction term with the female dummy reveals, the reduction in predicted mortality from vaccine-preventable diseases led to statistically significantly larger increases in female life expectancy than in male life expectancy. In contrast, the corresponding coefficient on the interaction effect with predicted mortality from non-vaccine-preventable diseases is statistically insignificant. This result is, thus, in line with the findings of the medical studies summarized in Section 2 which find vaccine efficacy to be higher among females.

The estimation results of column 3 imply that between 1940 and 1980 female life expectancy increased more than male life expectancy in countries where the mortality improvements were largely the result of vaccination campaigns. In countries where these improvements were driven by the introduction of new drugs or other health interventions, however, this does not appear to be the case. Comparing the magnitudes of these differences we see that, as a consequence of the mortality reductions from vaccine-preventable diseases, male life expectancy rose on average by 4.2 years and female life expectancy by 6 years. This means that the effect for females is 43% higher than that for males. Looking at the corresponding magnitudes for non-vaccine-preventable diseases instead, we find that the mortality reductions led to increases in life expectancy by on average 1.2 years for males and 1.55 years for females, with the difference between the two being statistically insignificant.[8]

In columns 4 and 5 we present the estimation results for the same regression specification but where the predicted mortality rates in 1980 are not assumed to be zero, but adjusted as described in Section 4. Specifically, in the estimation in column 4 we assume that mortality rates fell proportionately in line with the evolution of the global average, while in the estimation in column 5 we assume that the mortality rates fell to

---

[8]While the point estimate of -5.6 on the interaction term appears large, the actual magnitude of the effect is quantitatively not important as the variation in mortality rates for the NV-disease group in our sample is not very high.

the levels observed in the United States in 1980. In both cases the estimation results are very similar to the baseline result reported in column 3, confirming the pattern of a relatively stronger response of female life expectancy to changes in predicted mortality from vaccine-preventable diseases, but an insignificant gender-specific response to changes in predicted mortality from non-vaccine preventable diseases.

Turning to Panel B of Table 3, we can see the second-stage estimates of the 2SLS estimation that correspond to the first-stage estimations described above. The results in all cases are remarkably consistent.[9] Irrespective of the exact set-up in the first stage, the second-stage estimates suggest that improvements in life expectancy between 1940 and 1980 led to a significant increases in average years of schooling. Also when estimating equation (1) with OLS, as reported in column (6), we obtain a positive coefficient of similar magnitude. These estimates imply relatively higher increases in educational attainment of females compared to males, since the underlying life expectancy gains were higher for females than for males. In particular, the coefficient estimate of 0.115 in column (2) in combination with the predicted changes in life expectancy from the first stage regression imply an increase in schooling of on average 0.65 years for males and 0.9 years for females.[10] The resulting reduction of 0.25 years in the education gender gap can account for 80% of the actually observed reduction in the education gender gap over the sample period.

Comparing these estimates with previous work in the literature is also reassuring. Our second-stage coefficient estimate suggests that one extra year of life should lead to an increase in schooling by 0.115 years. This is similar to the effects found by Jayachandran and Lleras-Muney (2009) who report effect sizes between 0.11 and 0.15 years of schooling for each addition year of life. Alternatively, we can look at the implied elasticities of changes in educational attainment to changes in life expectancy. These elasticities are found to be between 0.6 and 1 by Jayachandran and Lleras-Muney (2009) based on data from Sri Lanka, and between 0.8 and 1.3 by Oster, Shoulson and Dorsey (2013) based on data from the United States. In our setup, given the initial levels of life expectancy and schooling, mortality reductions associated with the IET increased life expectancy by 16% for women and 12% for men. This in turn resulted in a 21% increase in average years of schooling for females and 14% for males, thus implying an elasticity of 1.18 for males and 1.34 for females.

We also tested whether the same improvements in life expectancy had differential effects on educational attainment of males and females by including in the second stage

---

[9]Also the reported F-statistics for the significance of the instruments in the first-stage yield in all cases values safely above the critical threshold of 10 suggested by Staiger and Stock (1997).

[10]We should note here that, compared to the actual increases in schooling of 3.15 years for males and 3.5 years for females that we observe in our sample, the increases predicted by our estimation may appear small. Yet, this is natural and should be expected given the presence of several other factor that contributed to increases in schooling between 1940 and 1980.

specification an interaction term of the life expectancy variable with the female dummy. While the table does not report these results for brevity, in all cases we find this interaction term to statistically insignificant. This suggests that the observed differential increases in years of schooling of females and males are mostly due to differential improvements in life expectancy and not to differential responses in schooling to given increases in life expectancy.

6   Robustness Checks

Having demonstrated the quantitative importance and the statistical significance of the link between the differential improvements in life expectancy across genders and the closing of the education gender gap in our baseline regressions, we proceed in this section to establish its robustness. For this purpose, we first explore the robustness of our first-stage estimation by looking at mortality from different sub-groups of infectious diseases and by employing gender-specific mortality rates. We then proceed to assess the robustness of our second-stage estimation by adjusting our country sample, considering life expectancy at different ages and using alternative measures of educational attainment. Finally, we also compare the relative importance of our explanation for the closing of the education gender gap with alternative ones proposed in the literature.

6.1   Robustness of First-Stage Estimation

If increases in life expectancy resulting from the introduction of vaccines are stronger for females than for males, as established in our first-stage estimation, then this pattern should be observable also for individual diseases as well as sub-groups of diseases. With that in mind in columns 1, 2, and 3 of Table 4 we repeat our first-stage estimation focusing on changes in mortality rates from the three most important infectious diseases of that time: malaria, pneumonia, and tuberculosis. Doing so is instructive as these three diseases together account for 87% of mortality from the 13 infectious diseases in 1940. In each of the three columns, we report the effect of the predicted changes in mortality from one of the three diseases, as indicated in the top part of the table, while controlling at the same time for the predicted changes in mortality rates from the remaining twelve diseases with the *Residual Mortality* variable.

[Insert Table 4 around here]

Comparing the estimation results across the three columns, we find that the interaction term between the female dummy and the predicted changes in mortality is statistically insignificant for the case of malaria, but negative and statistically significant for the case of pneumonia and tuberculosis. These results are in line with the expectation

16

that the relatively larger gains in female life expectancy over this period were driven by a stronger immune response of females to vaccination. As we explain in greater detail in Section C of the Appendix, both pneumonia and tuberculosis are diseases which after 1940 largely became preventable with vaccines. Malaria, on the other hand, was primarily controlled by newly developed insecticides, such as DDT, and to this date no effective vaccine exists.

Given the importance of these three diseases over our sample period, an important test is to ensure that our results are not only driven by a differential response of females to specific treatment methods for malaria, pneumonia and tuberculosis. For this purpose in column 4 we focus on the remaining ten diseases, splitting them once again into a group of diseases that became vaccine-preventable after 1940 and a group of diseases that did not. In this setup we also control at the same time for predicted mortality from the three excluded diseases in the Residual Mortality variable. As the results demonstrate, we see again differential changes in female life expectancy resulting from predicted reductions in mortality from diseases of the V-group, even with pneumonia and tuberculosis excluded, but not for diseases of the NV-group.

A related concern regarding our first-stage estimation is that the observed differential life expectancy gains for females and males may be due to variation in the causative agent behind the diseases rather than the role of vaccination and other methods of disease control. While the medical literature has not documented any differences in the immune response of females and males across different types of causative agents, we nevertheless test for this by separating the 13 diseases with respect to their causative agent (bacteria, viruses, parasites) as well as their main method of control. As all viral diseases in our data set (influenza, measles, smallpox) are vaccine-preventable and malaria is the only parasitic disease, for this robustness check we focus on just bacterial diseases. In column 5 we separate the nine bacterial diseases in our data set into vaccine- and non-vaccine-preventable ones and interact the predicted mortality rates from these two groups of diseases with the female dummy. In this setup again we control for the predicted mortality rates from the remaining four non-bacterial diseases with the Residual Mortality variable. Once again we find evidence that predicted mortality reductions resulting from the introduction of vaccines were associated with larger gains in life expectancy for females than for males, while the responses in life expectancy to other methods of disease control were uniform across genders. Thus, we can safely conclude that the patterns that we observe are not driven by the causative agent behind the different infectious diseases.

A final concern regarding our first-stage estimation may be the fact that the employed predicted mortality rates are not gender-specific. While disease-specific mortality rates in 1940 are not available by gender for all 75 countries in our main sample, as we already mentioned in Section 4, such information is available for a small sample of 22 countries. Based on the data reported by Preston et al. (1972) we construct gender-specific mortality

rates in 1940 for two groups of infectious diseases: a group of vaccine-preventable diseases that includes influenza, pneumonia and tuberculosis and a residual group containing all other infectious diseases. The robustness of our main finding using these data is explored in columns 5 and 6. In both cases we assume, in line with our baseline estimation, that the mortality rates from all diseases would fall to zero by 1980.

In the estimation in column 5 we employ the data for this small sample of countries, but ignore gender differences in mortality rates by assigning the country-wide mortality rates to both males and females, just like in our baseline specification. The estimation in column 6 employs the actual gender-specific mortality rates for males and females. The estimation results in both cases are similar.[11] Controlling for the predicted reductions in mortality from the diseases of the residual group, the predicted declines in mortality from influenza, pneumonia and tuberculosis, which became preventable with vaccines, had a significantly larger effect on female life expectancy. The corresponding absolute magnitudes of the effects are slightly higher when we employ the gender-specific mortality rates, but are qualitatively similar.

This last set of results suggest that the fact that we cannot employ gender-specific mortality rates in our main specification, if anything, will lead to an underestimation of the differential increases in female life expectancy that resulted from the IET. The underestimation is caused by the fact that female mortality rates in 1940 were already lower than the male rates. As discussed in Section 2, this has to do with the stronger immune response of women to pathogens, making them less susceptible to contract infectious diseases and having less severe disease outcomes. Using the average mortality rates to construct our mortality instrument will, thus, lead to an underestimation of the actual mortality reductions experienced by males and an overestimation of the mortality reductions for females. As a consequence, any resulting bias in our estimation will attenuate our female-specific coefficient of interest, β^{Vf}, toward zero.

6.2 Robustness to Sample Composition and Life Expectancy Measurement

Having shown that the first-stage estimation results do not hinge on the exact way in which we specify our predicted mortality instrument, we proceed to assess the robustness of our overall 2SLS estimation. A crucial assumption behind our estimation is that the mortality instrument can be reasonably excluded from our second-stage specification. Practically this means that the predicted changes in mortality rates from infectious diseases that occurred between 1940 and 1980 as a result of the IET were exogenous to the

---

[11]Due to the small sample size, in this setup we are forced to reduce the number of included fixed effects by employing separate fixed effects for gender and country rather than one for each gender-country pair.

potential of men and women in our sample of countries to attain a given level of education. While this assumption seems plausible for most countries in our sample, it may not entirely hold for the countries that were actively involved in the development of the key medical and chemical innovations that led to the IET. This group of countries includes France, Germany, Switzerland, the United Kingdom and the United States. With that in mind, in column 1 of Table 5, we omit these countries from the sample.[12] As we can see, this modification of our country sample does not affect our main findings. The estimated coefficients for all variables are very similar to those reported in Table 3 and we continue to observe a relatively stronger effect of the predicted mortality reductions from vaccine-preventable diseases on the life expectancy of females.

[Insert Table 5 around here]

Another concern regarding our estimates has to do with the quality of the data. As for some of the 75 countries in our sample the main data may not be very reliable, in columns 2 to 4 of Table 5 we omit one by one certain groups of countries for which data quality may be questionable. Specifically, in column 2 we omit African and Middle Eastern countries, in column 3 small Latin American and Caribbean countries, and in column 4 Eastern European countries. In all cases we can see from the resulting estimates that omitting these groups of countries does not alter our findings in any significant way, neither in the first stage nor in the second stage.

Up to this point our analysis has focused solely on the link between life expectancy at birth and educational attainment. We do so, as life expectancy at birth is the most common indicator of the overall mortality environment in a given country. Yet, mortality improvements may raise educational attainment in different ways depending on the exact age at which these improvements are realized. With that in mind it is instructive to consider the relative role played by mortality improvements in childhood versus adulthood. Hence, in columns 5 and 6 of Table 5, we replace life expectancy at birth with life expectancy at the ages of 5 and 20 respectively. Doing so leads to a sizeable decrease in our sample size as these alternative measures are not available for all countries in our sample. As the estimates reveal, though, this does not alter the qualitative nature of our previous findings.

In particular, we should note here that while the second stage results in Panel B suggest that improvements in life expectancy at higher ages have a larger effect on educational attainment, this is not actually the case. The higher estimated coefficients are a consequence of the fact that the changes in life expectancy at age 5 or 20 were smaller than those at birth. Expressed in terms of standard deviation changes, the magnitudes

---

[12]Omitting these countries also takes care of the concern that medical research at that time may have been biased toward finding vaccines against the diseases most prevalent in these countries.

of the effects are very similar to that in the baseline regression with a one standard deviation increase in predicted life expectancy raising educational attainment between 0.24 and 0.26 standard deviations.[13]

6.3 Robustness to Alternative Measures of Educational Attainment and Time Trends

As a further check on the robustness of our estimation results, it is instructive to investigate the relationship between life expectancy improvements and alternative measures of educational attainment. This is what we focus on in Table 6. Specifically, in column 1 we begin by looking at how the improvements in life expectancy between 1940 and 1980 affected the educational attainment of the cohorts of males and females who started school 5 years later, namely in 1945 and 1985. These are cohorts that started their formal schooling after World War II was over and, hence, their schooling should not have been affected by disruptions caused by the war. As the estimated coefficient in Panel B reveals, the effect of life expectancy on schooling for these two cohorts is very similar to the baseline case. Thus, our finding does not seem to hinge on wartime disruptions in the education systems that occurred in the countries involved in World War II.

[Insert Table 6 around here]

Having established this result, in columns 2 to 4 we proceed to assess whether the link between life expectancy at birth and educational attainment operates more at the primary schooling level or at the high level. For this purpose we use as our dependent variable in the second-stage estimation, instead of average years of schooling, the share of the respective cohort in the population that has completed primary education in column 2, secondary education in column 3, and tertiary education in column 4. While the point estimates are not directly comparable between these three cases and our baseline setup where we focus on average years of schooling, qualitatively the results are in line with the previous ones showing a significant positive effect of life expectancy improvements on educational attainment at all levels of education. Specifically, the estimates imply that the improvements in life expectancy observed between 1940 and 1980 led on average to a 19 percentage point increase in the primary school completion rates of males and to a 21 percentage point increase in that of females. For secondary education, the corresponding increases in the completion rates for males and females are 2.3 and 1.8 percentage points respectively, while at the tertiary level they are 0.7 and 0.3 percentage points. These figures suggest that the global health improvements over the post-war period had the

---

[13]Repeating our baseline 2SLS estimation with life expectancy at birth but restricting the sample to those countries for which life expectancy data at age 5 or 20 are available produces almost the same point estimates as in Table 3. Thus, the difference in the estimated coefficients in columns 5 and 6 compared to Table 3 are not the result of changes in the sample composition.

largest impact on primary education. This is not surprising given that many of the new treatment methods for infectious diseases benefited primarily young children who may have otherwise been forced to drop out of school at a young age or never started school to begin with.[14] Given that among cohorts who started school around 1940 less than 15% of the population completed secondary education and less than 5% completed tertiary schooling, the relatively larger increases for females in terms of primary schooling completion rates over the subsequent 40 years appear as the main driver behind the overall observed fall in the overall education gender gap.

An important concern regarding all the results presented so far is that they could be driven by pre-existing trends in the schooling data that happen to coincide with time trends in the life expectancy data but may be unrelated to the life expectancy improvements triggered by the IET. To check that this is not case, in column 5 of Table 6 we lag our average years of schooling data by 40 years and include it as a control variable in both our first- and second-stage specification. In both cases we find lagged schooling to be statistically insignificant and to not affect our main coefficients of interest. As a further test for pre-existing trends in the data, in column 6 we conduct a falsification test by replacing our baseline schooling variable with its 40th-year lag. If post-war life expectancy improvements can predict the rise in schooling observed between 1900 and 1940, this would indicate that our previously estimated relationship may be spurious and that the post-1940 increases in educational attainment would have occurred even in the absence of the IET. Reassuringly, though, we find the relationship between the two to be weak and statistically not different from zero. This should makes us confident that pre-existing trends in the schooling data are not driving the results.

6.4 Robustness to Alternative Explanations for the Rise in Female Schooling

The existing literature on the education gender gap, summarized in the introduction, has already highlighted several factors that may account for the relative rise in female educational attainment over the past decades. As the relative importance of these factors has not been analyzed before in a cross-country setting, it is instructive to consider them and compare them with the improvements in life expectancy that our analysis emphasizes. This comparison is offered in Table 7 where we include in our baseline specification various additional control variables that reflect these alternative explanations for the rise in female schooling. Further details on the construction of these variables are provided in Section B of the Appendix.[15]

---

[14]This finding is very much in line with evidence summarized in Almond and Currie (2011) regarding the large effects of disease exposure in early childhood on schooling outcomes.

[15]As the required data for the construction of some of these additional control variables are available only for a small set countries, the sample size in these regressions varies and occasionally we are forced

[Insert Table 7 around here]

In column 1, we start by controlling for each country's level of GDP per capita in the respective year in both the first- and second-stage of our estimation. Economic development is bound to have a positive effect on both life expectancy and schooling levels of males and of females, as it helps in correcting pre-existing gender imbalances (Lagerlof, 2003; Doepke and Tertilt, 2009).[16] As the estimation results reveal, controlling for the level of economic development does not alter our findings. The coefficient on GDP per capita is not statistically different from zero in the first-stage estimation. Moreover, its inclusion does not change the previously established pattern in the first-stage estimation with predicted mortality reductions from vaccine-preventable infectious diseases leading to relatively stronger increases in female life expectancy. In the second-stage estimation, GDP per capita is positively associated with average years of schooling. Yet, its inclusion appears to increase the quantitative importance of the effect of life expectancy on educational attainment.

In column 2, we consider the role of voting rights for women, which many authors in the literature have emphasized.[17] To control for this channel we include in our specification, a dummy variable reflecting whether women had the right to vote in the respective sample year. As we can see in Panel A, this variable is positively associated with life expectancy but its inclusion does not affect the link between mortality improvements associated with the IET and increases in life expectancy. The variable is also positively correlated with average years of schooling, as we can see in Panel B. Yet, the estimated coefficient is in this case statistically insignificant and its inclusion does not overturn the positive effect of life expectancy on educational attainment.

In column 3, we investigate the role played by changes in labor force participation of males and females over our sample period. This is motivated by the argument made by Goldin (2006) that improved labor market conditions for females increased their returns to schooling compared to men. While this mechanism may have been important in the case of some developed countries, such as the United States, in the context of our broad country-sample it does not appear to be quantitatively important. Moreover, controlling for it does not alter the patterns observed in our baseline regressions.[18]

In columns 4 and 5 we explore the role that characteristics of the marriage market played in the educational choices made by males and females. This relates to the arguments regarding the importance of education in the marriage market presented by

---

to reduce the number of included fixed effects. The exact nature of fixed effects employed in each specification is indicated in the second row from the bottom in Table 7.

[16]See also Duflo (2012) for a detailed survey of the literature.

[17]See Doepke, Tertilt and Voena (2012) for a summary of the literature.

[18]While the size of the estimated coefficients in the first and second stage differ from those shown in Table 3, this is not a consequence of controlling for labor force participation, but solely due to the change in sample size.

Chiappori et al. (2009) in the case of developed countries as well as Ashraf et al. (2016) in the case of developing countries. To capture this idea, we introduce two control variables that reflect the conditions in the marriage market and the relative status of women in a marriage. The first one is the average age at marriage of the male and female population. The second one is the average age difference at marriage between brides and grooms, which captures the relative power of women in the household, as discussed by de Moor and van Zanden (2009). Controlling for these two variables does not change our main findings regarding the link between life expectancy and schooling. The average age at marriage is positive related to life expectancy, as we see in column 4, but beyond this the inclusion of this variable has no further implications.[19]

In column 6 we assess the importance of changes in fertility behavior over our sample period. Theories about the demographic transition clearly link fertility rates with life expectancy and educational attainment.[20] With that in mind, we add fertility rates as a control in our specification. As the regression results indicate, higher fertility rates are associated with lower levels of life expectancy. Controlling for this effect, however, does not affect the main regression coefficients, neither in the first stage nor in the second stage.

While exploring alternative mechanisms that may account for the empirical patterns reflected in our regressions, it is important to consider also other factors that may have influenced male and female mortality in our sample countries but are unrelated to infectious diseases. Given our time period of interest, factors that are particularly important are improvements in maternal mortality, highlighted by Albanesi and Olivetti (2016), as well as changes in mortality from cancer and cardiovascular diseases, emphasized by Deaton (2003). In columns 7 and 8, we control for these factors using available data on maternal mortality rates and mortality from cancer and cardiovascular diseases, which exhibit a strong male bias. Controlling for maternal mortality leaves our main findings intact. Controlling for mortality rates from cancer and cardiovascular diseases changes quantitatively the magnitudes of our estimated effects, but this is due to the substantial reduction in the sample size, not to the inclusion of the control variables. Qualitatively, we continue to observe the previously established pattern that improvements in the mortality environment associated with the IET had larger effects on the life expectancy of females than of males. We also continue to find positive effects of life expectancy on educational attainment in the second-stage regression.

---

[19] Similar to the case of column 3, the change in the point estimates of the regression coefficients in columns 4 and 5 is largely driven by the changing sample size.

[20] As parents face a trade-off between the quantity and the quality of their children, they are more likely to invest in their education in a low fertility environment. At the same time, a lower level of fertility also implies that mothers have more opportunities to participate in the labor market, which in turn would increase the returns to their own human capital investment. See Murphy (2015) for evidence on the link between fertility and the education gender gap. For an overview of the theories regarding the demographic transition see Galor (2011).

Beyond the aforementioned factors, of course, there are also other factors that may have played a role in the relative rise of female educational attainment and the closing of the education gender gap, such as the introduction of the birth control pill, stressed by Goldin and Katz (2002), or the decline in the price of household appliances, highlighted by Greenwood et al. (2016). In the context of our analysis, though, the effect of these factors are largely captured by the year fixed effects, as the diffusion across most countries largely took place between 1940 and 1980.[21] Gender norms as well as other cultural and institutional attributes of countries are also expected to influence the health and educational outcomes of males and females. Yet, as these factors have been shown to be highly persistent (Alesina et al., 2013; Hansen et al., 2015), they are largely captured by our gender-country fixed effects. Moreover, as discussed in Section 5, we find no evidence that the same health improvements for women and men were associated with differential increases in their educational attainment. Thus, it is unlikely that our results are driven by factors influencing the relative returns to education of females and males conditional on experiencing the same gains in terms of health, such as differential changes in school accessibility.

7 Implications for Economic Development

So far we have established with our regression analysis that: (a) the global reductions in mortality from infectious diseases associated with the IET led to substantial improvements in life expectancy, (b) these improvements led to sizeable increases in educational attainment, and (c) females experienced larger improvements in life expectancy and as a consequence their schooling levels rose more than those of men. In this section we explore the broader implications that these mortality reductions had for economic development.

An extensive body of literature, summarized by Deaton (2003) and Weil (2014), has highlighted various channels through which improvements in health affect the process of economic development and vice versa. Yet, in the context of the IET Acemoglu and Johnson (2007) have already documented that the large improvements in life expectancy did not materialize in higher levels of GDP per capita. As these authors demonstrate, this is because the increases in aggregate GDP were offset by increases in the size of the population triggered by the lower mortality rates. Cervellati and Sunde (2011) have further explored this pattern and have shown that the effects of the life expectancy improvements on GDP per capita were not uniform across countries. In countries that had already experienced the demographic transition by 1940 and where fertility was already low, the IET led to increases GDP per capita, as the effect on the numerator of GDP

---

[21]The birth control pill, for example, is a case in point here. It was first approved by the FDA in 1957 and by 1990 the year in which we observe the educational attainment of our 1980 school cohort, it was available for use in all 75 of the countries in our sample apart from Myanmar (Finlay et al., 2012).

per capita was larger than that on the denominator. On the other hand, in countries that had not yet experienced the demographic transition the effect of life expectancy improvements on GDP per capita was largely negative.

Given our finding of non-uniform gains in life expectancy between males and females associated with the IET, it is instructive to assess whether the life expectancy improvements of males and females had differential impacts on per capita GDP. In fact, there is already some work in the literature suggesting this possibility. The models of the demographic transition proposed by de la Croix and Vander Donckt (2010) as well as Bloom et al. (2015) highlight how improvements in female health tend to be more conducive to economic development compared to improvements in male health. This is because health improvements for males and females, while they are conducive to educational attainment, have differential effects on fertility and labor force participation. An improvement in male health generates primarily an income effect on households, which increases both consumption and desired fertility. An improvement in female health, on the other hand, has beyond an income effect also a substitution effect, which tends to lower fertility and raise labor force participation. This is due to the fact that women naturally devote a larger share of their time to child-rearing activities.

To our knowledge, the existing literature has not offered an empirical test for these differential effects of male and female health improvements on economic development. This is largely due to the fact that male and female life expectancy are very highly correlated and their distinct effects cannot be estimated within the same regression. Our empirical setup, however, which links the differential improvements in male and female life expectancy to the exogenous mortality reductions from vaccine- and non-vaccine-preventable diseases implied by the initial disease environment, allows us to indirectly test the link between male and female life expectancy and economic development. Specifically, this can be done by estimating the following reduced-form specification:

$$\ln y_{ct} = b^V \sum_{d \in V} M_{dct} + b^{NV} \sum_{d \in NV} M_{dct} + d_c + d_t + e_{ct}. \tag{3}$$

Here y_{ct} denotes the level of GDP capita in country c in year t, while the terms $\sum_{d \in V} M_{dct}$ and $\sum_{d \in NV} M_{dct}$ correspond to the predicted mortality rates for the groups of vaccine-preventable (V) and non-vaccine-preventable (NV) infectious diseases. The specification includes country and year fixed effects, d_c and d_t, and is estimated in long-differences between 1940 and 1980 using the same sample of countries as in our baseline regressions.[22] Given the results of our analysis so far and the hypothesis that female health improvements are more conducive to economic development, we would expect the coefficient b^V to be

---

[22]Out of the 75 countries in our main sample, we have to drop six for which GDP per capita is not available in 1940.

smaller than b^{NV} as reductions in predicted mortality from vaccine-preventable infectious diseases would lead to larger improvements in female health than reductions in predicted mortality from non-vaccine-preventable ones.

The results of estimating equation (3) can be seen in Table 8. In column 1, we estimate first a simple specification in which we do not split the diseases into different groups, but look at how the overall reductions in predicted mortality from the 13 infectious diseases affected per capita GDP levels. The estimated coefficient is positive and statistically significant. This implies that the reductions in mortality were associated with a drop in GDP per capita and the drop was on average larger in countries where mortality fell more. This estimate, thus, mirrors the findings of Acemoglu and Johnson (2007) that global improvements in health associated with the IET did not lead to improvements in income.

[Insert Table 8 around here]

Having shown this, in column 2 we separate the overall predicted mortality reductions between those coming from the V and NV disease groups. As the estimation results indicate, there is a positive association between changes in mortality and GDP per capita for the latter group, but an insignificant association for the former group. Hence, drops in GDP per capita were on average observed in countries where the mortality reductions came primarily from non-vaccine-preventable diseases. In countries where the mortality reductions came primarily from vaccine-preventable diseases, which are the countries where on average female life expectancy rose more than male life expectancy, mortality reductions did not have a negative effect on GDP per capita.

In column 3 we explore this pattern further by separating the countries in our sample into those that had already experienced the demographic transition (DT) by 1940 and those that had not. We do so following the classification of Cervellati and Sunde (2011).[23] Starting with our simple specification of column 1 and interacting the overall change in predicted mortality with a post-DT dummy variable, we find that the effect of the mortality reductions on per capita GDP was different for pre- and post-transitional countries. While in the former group of countries mortality reductions were associated with a fall in GDP per capita, in the latter group they contributed to a rise in per capita GDP. This confirms the finding of Cervellati and Sunde (2011) that countries which had already undergone the demographic transition by 1940 experienced a rise in income levels as a result of the post-1940 improvements in the health environment.

Column 4 combines the estimation setups of columns 2 and 3 by interacting the post-DT dummy with the predicted mortality reductions coming from the V and NV

---

[23]Specifically, we classify countries as post-transitional if by 1940 life expectancy at birth exceeded 50 years, there was a sustained decline in fertility, or the crude birth rate had fallen below 30/1000.

disease groups. As the estimation results indicate, there is a negative and statistically significant association between changes in GDP per capita and changes in mortality in post-transitional countries, but this relationship only applies to mortality from vaccine-preventable diseases. For non-vaccine-preventable diseases, the association between changes in GDP per capita and changes in mortality is positive and statistically significant in both pre- and post-transitional countries. Noting that the countries where the mortality reductions were primarily driven by vaccination were also the ones where on average females gained more than males in terms of life expectancy, these results suggest the existence of a positive association between female health improvements and economic development once a country has achieved the demographic transition. On the other hand, no such positive association can be observed for male health improvements.

8 Conclusion

In this paper, we establish a connection between two important global developments that took place over the second half of the 20th century: the unprecedented improvements in life expectancy and the sharp rise in the relative educational attainment of women. In particular, we argue that these changes were to a large extent driven by the spread of western medical technologies around the globe. Following the coordinated efforts of the United Nations and the World Health Organization after World War II to promote the use of vaccines and antibiotics, both rich and poor countries witnessed large reductions in mortality rates from previously highly infectious diseases. While the effect that these developments had on life expectancy across countries is well documented, the differential nature of the life expectancy gains for males and females and their consequences for the closing of the education gender gap are a novel contribution of our analysis.

To establish this connection econometrically and determine the direction of causality, we rely on the exogenous nature of the mortality reductions experienced by individual countries. This exogeneity assumption is justified by the fact that the mortality reductions were a product of medical research in a small group of developed countries. We, furthermore, exploit biological differences in the immune response of males and females to vaccination in order to establish the differential effect that the introduction of the new medical technologies had on life expectancy across genders. As has been recently documented in the medical literature, females are better protected from a disease after vaccination since vaccine efficacy is higher for them compared to males. Hence, the improvements in the mortality environment that were achieved after World War II thanks to the introduction of new and improved vaccines were bound to be larger for females than for males.

Following this empirical strategy, we document that in countries were female life expectancy rose faster, as a result of the pattern of mortality reductions, there was also

a more rapid increase in female educational attainment. Moreover, we show that this mechanism can explain a substantial share of the decline in the education gender gap observed over this period. In additions to that, we present evidence that the countries experiencing these relatively larger increases in female life expectancy and education also saw increases in terms of their GDP per capita levels. Thus, the relatively larger health improvements for females appear to have had broader macroeconomic effects beyond raising female educational attainment.

Overall our findings suggest that public health policies, particularly those targeted at the female population, can be an important mechanism to empower women and promote gender equality. As poor health outcomes early in life may lead women to underperform in school and in the labor market later in life, policies targeted at improving these health outcomes for females can yield sizeable economic and social benefits. In this respect our analysis resonates well with an extensive literature that has documented a strong connection between the health status of children and their subsequent performance in school and in the labor market (Almond and Currie, 2011). The results of our analysis also support the view that female empowerment can have broader benefits in terms of economic development (Duflo, 2012). At the same time, though, our findings suggest that these benefits will take time before they will materialize and they may require large-scale policy commitments, such as those made by the UN and the WHO in the 1940s to eradicate infectious diseases.

References

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6): 925–985.

Albanesi, Stefania, and Claudia Olivetti. 2016. "Gender Roles and Medical Progress." *Journal of Political Economy*, 124(3): 650–695.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics*, 128(2): 469–530.

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*. Vol. 4B, , ed. David Card and Orley Ashenfelter, Chapter 15, 1315–1486. Amsterdam:Elsevier.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. 2016. "Bride Price and Female Education." NBER Working Paper No. 22417.

Baghdadi, J.E., M. Orlova, A. Alter, B. Ranque, M. Chentoufi, F. Lazrak, ..., and L. Abel. 2006. "An Autosomal Dominant Major Gene Confers Predisposition to

Pulmonary Tuberculosis in Adults." *Journal of Experimental Medicine*, 203(7): 1679–1684.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Dhruva Kothari. 2010. "Improving Immunisation Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunisation Campaigns with and without Incentives." *British Medical Journal*, 340(7759): 1291.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104: 184–198.

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topolova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science*, 335(6068): 582–586.

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares. 2005. "The Quantity and Quality of Life and the Evolution of World Inequality." *American Economic Review*, 95(1): 277–291.

Ben-Porath, Yoram. 1967. "The Production of Human Capital in the Life Cycle of Earnings." *Journal of Political Economy*, 75(4): 352–365.

Bernin, H., and H. Lotter. 2014. "Sex Bias in the Outcome of Human Tropical Infectious Diseases: Influence of Steroid Hormones." *Journal of Infectious Diseases*, 209(Suppl 3): S107–S113.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics*, 122(1): 73–117.

Bloom, David.E., Michael Kuhn, and Klaus Prettner. 2015. "The Contribution of Female Health to Economic Development." NBER Working Paper No. 21411.

Boucekkine, Raouf, David de la Croix, and Omar Licandro. 2002. "Vintage Human Capital, Demographic Trends, and Endogenous Growth." *Journal of Economic Theory*, 104: 340–375.

Cernetich, A., L.S. Garver, and A.E. Jedlicka. 2006. "Involvement of Gonadal Steroids and Gamma Interferon in Sex Differences in Response to Bloodstage Malaria Infection." *Infection and Immunity*, 74(3): 190–203.

Cervellati, Matteo, and Uwe Sunde. 2005. "Human Capital Formation, Life Expectancy, and the Process of Development." *American Economic Review*, 95(5): 1653–1672.

Cervellati, Matteo, and Uwe Sunde. 2011. "Life Expectancy and Economic Growth: The Role of the Demographic Transition." *Journal of Economic Growth*, 16(2): 99–133.

Cervellati, Matteo, and Uwe Sunde. 2015. "The Economic and Demographic Transition, Mortality, and Comparative Development." *American Economic Journal: Macroeconomics*, 7(3): 189–225.

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss. 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713.

Cook, F.I. 2008. "Sexual Dimorphism of Humoral Immunity with Human Vaccines." *Vaccine*, 26: 3551–3555.

Cutler, David M., Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives*, 20(3): 97–120.

Davila, S., M.L. Hibberd, R.H. Dass, E. Hazel, E. Wong, E. Sahiratmadja, ..., and M. Seielstad. 2008. "Genetic Association and Expression Studies Indicate a Role of Toll-Like Receptor 8 in Pulmonary Tuberculosis." *PLoS Genetics*, 4(10): e1000218.

Deaton, A. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature*, 41(1): 113–158.

de la Croix, David, and Marie van der Donckt. 2010. "Would Empowering Women Initiate the Demographic Transition in Least-Developed Countries?" *Journal of Human Capital*, 4(2): 85–129.

de la Croix, David, and Omar Licandro. 2012. "The Child is the Father of the Man: Implications for the Demographic Transition." *Economic Journal*, 123(567): 236–261.

de Moor, Tine, and Jan Luiten van Zanden. 2009. "Girl Power: The European Marriage Pattern and Labour Markets in the North Sea Region in the Late Medieval and Early Modern Period." *Economic History Review*, 63(1): 1–33.

Doepke, Matthias, and Michele Tertilt. 2009. "Women's Liberation: What's in It for Men?" *Quarterly Journal of Economics*, 124(4): 1541–1591.

Doepke, Matthias, Michele Tertilt, and Alessandra Voena. 2012. "The Economics and Politics of Women's Rights." *Annual Review of Economics*, 4: 339–372.

Duflo, E. 2012. "Women Empowerment and Economic Development." *Journal of Economic Literature*, 50(4): 1051–1079.

Fernández, R. 2013. "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century." *American Economic Review*, 103(1): 472–500.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics*, 119(4): 1249–1299.

Fernández, Raquel, and Joyce Cheng Wong. 2011. "The Disappearing Gender Gap: The Impact of Divorce, Wages, and Preferences on Education Choices and Women's Work." NBER Working Paper No. 17508.

Finlay, Jocelyn E., David Canning, and June Y. T. Po. 2012. "Reproductive Health Laws Around the World." PGDA Working Paper No. 96.

Fish, E.N. 2008. "The X-Files in Immunity: Sex-Based Differences Predispose Immune Responses." *Nature Reviews Immunology*, 8(9): 737–744.

Franconi, F., S. Brunelleschi, L. Steardo, and V. Cuomo. 2007. "Gender Differences in Drug Responses." *Pharmacological Research*, 55(2): 81–95.

Furman, D., B.P. Hejblum, N. Simon, V. Jojic, C.L. Dekker, R. Thiebaut, R.J. Tibshirani, and M.M. Davis. 2013. "Systems Analysis of Sex Differences Reveals an Immunosuppressive Role for Testosterone in the Response to Influenza Vaccination." *Proceedings of the National Academy of Sciences*, 111(2): 869–874.

Galor, Oded. 2011. *Unified Growth Theory.* Princenton, NJ:Princeton University Press.

Galor, Oded, and David N. Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review*, 86(3): 374–387.

Giefing-Kroell, C., P. Berger, G. Lepperdinger, and B. Grubeck-Loewenstein. 2015. "How Sex and Age Affect Immune Responses, Susceptibility to Infections, and Response to Vaccination." *Aging Cell*, 14(3): 309–321.

Goldin, Claudia. 1995. "The U-Shaped Female Labor Force Function in Economic Development and Economic History." In *Investment in Women's Human Capital and Economic Development.* , ed. Schultz T. P., 61–90. Chicago, IL:University of Chicago Press.

Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review*, 96(2): 1–21.

Goldin, Claudia., and Lawrence F. Katz. 2002. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy*, 110(4): 730–770.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives*, 20(4): 133–156.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2016. "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation." *American Economic Journal: Macroeconomics*, 8(1): 1–41.

Grossman, C.J. 1985. "Interactions Between the Gonadal Steroids and the Immune System." *Science*, 227(4684): 257–261.

Guerra-Silveira, F., and F. Abad-Franch. 2013. "Sex Bias in Infectious Disease Epidemiology: Patterns and Processes." *PLoS ONE*, 8(4): e62390.

Hamilton, J.B., and G.E. Mestler. 1969. "Mortality and Survival: Comparison of Eunuchs with Intact Men and Women in a Mentally Retarded Population." *Journal of Gerontology*, 24(4): 395–411.

Hansen, Casper W. 2013. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Health Economics*, 32(6): 1142–1152.

Hansen, Casper W., and Holger Strulik. 2017. "Life Expectancy and Education: Evidence from the Cardiovascular Revolution." *Journal of Economic Growth*, 22(4): 421–450.

Hansen, Casper W., P.S. Jensen, and C.V. Skovsgaard. 2015. "Modern Gender Roles and Agricultural History: The Neolithic Inheritance." *Journal of Economic Growth*, 20(4): 365–404.

Hazan, Moshe, and Hosny Zoabi. 2006. "Does Longevity Cause Growth? A Theoretical Critique." *Journal of Economic Growth*, 11(4): 363–376.

Hazan, Moshe, and Hosny Zoabi. 2015. "Sons Or Daughters? Sex Preferences And The Reversal Of The Gender Educational Gap." *Journal of Demographic Economics*, 81(2): 179–201.

Jayachandran, Seema, and Andriana Lleras-Muney. 2009. "Life Expectancy and Human Capital Investements: Evidence from Maternal Mortality Declines." *Quarterly Journal of Economics*, 124(1): 349–397.

Kalemli-Ozcan, Sebnem. 2002. "Does the Mortality Decline Promote Economic Growth?" *Journal of Economic Growth*, 7(4): 411–439.

Klein, S.L., A. Jedlicka, and A. Pekosz. 2010. "The Xs and Y of Immune Responses to Viral Vaccines." *Lancet Infectious Diseases*, 10(5): 338–349.

Klein, S. L, and A. Pekosz. 2014. "Sex-based Biology and the Rational Design of Influenza Vaccination Strategies." *Journal of Infectious Diseases*, 209(Suppl 3): S114–S119.

Klein, S.L., I. Marriott, and E.N. Fish. 2015. "Sex-based Differences in Immune Function and Responses to Vaccination." *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 109(1): 9–15.

Krucken, J., M.A. Dkhil, J.V. Braun, R.M. Schroetel, M. El-Khadragy, P. Carmeliet, H. Mossmann, and F. Wunderlich. 2005. "Testosterone Suppresses Protective Responses of the Liver to Blood-Stage Malaria." *Infection and Immunity*, 73(1): 436–443.

Lagerlof, Nils Petter. 2003. "Gender Equality and Long-Run Growth." *Journal of Economic Growth*, 8(4): 403–426.

Libert, C., L. Dejager, and I. Pinheiro. 2010. "The X Chromosome in Immune Functions: When a Chromosome Makes the Difference." *Nature Reviews Immunology*, 10: 594–604.

Lozano, R., M. Naghavi, K. Foreman, S. Lim, K. Shibuya, V. Aboyans, ..., and C.J.L. Murray. 2012. "Global and Regional Mortality from 235 Causes of Death for 20 Age Groups in 1990 and 2010: A Systematic Analysis for the Global Burden of Disease Study 2010." *Lancet*, 380(9859): 2095–2128.

Lucas, Adrienne M. 2010. "Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics*, 2(2): 46–71.

Muenchhoff, M., and P.J.R. Gouldner. 2014. "Sex Differences in Pediatric Infectious Diseases." *Journal of Infectious Diseases*, 209(Suppl 3): S120–S126.

Murphy, Tommy E. 2015. "Old Habits Die Hard (Sometimes) Can Departement Heterogeneity Tell us Something about the French Fertility Decline?" *Journal of Economic Growth*, 20(2): 177–222.

Neyrolles, O., and L. Quintana-Murci. 2009. "Sexual Inequality in Tuberculosis." *PLoS Medicine*, 6(12): 1–6.

Nhamoyebonde, S., and A. Leslie. 2014. "Biological Differences Between the Sexes and Susceptibility to Tuberculosis." *Journal of Infectious Diseases*, 209(Suppl 3): S100–S106.

OECD. 2015. *The ABC of Gender Equality in Education: Aptitude, Behavior, Confidence.* PISA, OECD Publishing.

Oster, E., I. Shoulson, and R. Dorsey. 2013. "Limited Life Expectancy, Human Capital and Health Investements: Evidence from Huntington Disease." *American Economic Review*, 103(5): 1977–2002.

Owens, P.F. 2002. "Sex Differences in Mortality Rate." *Science*, 297(5589): 2008–2009.

Poland, G.A., I. Ovsyannikova, and R.M. Jacobson. 2008. "Genetic Susceptibility to Infectious Diseases." , ed. R.A. Kaslow, J.M. McNichol and A.V.S. Hill, Chapter Genetics and Immune Response to Vaccines. New York:Oxford University Press.

Preston, Samuel H., Nathan Keyfitz, and Robert Schoen. 1972. *Causes of Death: Life Tables for National Populations.* New York, NY:Seminar Press.

Reijnders, Laurie S. M. 2018. "Wealth, Wages and Wedlock: Explaining the College Gender Gap Reversal." *Scandinavian Journal of Economics*, forthcoming.

Soares, Rodrigo R. 2005. "Mortality Reductions, Educational Attainment, and Fertility Choice." *American Economic Review*, 95(3): 580–601.

Soares, Rodrigo R. 2006. "The Effect of Longevity on Schooling and Fertility: Evidence from the Brazilian Demographic and Health Survey." *Journal of Population Economics*, 19(1): 71–97.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regressions with Weak Instruments." *Econometrica*, 65(3): 557–586.

Stanberry, L.R., S.L. Spruance, A.L. Cunnigham, D.I. Bernstein, A. Mindel, S. Sacks, ..., and G. Dubin. 2002. "Glycoprotein-D-Adjuvant Vaccine to Prevent Genital Herpes." *New England Journal of Medicine*, 347(21): 1652–1661.

Svanberg, L. 1981. "Effects of Estrogen Deficiency in Women Castrated When Young." *Acta Obstetricia et Gynecologica Scandinavica*, 106(S106): 11–15.

Tsuyuguchi, K., K. Suzuki, H. Matsumoto, E. Tanaka, R. Amitani, and F. Kuze. 2001. "Effect of Oestrogen on Mycobacterium Avium Complex Pulmonary Infection in Mice." *Clinical and Experimental Immunology*, 123(3): 428–434.

United Nations. Various Editions. "Demographic Yearbook." New York, NY.

United Nations Population Division. 2017. "World Population Prospects." esa.un.org/unpd/wpp/.

Weil, David N. 2014. "Health and Economic Growth." In *Handbook of Economic Growth*. Vol. 2B, , ed. Philippe Aghion and Steven N. Durlauf, Chapter 3, 623–682. Amsterdam:Elsevier.

World Bank. 2012. *Gender Equality and Development.* World Development Report.

Wunderlich, F., W.P. Benten, and M. Lieberherr. 2002. "Testosterone Signaling in T Cells and Macrophages." *Steroids*, 67(6): 535–538.

Yamamoto, Y., H. Saito, T. Setogawa, and H. Tomioka. 1991. "Sex Differences in Host Resistance to Mycobacterium Marinum Infection in Mice." *Infection and Immunity*, 59(11): 4089–4096.

Figure 1: Global Evolution of the Education Gender Gap
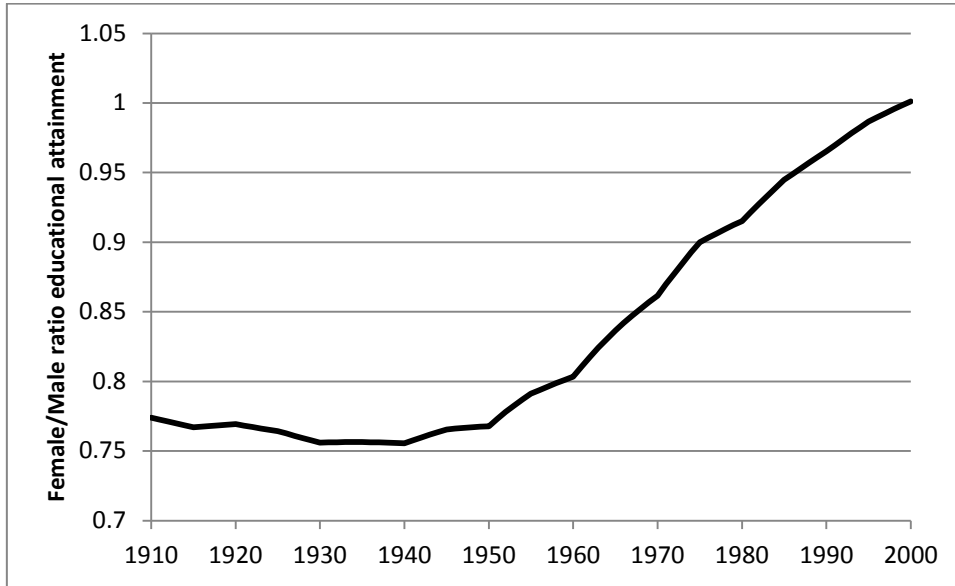
Notes: This figure depicts the global evolution of the education gender gap measured as the male-to-female ratio of average years of schooling. The data are from Barro and Lee (2013) and refer to the schooling levels of the cohort that was 15 to 19 years old in the respective year. The data series shows the simple averages across a sample of all 146 countries covered in the data set.

Figure 2: The Closing of the Education Gender Gap in Selected Countries
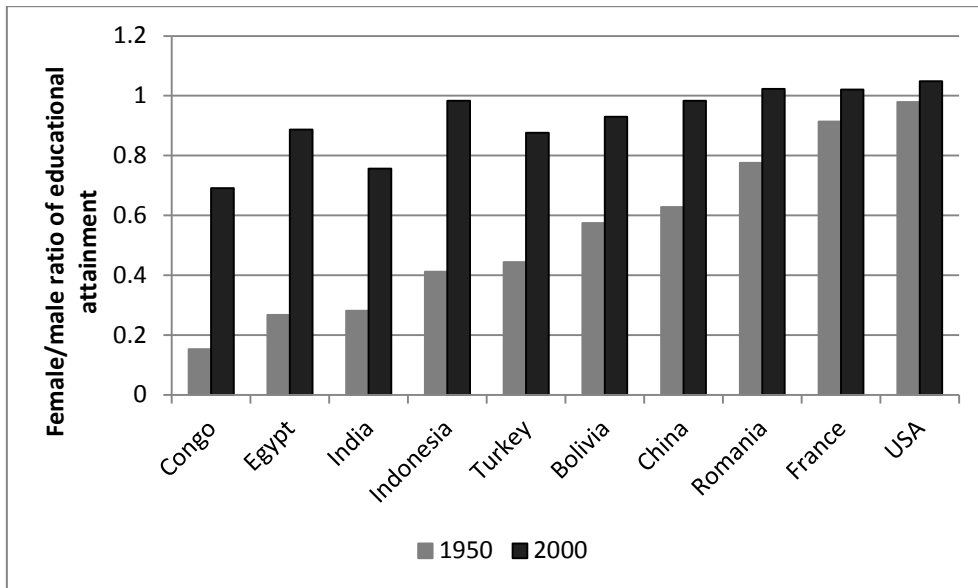
Notes: This figure depicts the change in the education gender gap measured as the male-to-female ratio of average years of schooling between 1950 and 2000 in selected countries. The data are from Barro and Lee (2013) and refer to the schooling levels of the cohort that was 15 to 19 years old in the respective year.

Table 1: Life Expectancy and Educational Attainment in Europe

Year	Life expectancy at birth, Females	Life expectancy at birth, Males	Avg. years of schooling, Females	Avg. years of schooling, Males
1900	49.8	46.9	5.6	5.9
1913	55.1	52.2	5.6	6.0
1928	59.7	56.8	6.5	7.2
1939	63.7	59.6	7.3	8.1
1950	69.9	65.9	9.0	9.8
1964	74.7	69.1	11.2	11.4
1977	77.3	70.7	12.2	12.1
1990	79.4	72.8	12.5	12.1

Notes: This table reports the levels of life expectancy at birth and average years of schooling of males and females in selected years. The average years of schooling data reflect the respective schooling levels of the cohort that was 5-9 years old in the respective year, measured 10 years later. The reported values are simple averages across a sample of 11 European countries (Belgium, Czechoslovakia, Denmark, France, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, and United Kingdom.) for which life expectancy data go back to 1900. The life expectancy data were gathered from the Human Mortality Database and the Human Lifetable Database. The educational attainment data come from Barro and Lee (2013).

Table 2: Descriptive Statistics of Key Variables

Variable	Mean	S.D.	Min	Max
Years of schooling, females, 1940	4.26	2.54	0.146	9.17
Years of schooling, males, 1940	4.54	2.51	0.144	11.16
Years of schooling, females, 1980	7.71	2.17	3.13	11.61
Years of schooling, males, 1980	7.69	1.94	3.21	11.44
Mortality, overall, 1940	0.436	0.274	0.033	1.126
Mortality, vaccine group, 1940	0.363	0.241	0.03	0.909
Mortality non-vaccine group, 1940	0.064	0.14	0	0.853
Life expectancy at birth, females, 1940	49.61	12.17	28.44	70.72
Life expectancy at birth, males, 1940	46.8	11.19	25.46	66.97
Life expectancy at birth, females, 1980	69.64	7.5	51.51	79.6
Life expectancy at birth, males, 1980	64.07	6.52	48.64	73.56

Notes: This tables presents the descriptive statistics for the key variables in our analysis. The statistics reported refer to our main sample of 75 countries that we use in our baseline regression analysis. A precise definition of each variable and the exact data sources is provided in Section B of the Appendix.

Table 3: Life Expectancy and Educational Attainment: Baseline 2SLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1st stage results			Life expectancy at birth			
Mortality, overall	-15.39***	-12.938***				
	[2.108]	[2.167]				
Female x Mortality, overall		-4.900**				
		[2.154]				
Mortality, vaccine group			-11.538***	-13.606***	-11.211***	
			[2.41]	[2.842]	[2.465]	
Female x Mortality, vaccine group			-4.959**	-5.847**	-5.139*	
			[2.440]	[2.877]	[2.648]	
Mortality, non-vaccine group			-18.534***	-18.839***	-18.120***	
			[4.588]	[4.663]	[4.597]	
Female x Mortality, non-vaccine group			-5.618	-5.711	-6.008	
			[6.800]	[6.911]	[6.776]	
Panel B: 2nd stage results			Years of schooling			
Life expectancy	0.117***	0.115***	0.114***	0.114***	0.112***	0.100***
	[0.0444]	[0.0426]	[0.0415]	[0.0415]	[0.0426]	[0.0324]
Observations	300	300	300	300	296	300
Countries	75	75	75	75	74	75
1st stage F-statistic	77.35	43.94	22.64	22.64	21.48	-

Notes: This table presents our baseline estimates for the relationship between changes in life expectancy and educational attainment measured in terms of average years of schooling between 1940 and 1980. Columns (1) to (5) present 2SLS estimates. Column (6) shows the corresponding OLS estimate. Panel A shows the results of the first-stage estimation where life expectancy is instrumented by predicted mortality rates from different groups of infectious diseases, while Panel B shows the results of the second-stage estimation. Columns (1) to (3) employ the baseline mortality instrument where the 1980 mortality rates are set to zero in all countries. Column (4) assumes that mortality rates fell proportionately in line with the global average. In column (5) we assume that mortality rates fell to the level observed in the U.S. in 1980 (the U.S. is dropped from the sample for this reason). All regressions include country-sex and year fixed effects. Heteroskedasticity robust standard errors clustered at the country-sex level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 4: Robustness Checks for First-Stage Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Life expectancy at birth			
Source of Mortality	Malaria	Pneumonia	Tuberculosis	Minor diseases	Bacterial diseases	Average mortality	Sex-specific mortality
Mortality	-16.101***	-14.01***	3.808				
	[5.179]	[3.78]	[4.814]				
Female x Mortality	-8.882	-9.106**	-11.813**				
	[7.207]	[4.105]	[5.414]				
Mortality, V-group				-9.863	-9.969***	2.221	3.013
				[25.31]	[2.816]	[17.832]	[16.603]
Female x Mortality, V-group				-45.97	-5.531*	-5.502***	-7.512***
				[31.64]	[3.250]	[1.124]	[1.670]
Mortality, NV- group				-39.41	-43.496*		
				[31.30]	[25.792]		
Female x Mortality, NV- group				-8.47	-13.75		
				[48.97]	[39.254]		
Residual Mortality	-14.105***	-13.319***	-20.38***	-15.22***	-20.258***	-65.448*	-64.957*
	[2.158]	[3.391]	[2.46]	[2.514]	[4.408]	[35.771]	[32.315]
Observations	300	300	300	300	300	88	88
Countries	75	75	75	75	75	22	22

Notes: This table explores the robustness of our first-stage estimates to using different groupings of the mortality rates from the infectious diseases and to employing sex-specific mortality data, which are available for a limited sample of 22 countries. The specific disease or sub-group of diseases a regression focuses on is indicated at the top of each column. The exact details and the rationale behind the different groupings are explained in the text. All regressions include country-sex and year fixed effects, except for col. 6 and 7 where we employ separate fixed effects for country, sex and year due to the small sample size. Heteroskedasticity robust standard errors clustered at the country-sex level (col. 1-5) or country-level (col. 6, 7) are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 5: Robustness Checks for Sub-Samples and Life Expectancy at Higher Ages

	(1)	(2)	(3)	(4)	(5)	(6)
	Omit countries involved in drug development	Omit African & Middle Eastern countries	Omit small Latin American countries	Omit Eastern European countries	Life expectancy at age 5	Life expectancy at age 20
Panel A: 1st stage results	Life expectancy at birth				LE at age 5	LE at age 20
Mortality, V-group	-10.059***	-11.44***	-11.692***	-11.672***	-4.767*	-3.765
	[2.451]	[2.555]	[2.507]	[2.504]	[2.694]	[2.323]
Female x Mortality, V-group	-4.853*	-5.154**	-4.852**	-4.623*	-5.551**	-3.246
	[2.453]	[2.654]	[2.437]	[2.580]	[2.678]	[2.216]
Mortality, NV-group	-17.218***	-18.758***	-21.431***	-18.093***	-22.467*	-17.001*
	[4.665]	[4.588]	[5.950]	[4.692]	[12.496]	[9.468]
Female x Mortality, NV-group	-5.704	-5.341	-6.569	-5.858	-1.019	-1.003
	[6.856]	[6.811]	[9.123]	[6.854]	[14.961]	[11.138]
Panel B: 2nd stage results	Years of schooling					
Life expectancy	0.074*	0.115***	0.138***	0.0991**	0.193***	0.292***
	[0.0400]	[0.0425]	[0.0417]	[0.0429]	[0.0651]	[0.0944]
Observations	280	272	280	268	204	212
Countries	70	68	70	67	51	53
1st stage F-statistic	18.69	20.35	22.88	19.26	11.6	7.16

Notes: This table explores the robustness of our second-stage estimates to changes in the sample composition and employing data on life expectancy at higher ages. Col. 1 omits France, Germany, Switzerland, the United Kingdom and the United States. Col. 2 omits Algeria, Egypt, Iran, Iraq, Morocco, South Africa, Tunisia. Col. 3 omits Barbados, Belize, Guyana, Jamaica, Trinidad and Tobago. Col. 4 omits Bulgaria, Czech Republic, Hungary, Poland, Romania, Russian Federation, Slovakia, Yugoslavia. All regressions include country-sex and year fixed effects. Heteroskedasticity robust standard errors clustered at the country-sex level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 6: Robustness Checks with Alternative Measures of Educational Attainment and Accounting for Time Trends

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1st stage results	Life expectancy at birth					
Mortality, vaccine group	-11.538***	-11.538***	-11.538***	-11.538***	-11.52***	-11.538***
	[2.41]	[2.41]	[2.41]	[2.41]	[2.379]	[2.41]
Female x Mortality, vaccine group	-4.959**	-4.959**	-4.959**	-4.959**	-4.544*	-4.959**
	[2.440]	[2.440]	[2.440]	[2.440]	[2.418]	[2.440]
Mortality, non-vaccine group	-18.534***	-18.534***	-18.534***	-18.534***	-18.282***	-18.534***
	[4.588]	[4.588]	[4.588]	[4.588]	[4.742]	[4.588]
Female x Mortality, non-vaccine group	-5.618	-5.618	-5.618	-5.618	-5.628	-5.618
	[6.800]	[6.800]	[6.800]	[6.800]	[7.11]	[6.800]
Lagged years of schooling					0.582*	
					[0.344]	
Panel B: 2nd stage results	Schooling 1945/1985	Primary schooling	Secondary schooling	Tertiary schooling	Control for lagged schooling	Lagged schooling as DV
Life expectancy at birth	0.114***	0.0692***	0.0378**	0.0561*	0.122***	0.0317
	[0.0421]	[0.0123]	[0.0152]	[0.0297]	[0.0402]	[0.0235]
Lagged years of schooling					-0.230	
					[0.153]	
Observations	300	300	300	300	300	300
Countries	75	75	75	75	75	75
1st stage F-statistic	22.64	22.64	22.64	22.64	21.86	22.64

Notes: This table explores the robustness of our second-stage estimates to employing alternative measures of a cohort's level of human capital and accounting for time trends in the education data. All regressions include country-sex and year fixed effects. Heteroskedasticity robust standard errors clustered at the country-sex level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 7: Robustness Checks for Aternative Explanations for the Rise in Female Schooling

Additional control variable	Log GDP per capita	Voting rights	Labor force participation	Marriage age	Age difference at marriage	Fertility rate	Maternal mortality	Cancer/Cardio-vascular diseases
Panel A: 1st stage results				Life Expectancy at Birth				
Mortality, V-group	-15.670***	-9.251***	-11.671*	-21.125**	-19.633**	-11.127***	-12.589***	-10.191
	[2.941]	[2.237]	[6.231]	[9.194]	[8.499]	[2.575]	[2.508]	[9.137]
Female x MortalityV-group	-5.448*	-4.959**	-4.571***	-4.892***	-5.954***	-4.945*	-5.087*	-3.942***
	[3.033]	[2.463]	[1.118]	[0.910]	[1.022]	[2.591]	[2.168]	[1.427]
Mortality, NV-group	-5.060	-14.095***	-96.745**	-11.444	-8.373	-19.319***	-22.36***	-455.46**
	[13.328]	[4.839]	[46.318]	[38.867]	[37.659]	[4.868]	[4.359]	[226.55]
Female x Mortality, NV-group	-2.753	-5.618	-11.427	-2.360	-6.769	-6.368	-7.330	37.308
	[19.317]	[7.719]	[12.074]	[5.816]	[6.550]	[7.432]	[6.298]	[56.073]
Additional control variable	-1.099	4.982***	0.024	0.634*	0.692	-0.183**	0.033	-0.012**
	[1.426]	[0.850]	[0.036]	[0.340]	[0.462]	[0.079]	[0.259]	[0.0057]
Panel B: 2nd stage results				Years of schooling				
Life expectancy at birth	0.136***	0.095*	0.212***	0.155***	0.148***	0.194***	0.155***	0.292***
	[0.039]	[0.053]	[0.047]	[0.050]	[0.053]	[0.029]	[0.029]	[0.090]
Additional control variable	1.214*	0.578	-0.01	0.112	0.046	0.0002	-0.085	0.0036
	[0.469]	[0.500]	[0.009]	[0.123]	[0.173]	[0.023]	[0.091]	[0.0023]
Observations	268	300	176	160	160	256	252	84
Countries	67	75	44	40	40	64	63	21
Fixed effects	ctry-sex, year	ctry-sex, year	ctry, sex, year	ctry, sex, year	ctry, sex, year	ctry-sex, year	ctry-sex, year	ctry, sex, year
1st stage F-stat.	20.91	16.83	28.44	26.11	21.05	18.13	22.23	8.98

Notes: This table explores the robustness of our main results when considering alternative explanations for the rise in female schooling. The regressions include the indicates fixed effects. Heteroskedasticity robust standard errors clustered at the country-sex or country level (in correspondence with the employed fixed effects) are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 8: Effects of Female and Male Health Improvements on per Capita GDP

Dependent Variable:	Log GDP per capita			
	(1)	(2)	(3)	(4)
Mortality, overall	0.530***		0.404***	
	[0.153]		[0.144]	
Mortality x post, overall			-0.941***	
			[0.197]	
Mortality, V- group		0.213		0.216
		[0.235]		[0.212]
Mortality x post, V- group				-0.933***
				[0.297]
Mortality, NV- group		2.082**		1.329*
		[0.962]		[0.753]
Mortality x post, NV- group				3.264
				[13.864]
Observations	138	138	138	138
Countries	69	69	69	69
Within R-squared	0.87	0.87	0.89	0.89

Notes: This table presents reduced-form estimates of the effects of the predicted mortality reductions from different groups of infectious diseases on GDP per capita between 1940 and 1980. All regressions include country and year fixed effects. Heteroskedasticity robust standard errors clustered at the country level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

The International Epidemiological Transition
and the Education Gender Gap

Online Appendix

Mariko J. Klasing   Petros Milionis

University of Groningen[*]

March 2018

This appendix includes supplementary material to the paper and is meant to be published online. Section A provides a list of the countries included in our samples. Section B outlines the data sources and descriptions for all the variables that we employ in our analysis. Section C presents the information based on which the classification of the infectious diseases into the vaccine-preventable and non-vaccine-preventable groups were made.

---
[*]Faculty of Economics & Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail Addresses: m.j.klasing@rug.nl, p.milionis@rug.nl.

Section A: List of Countries in our Sample

Main Sample

Our main sample consists of the following 75 countries:

Algeria, Argentina, Australia, Austria, Bangladesh, Barbados, Belgium, Belize, Bolivia, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Germany, Greece, Guatemala, Guyana, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Italy, Jamaica, Japan, Korea, Malaysia, Mauritius, Mexico, Morocco, Myanmar, Netherlands, New Zealand, Nicaragua, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Portugal, Poland, Romania, Russian Federation, Slovakia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Taiwan, Thailand, Trinidad and Tobago, Tunisia, U.S.A., United Kingdom, Uruguay, Venezuela, Vietnam, Yugoslavia.

Small Sample

Our small sample consists of the following 22 countries for which we have gender-specific mortality rates in 1940 from infectious diseases as well as from cancer and cardiovascular diseases:

Australia, Canada, Chile, China, Czech Republic, Denmark, France, Greece, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Slovakia, South Africa, Spain, Sweden, Switzerland, Taiwan, U.S.A., United Kingdom.

Section B: Data Description and Sources

This section provides details on all the variables that we employ in our empirical analysis. It explains how each variable is constructed and based on which sources.

Average years of schooling for each gender and age-cohort are taken from Barro and Lee (2013). We focus on the cohorts of individuals that were between 5 and 9 years old in the respective sample years, measured 10 years later. Thus, the 1940 values reflect the schooling levels of the 15-19 years old in 1950, while the 1980 values reflect the schooling levels of the 15-19 years olds in 1990.

Primary schooling completion rates are also from Barro and Lee (2013). They correspond to the share of a given schooling cohort that eventually completed primary education. The respective cohorts we focus on are the individuals that were between the age of 5 and 9 years in 1940 and 1980 respectively, measured 20 years later at the age 25 to 29 years.

Secondary schooling completion rates are also from Barro and Lee (2013). They correspond to the share of a given schooling cohort that eventually completed secondary education. The respective cohorts we focus on are the individuals that were between the age of 5 and 9 years in 1940 and 1980 respectively, measured 20 years later at the age 25 to 29 years.

Tertiary schooling completion rates are also from Barro and Lee (2013). They correspond to the share of a given schooling cohort that eventually completed tertiary education. The respective cohorts we focus on are the individuals that were between the age of 5 and 9 years in 1940 and 1980 respectively, measured 20 years later at the age 25 to 29 years.

Life expectancy at birth in 1940 is primarily based on information reported in the 1948, 1949-1950, 1951 editions and the retrospective section of the 1967 edition of the Demographic Yearbook published by the United Nations. In all cases we use the data reported in the most recent edition. These data are supplemented with the information provided by Preston (1975) and by the Federal Security Agency in their Summary of International Vital Statistics 1937-44. The respective information for 1980 was obtained from the electronic version of World Population Prospects of the UN Population Division. As in some cases the available data are not exactly for the year 1940, we obtain an estimate for 1940 by linearly interpolating the data between the closest year before and after 1940. For countries that were already involved in World War II by 1940, we use life expectancy data for the closest available year prior to and after their involvement in the war and interpolate the value for 1940. This way our life expectancy figures in 1940 should not reflect the effect of the war but the normal mortality environment in each country.

Life expectancy at age 5 in 1940 and 1980 are drawn from the same sources as the life expectancy at birth data. These are specifically the UN Demographic Yearbooks for 1940 and the World Population Prospects database for 1980. Missing values are also interpolated based on the same procedure.

Life expectancy at age 20 in 1940 and 1980 are drawn from the same sources as the life expectancy at birth data. These are specifically the UN Demographic Yearbooks for 1940 and the World Population Prospects database for 1980. Missing values are also interpolated based on the same procedure.

Mortality rates in 1940 by disease are obtained from Acemoglu and Johnson (2007), who report mortality rates separately for 13 infectious diseases. For the year 1980, we do not use actual mortality rates, but predicted ones. Following Acemoglu and Johnson (2007) we take the 1980 mortality rates from all infectious diseases to be zero. This way our mortality variable does not reflect the actual changes in mortality rates in each country, but the predicted changes resulting from the introduction of the new medical technologies that provided effective control of infectious diseases. We furthermore construct two alternative variants of this variable where the predicted mortality rates in 1980 are based either on U.S. mortality rates in 1980 or on the country-specific rates in 1940 scaled down by the average rate of decline in mortality from the respective diseases observed at the global level.

Overall mortality correspond to the sum of the predicted mortality rates from the 13 infectious diseases. In our baseline setup this value is zero in 1980 by construction.

Mortality from vaccine-preventable diseases (V-group) corresponds to the sum of the mortality rates from the following seven diseases: diphtheria, influenza, measles, pneumonia, smallpox, tuberculosis, whooping cough. In our baseline setup this value is zero in 1980 by construction.

Mortality from non-vaccine-preventable diseases (NV-group) corresponds to the sum of the mortality rates from the following six diseases: cholera, malaria, plague, scarlet fever, typhoid fever, typhus. In our baseline setup this value is zero in 1980 by construction.

Mortality from the 10 remaining diseases correspond to the sum of the mortality rates from the following ten diseases: cholera, diphtheria, influenza, measles, plague, scarlet fever, smallpox, typhoid fever, typhus, whooping cough. In our baseline setup this value is zero in 1980 by construction.

Mortality from bacterial diseases correspond to the sum of the mortality rates from the following nine diseases: cholera, diphtheria, plague, pneumonia, scarlet fever, typhoid fever,

tuberculosis, typhus, whooping cough. In our baseline setup this value is zero in 1980 by construction.

Gender-specific mortality rates in 1940 are from Preston et al. (1972). They are available for the above-mentioned 22 countries of the small sample. The authors provide information on 11 distinct mortality causes, not all of which are infectious diseases. Based on this information we construct mortality rates for a group of vaccine-preventable infectious diseases that includes influenza, pneumonia and tuberculosis and a residual group of infectious diseases. As the reported mortality rates do not always correspond exactly to the year 1940, we interpolate the missing values between the closest year available before and after 1940 assuming a linear trend. For 1980 we assume again that the mortality rates for both groups of diseases were zero.

GDP per capita data are taken from the Maddison Project. For the countries for which data are not available in 1940 we interpolate the missing values assuming a constant growth rate between the years for which data are available.

Women's voting rights are measured with a dummy variable which is coded as 1 if women had the right to vote in a given country and year and zero otherwise. The coding of this variable is based on the information provided in Ramirez et al. (1997) who document for 133 countries between 1890 and 1990 when women acquired suffrage.

Labor force participation rates for 1940 correspond to the activity rates relative to the overall population and come from the 1948 and the 1949-1950 editions of the UN Demographic Yearbook. In cases where data for the year 1940 are not available, we use the data for the year closest to 1940. In most cases, this is the year 1941. The rates for 1980 are based on information reported in the electronic database of labour statistics of the International Labour Organization (ILOSTAT). As the database reports activity rates relative to the working-age population, we make these rates comparable to the 1940 rates by converting them to shares of the overall population. We do so by using the corresponding population data from the World Population Prospects database.

Average age at marriage data for the male and female population are constructed based on the information reported in the UN Demographic Yearbook on the number of marriages by age of the bride and groom. The 1940 data are based on information reported in the 1948, 1949-50 and 1958 editions of the Yearbook, while the 1980 values are based on the 1982 edition. If information in the exact years of interest is not available, we use the data in the years closest to 1940 and 1980 respectively. As the data are typically reported for 5-year age groups, running from the age of 15-19 years onward, we take the mid point of each age bracket and calculate the average age based on that. For marriages that take place in the last age group, 60 plus, we assume that they happen at the age of 62.

Average age difference at marriage between men and women is calculated as the difference between the average age of the bride and that of the groom. The underlying data are the average age at marriage data described above.

Fertility rates are measured as the number of life births per 1,000 individuals. The data for 1940 come from the 1949-1950 edition of the UN Demographic Yearbook. The 1980 values are taken from the 1981 edition of the UN Demographic Yearbook.

Maternal mortality rates in 1940 are taken from World Health Organization (1951), Federal Security Agency (1947) and the 1951 and 1957 editions of the UN Demographic Yearbook. In case of revisions in the UN data, we use the most recently reported data. Missing values are interpolated using the two closest data points before and after 1940 assuming a linear time trend. The 1980 mortality rates are taken from the 1981, 1982, 1983 and 1985 editions of the UN Demographic Yearbooks. Data are available for 63 out of the 75 countries in our data set and are by definition zero for the male population.

Mortality rates from cancer and cardiovascular diseases in 1940 are taken from Preston et al. (1972) and are available for 21 countries. Missing values are interpolated using the two closest data points before and after 1940 assuming a linear time trend. The corresponding 1980 figures are taken from the 1981, 1982, 1983 and 1985 editions of the UN Demographic Yearbook.

The **Post-demographic transition dummy** is coded based on the criteria provided by Cervellati and Sunde (2011). Specifically a country is classified as having gone through the demographic transition if by 1940 life expectancy at birth exceeded 50 years, there was a sustained decline in fertility, or the crude birth rate had fallen below 30/1000. Based on this classification the post transitional countries consist of Argentina, Australia, Austria, Belgium, Bulgaria, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Spain, Sweden, Switzerland, U.S.A., United Kingdom, Uruguay, Yugoslavia.

Section C: Disease Classification

As part of our analysis we classify the different infectious diseases into two groups: a group of vaccine-preventable diseases (*V-group*) and a group of non-vaccine-preventable diseases (*NV-group*). We perform this classification based on specific information regarding the nature of each disease and the medical breakthroughs related to its treatment that took place as part of the international epidemiological transition.

Cholera: Cholera is an acute diarrheal disease caused by the bacterium Vibrio cholerae and is spread by drinking contaminated water or eating contaminated food. The first effective vaccine against cholera was developed in 1896 by Wilhelm Kolle. However, the vaccines available since then, including the modern ones developed in the 1990s, provide only limited immunity. This is typically for no more than one year (Parish, 1965; WHO, 2010). The most important mechanism for controlling cholera is the provision of clean drinking water and improved sanitation. Since cholera-induced diarrhea results in severe dehydration within 3-4 hours, the development of more effective oral rehydration therapy in the 1950s provided a major breakthrough in the treatment of cholera. Furthermore, antibiotics developed in the 1940s shorten the course of the disease and reduce the symptoms (Kiple, 2003). We, thus, code cholera as a non-vaccine-preventable disease given the ineffectiveness of the available vaccines.

Diphtheria: Diphtheria is caused by the Corynebacterium diphtheriae and is spread by air or touch. Kitasano and Behring developed in 1890 the first antitoxin from the serum of animals that had been infected with the toxins secreted by the bacilli. They showed that this antitoxin could be used to treat diphtheria patients. In 1913 Behring reported first trials of successful immunization with toxin-antixtoxin mixtures (Baker and Katz, 2004). However, even though toxin-antitoxin mixtures had been successfully used for vaccination in some U.S. cities in the 1920s, it was generally difficult to employ these mixtures. This was because it was unclear how toxins and antitoxins should be balanced to induce an effective and at the same time safe immunization (Kiple, 2003). The key event in the development of safe vaccines was the development of a chemically modified toxin, called toxoid, by Leon Ramon of the Pasteur Institute. By using heat and formalin, the toxin's toxicity was reduced without affecting its immunizing power. This vaccine was approved for children in France in 1927 and made compulsory for children in France in 1938 (Parish 1965), but in other countries the vaccine was only adopted much later. Mass immunization at the global scale did not start until after World War II and after the licensing of the combined diphtheria-tetanus vaccine in 1947 (Baker and Katz, 2004). Large-scale vaccinations were further improved and facilitated with the development of the jet injector in the 1960s and the invention of freeze drying during WWII, which greatly increased the storability of vaccines and facilitated vaccine transportation (Kiple, 2003). Given that large-scale immunization campaigns outside of France did not start before the mid 1940s, we code diphtheria as a

vaccine-preventable disease.

Influenza: Influenza is caused by three strands of the influenza virus, types A, B, and C. It is passed by breath-born droplets, such as sneezing and coughing. The type A virus was first isolated and identified in 1934 and type B in 1940 (Kiple, 2003). The first vaccine was developed during World War II to protect soldiers and production was expanded to civilian markets in 1946 (Hoyt, 2006). While the development of antibiotics in the 1940s also played a role in reducing mortality from secondary bacterial infections, the primary control mechanism for influenza is vaccination. Hence, we place influenza in the group of vaccine-preventable diseases.

Malaria: Malaria is caused by parasites transmitted by the female mosquito of the Anopheles type. The causal parasites were identified and the role of the Anopheles mosquito established in the 1890s. Since then the strategy generally applied to control malaria has been the control of mosquitoes. This became much more effective with the discovery of DDT residual insecticides in the late 1940s (Kiple, 2003). Effective vaccines still do not exist to this date. Given this, we assign malaria to the group of non-vaccine-preventable diseases.

Measles: Measles are a viral infection, principally occurring among children. The disease is transmitted by air and through touch. Measles vaccines were first produced in 1958 and licensed in 1963. A single dose of the measles virus vaccine provides life-long immunity in over 95% of cases (Kiple, 2003). First steps in the development of the measles vaccine were the discovery that the disease was caused by a virus in 1911. Yet, the isolation of the measles virus did not occur until 1954. Thus, we assign measles to the group of vaccine-preventable diseases.

Plague: Plague is caused by the bacterium Yersinia pestis and is transmitted to humans through the bite of infected fleas. First vaccines were developed in the late 1890s, but the exact nature of the disease and the role of the flea in transmitting the disease from infected rats and other wild rodents were not fully studied until 1914 (Parish, 1965). After this was established, the elimination of rats and fleas became the key step in the prevention of plague. While vaccines developed in the 1890s were effective in reducing incidences of plague and plague mortality, protection only lasted for a few months (Parish, 1965). With the development of antibiotics, especially streptomycin, in the 1940s, vaccination became of secondary importance (Kiple, 2003). Given the limited protection vaccines afforded and the fact that they had been available since the 1890s, we code plague as a non-vaccine-preventable disease.

Pneumonia: Pneumonia is caused by a variety of infectious agents, including viruses and bacteria. For viral pneumonia, viral vaccines, in particular influenza vaccines developed in the 1940s, are effective in preventing infections (Ruuskanen et al., 2011). These became available for civilian markets in 1946 (Hoyt, 2006). For bacterial pneumonia, which is caused by the bacterium

8

Streptococcus pneumoniae, both vaccines and antibiotics have been used for disease control. Specifically, the bacterium was first isolated by Louis Pasteur in 1880 and the identification of other causative agents followed gradually. Vaccines against bacterial pneumonia became first available 1945, but since the discovery of penicillin and other antibiotics they were not widely used and not further developed (Kiple, 2003). This is because antibiotics were more effective than the original vaccines for treating and reducing mortality from bacterial pneumonia. Vaccines against bacterial pneumonia, however, were further developed in the late 1970s and early 1980s. This resulted in the licensing of more effective vaccines in 1977 and 1983, the final breakthrough being the development of pneumococcal conjugate vaccines in 2000 (Klein and Plotkin, 2007). We assign pneumonia to the vaccine-preventable diseases, as key vaccines became available in the 1940, albeit only partially effective ones.

Scarlet Fever: Scarlet fever is caused by certain strains of group A Streptococcus bacteria. The bacteria were first isolated in 1887. Efforts to produce a streptococcus vaccine followed, but the results were not convincing. To this date no effective vaccine exists. However, the disease is effectively treated with antibiotics, including penicillin, developed in the 1940s (Kiple, 2003). Given this, we assign scarlet fever to group of non-vaccine-preventable diseases.

Smallpox: Smallpox was an infectious disease caused by the viruses Variola major and Variola minor. An effective vaccine was developed in 1798 by Edward Jenner, which was derived from cowpox, a related disease found among animals (Kiple, 2003). However, large-scale immunization was not possible until the 1950s when free-drying, invented in the 1940s, was adopted for mass production of vaccines (Fenner, 2011). Until then the storability of the vaccine was very limited and temperature fluctuations frequently caused the vaccine to lose its potency. The invention of the jet injector in the 1960s was another important development that significantly facilitated the administration of the vaccine. The combination of these technologies resulted in the eradication of the disease by the late 1970s. We, therefore, assign smallpox to the vaccine-preventable diseases.

Tuberculosis: Tuberculosis is caused by the bacterium Mycobacterium tuberculosis and is transmitted through air. The bacterium was discovered in 1882 by Robert Koch. The first effective vaccine was developed by the French Pasteur Institute and became available for use in 1921 (Parish, 1965). This is the BCG (Bacille Calmette Guerin) vaccine which is still in use today. However, mass immunization did not start until after the late 1940s when UNICEF in cooperation with the Danish and Swedish Red Cross teams started mass immunization campaigns (Kiple, 2003). Until then, the vaccine had been primarily used in France where by 1933 about 20% of all children were vaccinated. Outside of France, though, only 500,000 children had been vaccinated at that time (Parish, 1965). This changed with the mass post-war immunization campaigns. In 1948-1951 alone 30 million people were tuberculin tested and 14 million were

vaccinated. The campaign was extended further under a joint UNICEF/WHO initiative which led to the vaccination of an additional 60 million people by 1956 (Comstock, 1994). Given the sensitivity of the BCG vaccine to sunlight and temperature fluctuations and the fact that it could only be stored for 10 days after initial preparation (Bonah, 2005), the invention of freeze-drying in the 1950s provided another major step forward in the fight against tuberculosis (Parish, 1965). It allowed for the vaccine to be stored under various climatic conditions for prolonged periods of time and to be transported easily over long distances. In addition, drugs developed in the 1940s and 1950s, in particular streptomycin in 1944, provide effective treatment and made most cases of the disease curable. Given that mass immunization campaigns did not start before the late 1940s, we code tuberculosis as a vaccine-preventable disease.

Thyphoid Fever: Typhoid fever is caused by the bacterium Salmonella typhi and is spread through contaminated water or food. Vaccines were developed before World War I, but they were not very effective. With the introduction of antibiotics in the 1940s, which provided effective treatment, the use of vaccines became of secondary importance in the control of the disease (Kiple, 2003). Hence, we code typhoid fever as a non-vaccine-preventable disease given that vaccines had already been available long before 1940 and were not very effective.

Typhus: Typhus is caused by Rickettsia bacteria transmitted by the human body louse Pediculus humanus corporis. The role of the louse was established in 1909 and the bacterium first demonstrated in 1916 (Parish, 1965). Public health measures such as good hygiene and sanitation played an important role in the control of the disease. First attempts in developing a vaccine to be used in humans were made in the 1930s. This was in the form of the Weigl vaccine, which used killed bacteria, and the Cox vaccine, which used live bacteria (Lindenmann, 2002). They were albeit difficult to manufacture and turned out to not be very effective. With the introduction of antibiotics, which provides effective treatment for typhus, and the development of DDT, which allowed for very efficient delousing, vaccines fell into disuse (Lindenmann, 2002). We assign typhus to the non-vaccine-preventable diseases given that vaccines were already available before 1940 and not much used after the introduction of broad-spectrum antibiotics.

Whooping Cough: Whooping cough, or pertussis, is caused by the bacterium Bordetella pertussis and occurs primarily among young children. The causative agent was first isolated in 1900 and the first effective vaccine was developed in the 1930s. There had been earlier attempts already in the 1920s to develop a pertussis vaccine. However, these early vaccines had very limited efficacy and became obsolete with the introduction of the vaccines developed by Kendrick and Eldering (Parish, 1965; Granstroem, 2011). First large-scale clinical trials were conducted in the U.S. in the late 1930s and early 1940s, which demonstrated the effectiveness and safety of the vaccine (Parish, 1965). Following that, the American Academy of Pediatrics approved the vaccine for routine use in 1943. However, it was not until the later 1940s that pertussis

vaccines were widely used when the combination vaccine with diphtheria and tetanus, DPT, was licensed (Parish, 1965). Therefore, we assign whooping cough to the group of vaccine-preventable diseases.

References

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6): 925–985.

Baker, J.P., and S.L. Katz. 2004. "Childhood Vaccine Development: An Overview." *Pedriatric Research*, 55(2): 347–356.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104: 184–198.

Bonah, C. 2005. "The 'Experimental Stable' of the BCG Vaccine: Safety, Efficacy, Proof, and Standards, 1921-1933." *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(4): 696–721.

Cervellati, Matteo, and Uwe Sunde. 2011. "Life Expectancy and Economic Growth: The Role of the Demographic Transition." *Journal of Economic Growth*, 16(2): 99–133.

Comstock, G.W. 1994. "The International Tuberculosis Campaign: A Pioneering Venture in Mass Vaccination and Research." *Clinical Infectious Diseases*, 19(3): 528–540.

Federal Security Agency. 1947. *Summary of International Vital Statistics, 1937-1944.* Washington, D.C.:U.S. Public Health Services, National Office of Vital Statistics.

Fenner, Frank. 2011. "Smallpox Eradication: the Vindication of Jenner's Prophecy." In *History of Vaccine Development.* , ed. Stanley A. Plotkin, Chapter 5, 27–32. New York, NY:Springer.

Granstroem, M. 2011. "The History of Pertussis Vaccination: From Whole-Cell to Subunit Vaccines." In *History of Vaccine Development.* , ed. S. Plotkin and B. Fantini. Paris:Elsevier.

Hoyt, K. 2006. "Vaccine Innovation: Lessons from World War II." *Journal of Public Health Policy*, 27(1): 38–57.

International Labour Organization. 2017. "Database of Labour Statistics (ILOSTAT)." www.ilo.org/ilostat.

Kiple, Kenneth F., ed. 2003. *The Cambridge Historical Dictionary of Diseases.* Cambridge:Cambridge University Press.

Klein, J.O., and S.A. Plotkin. 2007. "Robert Austrian: 1917-2007." *Clinical Infectious Diseases*, 45(1): 2–3.

Lindenmann, J. 2002. "Typhus Vaccine Developments from the First to the Second World War." *History and Philosophy of the Life Sciences*, 24(3-4): 467–485.

Maddison, Angus. 2013. "The Maddison-Project." http://www.ggdc.net/maddison/maddison-project/home.htm.

Parish, H.J. 1965. *A History of Immunization.* London:E. & S. Livingstone Ltd.

Preston, Samuel H. 1975. "The Changing Relation between Mortality and Level of Economic Development." *Population Studies*, 29(2): 231–248.

Preston, Samuel H., Nathan Keyfitz, and Robert Schoen. 1972. *Causes of Death: Life Tables for National Populations.* New York, NY:Seminar Press.

Ramirez, F.O., Y. Soysal, and S. Shanahan. 1997. "The Changing Logic of Political Citizenship: Cross-National Acquisition of Women's Suffrage Rights, 1890 to 1990." *American Sociological Review*, 62(5): 735–745.

Ruuskanen, O., E. Lahti, L.C. Jennings, and D.R. Murdoch. 2011. "Viral Pneumonia." *Lancet*, 377(9773): 1264–1275.

United Nations. Various Editions. "Demographic Yearbook." New York, NY.

United Nations Population Division. 2017. "World Population Prospects." esa.un.org/unpd/wpp/.

World Health Organization. 1951. *Annual Epidemiological and Vital Statistics, 1939-46.* Geneva:World Health Organization.

World Health Organization. 2010. "Cholera Vaccines: WHO Position Paper." *Weekly Epidemiological Record*, 85(13): 117–128.