UNIVERSITY of York

HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 20/01

The Education-Health Nexus: A Meta-Analysis

Xindong Xue

January 2020

The Education-Health Nexus: A Meta-Analysis

By

Xindong Xue[1]

School of Public Administration,
Zhongnan University of Economics & Law, China

December 6, 2019

[1] Correspondence: No.182, Nanhu Avenue, East Lake Hi-tech Development Zone, Wuhan 430073, China. Tel.: +86 27 8838 7901; fax: +86 27 8838 6936.
E-mail address: xuexindong@zuel.edu.cn.

The Education-Health Nexus: A Meta-Analysis

Xindong Xue

Abstract

Does education cause a better health? No consensus answer to this question has yet emerged. In this paper, I perform a meta-analysis of the extensive literature on the health effects of education. The final sample identifies 105 studies with 4,671 estimates. Overall, the health effect of education is not economically meaningful, although statistically significant. There is severe publication bias favoring the positive effect of education on health. Studies that do not control for endogeneity are prone to exaggerate the estimated effect. In addition, the effect becomes weaker for more recent studies. The results suggest that education may not be a feasible policy option for promoting population health.

JEL CODES: B49, C49, I10, I20, I31
KEYWORDS: Education, Health, Human capital, Meta-analysis, Research synthesis

1.Introduction

Since the seminal works of Schultz (1961), Becker (1964) and Grossman (1972), social scientists have generally agreed that education and health play a fundamental role in economic development and well-being. In 2015, United Nations listed education and health as important sustainable development goals by 2030 (United Nations, 2015). Given their intrinsic value for human development, an increasingly large body of literature has been devoted to examining the complementary relationship between education on health. Does education affect health? The empirical answer to this question has important implications. In theory, it can directly test the model of demand for health capital, which predicts that education improves health (Grossman, 1972). In policy, if education does have a large, beneficial effect on health, then educational interventions might serve as a more cost-effective tool for promoting population health than merely increasing public health spending.

A large, positive correlation between education and health has been observed extensively in many countries, no matter how education and health are measured.[1] However, there is a broad disagreement on how large the effect is and whether this effect is causal. While some studies report a significant education effect on health (e.g., Van Kipperslius, O'Donnell, & van Doorslaer, 2011; Oreopoulos, 2006, 2007; Lleras-Muney, 2005), some other studies find small or no effects (e.g., Meghir et al., 2018; Lager & Torssander, 2012; Clark & Royer ,2013; Behrman et al.,2011; Arendt, 2005; Albouy & Lequien, 2009; Braakmann, 2011). Moreover, some studies find mixed evidences across different health dimensions or sub-groups. Education works for some health indicators or only for some specific groups (e.g., James, 2015; Kemptner et al., 2011; Webbink et al., 2010).

Given the diversity of findings on the effects of education on health, the purpose of this paper is to employ meta-analysis to quantitively review the empirical literature in this regard.

[1] For a comprehensive review, see Grossman and Kaestner (1997), Grossman (2006), Cutler and Lleras-Muney (2008,2012).

Meta-analysis is a reliable and objective way to synthesize research findings and has been extensively used in the field of economics, especially in the case that the empirical literature lacks consensus (Neves et al., 2015; Stanley & Doucouliagos, 2012). It also employs multivariate regression to reveal the factors underlying the heterogeneity of estimates, and to establish whether there are any consistent and generalizable results which apply across contexts and methods (Anderson et al., 2018).

To date, there are only two meta-analytic studies on the education effects on health.[1] Furnee et al. (2008) is the first to conduct a meta-analysis on 88 estimates from 40 studies. Their results show that the quality-adjusted life years (QALY) weight of a year of education is 0.036. However, their study only uses self-reported health as health measure. Since health is a multi-dimensional concept, self-reported health alone cannot delineate the overall picture of the health effects of education. Another study by Hamad et al. (2018) provide a review on the quasi-experimental studies of compulsory schooling laws. Their findings indicate that education has an effect on most health outcomes—most beneficial, some negative. However, they don't make clear on what the effect size is. Since different studies use different health measures and estimation methodology, estimates may not be comparable in different studies.

Building on the previous studies, this paper seeks to answer three core questions. First, what is overall effect of education on health? Second, is there any publication bias in the current literature? Third, what are the factors explaining the heterogeneity of the estimates? This paper contributes to the existing meta-analytic literature in several ways. First, I carry out a systematic search and reporting procedure to extract standardized effect size estimates from different studies. The standardized effect size makes it possible to compare the different estimates in different studies. Second, the final sample consists of 4,671 estimates from 105 studies, representing the most comprehensive meta-analysis on the effects of education on

[1] Another closely related literature is Galama et al. (2018). They reviewed the experimental and Quasi-experimental evidence on the effects of education on health and mortality. However, they didn't conduct a formal meta-analysis.

health up to date. Third, I perform meta-regression analyses to reveal the factors behind the different estimated effects across studies. Fourth, as highlighted in some studies (e.g., Galama et al, 2018), heterogeneity may underlie the conflicting results on the effects of education on health. This study codes a detailed list of measures of health and education so as to explore the heterogeneous effects of education on health. Finally, I implement a variety of weighting procedures to calculate the mean effect, use clustered standard errors to eliminate serial correlations between estimates within studies, and conduct formal Funnel Asymmetry Tests (FATs) and Precision Effect Estimate with Standard Error (PEESE) to detect publication bias.

I find that the overall effect of education on health is not economically meaningful, although statistically significant. There is substantial publication bias favoring a positive impact of education on health. The meta-regression analysis indicates that studies that do not address endogeneity are prone to exaggerate the effects of education on health. And the effect becomes weaker for more recent studies. The results cast doubt on the policy initiatives to improve population health through educational interventions.

The rest of the paper proceeds as follows. Section 2 introduces the conceptual framework. Section 3 describes the meta-dataset. Section 4 conducts the preliminary analysis. Section 5 discusses publication bias. Section 6 explores the heterogeneity. Section 7 conducts robustness checks. Section 8 summarizes and concludes.

2. Conceptual Framework

In theory, there are several explanations why education may improve health. The first is the productive efficiency hypothesis, which states that people with higher education are more efficient producers of health because education increases the productive efficiency from the given quantities of health inputs (Grossman, 1972, 2006). The second is the allocative efficiency hypothesis, which argues that education can improve health through the optimal mix of health inputs. Better-educated people have more information on the deleterious effects of smoking and bad habits, so that they are more likely to have healthy lifestyles (Rosenzweig & Schulz, 1989). The third is that education improves health through channels such as better labor market opportunities, higher income, better living conditions, higher quality of care and living environment (Card, 1999; Cutler & Lleras-Muney, 2010).

In empirical literature, identifying the causal effects of education on health is plagued by endogeneity problems. The first originates from reverse causality. Healthier people usually have higher education (Behrman & Rosenzweig, 2004). Second, there may be omitted third variables including genetics, time preference and family background, which simultaneously cause both education and health to move in the same direction (Fuchs, 1982; Bijwaard et al., 2015). For example, people with low time preference usually value future returns more highly, thus investing more in education and health concurrently.

Several identification strategies have been implemented to disentangle the causal relationships between education and health. However, the results of these studies are not consistent. The most widely used strategy is instrumental variable (IV). For example, Lleras-Muney (2005), in her influential study, uses compulsory schooling laws as instruments for education and finds one additional year of compulsory schooling reduces 10-year mortality in the United States by as much as 6 percentage points. In contrast, Black et al. (2005), also using the compulsory schooling laws as instruments, find no effect of education on mortality among the US population. Furthermore, Buckles et al. (2016) exploits the exogeneous variation in years of college education caused by Vietnam draft and shows that education reduces mortality significantly. Finally, a recent study by Fletcher (2015) shows that the effects of education are not precisely estimated, though they appear to be large. It should also be noted that the variations in the results of instrumental variables studies may be partly due to the validity of instrument, as the exclusion condition of IV cannot be directly tested (Grossman, 2015).

The second approach to address endogeneity is Regression Discontinuity Design (RDD).

Compared to IV, RDD imposes weaker assumptions for identification. Oreopoulos (2006) employ compulsory schooling laws reform in UK as a discontinuity and finds a significant effect of education on both physical health and self-rated health. Clark and Royer (2013), however, use the same strategy and report no significant effect of education on mortality in Britain. Albouy and Lequien (2009) also find that education does not have significant effect on the survival rate of French population. Meghir et al. (2018) examine the consequence of compulsory education reform in Sweden and find no significant effect of the reform on mortality.

The third alternative approach is twin fixed effects estimation. The logic behind this approach is that twins share almost the same characteristics such as genetic inheritance, gender, family characteristics and innate ability. It is plausible that differences in education between twins are exogenously determined, so differences in outcomes across twins can be seen as the outcomes from differences in their education. Using the Danish twin samples, Behrman et al. (2011) find that education has no significant effect on mortality, while Van den Berg et al. (2012) and Madsen et al. (2010) report mixed findings which depend on the sample studied. On the basis of twin samples in Sweden, Lundborg et al. (2016) finds that education significantly reduces mortality. Furthermore, some studies show that education reduces the overweight for men (Webbink et al. 2010) but not for women (Webbink et al. 2010, Amin et al. 2013).

It can be seen from the prior literature that estimates on the effects of education on health vary greatly depending the methodology, health measures and sample used. In light of these conflicting findings, it becomes an empirical question to ascertain whether and to what extent education has an impact on health. Since there are large discrepancies in estimates across different studies, it is essential to synthesize these results and arrive at a general conclusion. The meta-analysis can address this question.

3. The Meta-Data Set

3.1 Search Strategy and Selection Criteria

Following the MAER-Net guidelines proposed by Stanley et al. (2013), I carry out a systematic search for the potential studies in the following online databases: *EconLit, Web of Knowledge, Google scholar, JSTOR, EBSCO, RePEc, IDEAS, SSRN, Scopus, ProQuest, NBER, IZA, OECD Library and World Bank Publications*. In line with Minasyan et al. (2019), reference snowballing techniques was also employed to collect articles identified through the search engine process. The search scope covers published journal articles, book chapters, conference proceedings, working papers, master theses, doctoral dissertations, research reports, and other technical reports. The combinations of the following key words are used: *"education", "schooling", "health", "mortality", "disease", "obesity", "BMI", "morbidity", "depression", "cognition", "life expectancy"* and *"survival"*. The search process was completed at the end of December, 2018.

The initial search identified 474 studies in total. To make the analysis consistent, I adopt a stepwise procedure to finalize the sample. The first step is to remove the duplicated studies, including working papers reporting the same estimates with the final published version. Second, I drop the studies which examined the role of health education, for example, the physical education or oral health education, which was not directly related to our subject. Third, I delete the studies that examine the effect of health on education. Forth, I restrict the sample to the effect education on own health, excluding the studies investigating the intergenerational effect of education on health. Fifth, I discard studies lacking sufficient information to calculate t-statistics (that is, standard errors, t-statistics, p-value, or 95% confidence Intervals). Sixth, I exclude the theoretical studies or systematic reviews which did not report econometric estimates. Lastly, I eliminate studies which includes interaction terms and quadratic specifications of the education variable in the regression specification because it is difficult to extract the partial estimated effect (Gunby, Jin & Reed, 2017). I collected 4,718 estimates from 107 studies. The detailed flowchart is shown in Appendix A. The studies are listed in Appendix B.

Table 1 reports the source of selected studies. The majority of the estimates (73.52%) is extracted from journal articles, followed by estimates from working papers (24.62%), conference proceedings (1.54%) and books (0.32%). With regard to the common journal outlets, the journal with highest percentage of estimates is *Social Science & Medicine (18.49%)*, an interdisciplinary journal devoted to the exploration of social science on health. The next journals in terms of frequency are *Journal of Health Economics* (10.83%), *Economics & Human Biology* (9 %), *Social Science Research* (7.69%), and *Economics of Education Review* (6.29%). Following these are a list of economic journals (*OECD Journal: Economic Studies, Health Economics, China Economic Review, Applied Economics*), population journal (*Demography*) and public health journal (*International Journal of Epidemiology*). This proves that education-health nexus is a broad, appealing subject in the field of multi-disciplines.

3.2 Exclusion of Outliers

Outliers may be a concern in the meta-analysis. Before proceeding to code the studies, I remove a few implausibly influential estimates from the dataset. Following Gallet & Doucouliagos (2017), I first estimate a FAT-PET and then remove observations with absolute value of the standardized residual greater than 3.5. As a consequence, the 47 outliers are removed from the dataset. The final sample consists of 105 studies with 4671 estimates. Note that the main results remain quantitatively similar without excluding the outliers.

Figure 1 displays the histogram of t-statistics used in the final sample. The distribution is right-skewed with a few large outliers, indicating many studies reporting positive estimates. Of these, 2183 (46.74%) recorded a significant, positive relationship between education and health at the 10% significance level or below. 2319 (49.65%) record no significant relationship. There are 169 estimates (3.62%) reporting a significant, negative relationship.

3.3 Coding Procedure

For each study in the sample, I code the data to extract a range of study characteristics. This includes: the study's author(s), publication year, publication status (e.g., journal, working paper, books), journal name, data type (cross-sectional or panel; individual level or aggregate level), countries studied, names of health variables, names of education variables, number of observations, and sample type (whole population or subsamples). I also record the regression coefficient and its associated standard errors or confidence intervals so as to derive the partial correlation coefficient (PCC).

It is worth noting that there are a number of studies not reporting the t-statistics or the standard error but p-values. I use the TINV function in Excel to calculate the t-statistics using the p-values and the degrees of freedom (Stanley & Doucouliagos, 2012). Similarly, there are some studies only reporting the levels of statistical significance with asterisk mark ***, **and *. I follow the rule of thumb to assume that p-value is 0.01 for ***, 0.05 for **, 0.1 for * and 0.5 for no asterisk mark. I then use these p-values and degrees of freedom to calculate the t-statistics.

In addition, some studies employ non-linear estimations (e.g.,Probit/logit, cox proportional hazard model) and report only the Odds Ratio (OR) with standard error or 95% confidence intervals. In these cases, I calculate t-statistics by $t_i = \frac{\ln(\widehat{\beta}_{1i})\cdot\widehat{\beta}_{1i}}{S.E._{\cdot i}}$ or $t_i = \frac{\ln(\widehat{\beta}_{1i})\cdot\widehat{\beta}_{1i}}{S.E._{\cdot i}}, where\ S.E._{\cdot i} = \frac{\ln(upperbound_i) - \ln(lowerbound_i)}{2*1.96}$.

The following ten types of estimation methods are coded: OLS, Feasible generalized least squares (FGLS), probit/logit, ordered logit or probit, Hierarchical Linear Model (HLM), instrumental variables (IV), Fixed Effects (FE), Regression Discontinuity Design (RDD), Experiment/Quasi experiment, other estimation techniques. A set of dummy variables are also generated to indicate whether some variables were controlled in the regression analysis, which include age, gender, race, marital status, income and occupation.

As previous mentioned, endogeneity is one of the uttermost concerns in the education and health literature. For this reason, I pay particular attention to estimates produced by estimation methodology addressing endogeneity (IV, FE, RDD, Experiment/Quasi experiment). I will examine whether endogeneity matters a lot for the effect of education on health.

I also code the methods used to derive *t*-values into a set of dummies (*tNormal, tCalculatedBypValue and tCalculatedByCI)*. The type of standard error associated with estimates are coded as *SEspherical* and *SEnonspherical.*

TABLE 3 presents the detailed information on the health measures in the sample. I divide the health measures into three broad categories: physical health, mental health and general health. Of physical health, the most frequent measure is mortality (25.37%). The second frequent measure of physical health is obesity (21.43%), which includes BMI, overweight, body size etc. The third frequent measure of physical health is the presence of a particular illness or disease such as hypertension, heart disease, diabetes (19.72%). The last measure of physical health is Activities of Daily Living (8.07%).

Of measures of mental health, depression is the common form of mental health problems (6.83%). Another type of mental health is cognition (0.77%). Many studies used instruments such as CES-D and malaise score to measure mental well-being. The majority of mental health measures is self-reported.

For measures of general health, it is common to elicit individual's general health through five-point scale subjective assessment: very poor, poor, fair, good and excellent. The general health is generally assessed through self-reporting.

TABLE 4 displays the education variables in the sample. 45.94% of education variable is coded as continuous years of schooling, and 54.06% of education variable is coded as categorical educational levels. Of categorical education levels, about a half (49.94%) is secondary education. Primary and tertiary educations account for 19.52% and 30.53% respectively.

4. Preliminary Analysis

4.1 Effect Sizes

The first step in meta-analysis is to extract the estimates from the literature. To do that, I focus on studies that estimate the effects of education on health with the following regression specification:

$$\mathrm{H} = \beta_0 + \beta_1 Educ + \sum_{k=2}^{i} \beta_k X_k + \text{error} \qquad (1)$$

Where H is a measure of health, $Educ$ is a measure of education, X_k is a vector of control variables, and β_1 is the parameter to be estimated. In meta-analysis, $\hat{\beta}_{1i}$ is the dependent variable, representing the estimated effect of education on health in study i.

However, $\hat{\beta}_{1i}$ cannot be directly used in meta-analysis because the estimates are not comparable across different studies. This is due to the differences in the measure of education and health, or estimation methods. For example, some studies used continuous years of schooling as a proxy for education, while others used categorical education levels. The health measures also differ across studies. Furthermore, because different estimation procedures are employed in different studies, the coefficients were interpreted in various ways. To circumvent this problem, I follow the common approach and use partial correlation coefficient (PCC) to convert estimates into a unitless, comparable measure. PCC is specified as follows:

$$PCC_i = \frac{t_i}{\sqrt{{t_i}^2 + df_i}}$$

Where t_i is t-statistics, df_i is degree of freedom. The value of PCC lies between -1 and 1.

The corresponding standard error of PCC_i is calculated by:

$$s.e.(PCC_i) = \sqrt{\frac{1 - PCC_i^{\ 2}}{df_i}}$$

Accordingly, PCC_i and $s.e.(PCC_i)$ will be used in the following meta-analysis.

Figure 2 depicts the distribution of PPC values in the sample. Compared to the distribution of t-statistics, the distribution of PCC displays a normal shape. The mean and median value in our sample are 0.027 and 0.015. According to the guidelines proposed by Doucouliagos (2011), PCC less than 0.07 can be considered to be small, with 0.17 or above considered to be moderate, and 0.33 or above large. Judged by this criterion, the effects of education on health is very small. Nevertheless, this is not surprising because almost half of the estimates (49.65%) in the sample is statistically insignificant.

4.2 Calculation of Mean effects

To compute the mean effect of education on health, I employ two commonly used estimators in meta-analysis: the fixed effects (FE) estimator and random effects (RE) estimator. [1]

The FE implicitly assumes that there is a single underlying true effect and the reason for different estimates across studies is because of sampling error. The FE estimator can be estimated by Weighted Least Square (WLS), with weights being the inverse of $s.e.(PCC_i)$. It can be formulated as follows:

$$\frac{PCC_i}{s.e.(PCC_i)} = \frac{\alpha_0}{s.e.(PCC_i)} + \frac{\epsilon_{i,}}{s.e.(PCC_i)}\ , i = 1,2, \dots, N \qquad (2)$$

Where N is the number of estimates in the dataset. α_0 measures the mean effect of education on health.

Some scholars argued that the assumption underlying FE is too restrictive. There may be not a single underlying effect, as each study may has its own true effect size. In meta-analysis,

[1] It should be noted that the terms "Fixed Effects" and "Random Effects" are not the same concepts as in the panel data in econometrics.

this heterogeneity can be accommodated by RE estimator.

In the RE estimator, the variation in the effect sizes consists of two parts: the sampling error and the heterogeneity in the true effect size. Assume the heterogeneity represented by τ^2 independent of sampling error, the variation can be expressed as:

$$\omega_i = \sqrt{s.e.(PCC_i)^2 + \tau^2} \qquad (3)$$

In this case, RE estimator is weighted average of all the estimates of the effect size in studies, with the weight given by ω_i. It can be obtained by:

$$\frac{PCC_i}{\omega_i} = \frac{\alpha_0}{\omega_i} + \frac{\epsilon_{i,}}{\omega_i}, i = 1,2,\dots,N \qquad (4)$$

Which method should be preferred in meta-analysis? There is no consensus among researchers yet (Reed, 2015; Doucouliagos & Paldam, 2013; Stanley & Doucouliagos, 2012). Although it is generally held that the assumption of RE model is more realistic, some scholars argue that the FE estimator produce less biased estimates if there is publication bias. As a consequence, both estimates will be reported in this study.

Another issue relates to the weight schemes in this study. In the dataset, the number of estimates per study varies greatly, ranging from 1 to 238, with a mean of 44. The preceding analysis implicitly gives more weight to studies with more estimates and less weight to studies with few estimates. In line with previous studies (Xue & Reed, 2019; Gunby, Jin & Reed, 2017), I employ two weighting schemes, i.e., equal weight to studies (weight 1) and equal weight to individual estimates (weight 2) in the following analysis.

Lastly, since there are more than one estimates in most studies in the sample, it is likely that the estimates are correlated within the study. This will lead to serial correlation and inefficient estimator. To mitigate this concern, I will estimate the model by Weighted Least Square (WLS) with heteroscedasticity cluster-robust standard errors, which accommodates the serial correlation between estimates within each study.

Table 5 reports the preliminary results of the mean effect of education on health. As is shown, the estimated effects of education on health lie between 0.012 to 0.03 and are significant at 5% level. Taking all estimates into account, the overall PCC between education and health is approximately 0.02. According to the guidelines for assessing the strength of a correlation coefficient proposed by Dougcouliagos (2011), the mean effect is far below 0.07, thus being very small, although statistically significant. It should be noted that the simple overall meta-analysis should be interpreted with caution in the case of publication bias. In next section, I will investigate whether there is publication bias and how it might affect the reported estimates in the current literature.

5. Publication Bias

In meta-analysis, publication bias poses a serious threat to the validity of analytic results. Publication bias refers to the fact that peer-reviewed journals are more likely to publish studies with significant results than studies with nonsignificant results, or some studies with insignificant results are seldom written out by authors. Publication bias will result in the incorrect estimates of the mean effect sizes.

An informal way to detect publication bias is the funnel plot, which plots the effect size on the x-axis and standard error on the y-axis (Egger et al., 1997). If there is no publication bias, the distribution of the standard error will be symmetric around the mean line. Publication bias introduces asymmetry into the funnel plot. In the presence of an upward bias, the scatter-dot will cluster on the right of the mean line, or vice versa.

Figure 4 shows that funnel plot of the PCC and standard error in the sample. It can be seen that as the standard error increases, the PCC values are skewed to the right, suggesting an upward publication bias towards a positive impact of education on health.

In addition to the visual representation of the publication bias, a formal way to test publication bias is Funnel-Asymmetry-Precision Test (FAT-PET) (Stanley, 2005; Stanley, 2008). FAT-PET is a simple meta-regression of the PCC on its standard error:

$$PCC_{ij} = \alpha_0 + \alpha_1 se(PCC_{ij}) + v_{ij} \quad (5)$$

Equation (5) can be estimated by FE estimator and RE estimator, as in Equation (2) and Equation (4).

Based on Monte Carlo simulations, Stanley (2008) further argues that α_0 in Equation (5) may be biased downward in the case that null hypothesis is rejected. To surmount this problem, I follow the suggestion of Stanley and Doucouliagos (2014) and replace the standard error with its square term. In this case, α_0 is called the Precision Effect Estimate with Standard Error (PEESE) which is specified as:

$$PCC_{ij} = \alpha_0 + \alpha_1 se(PCC_{ij})^2 + v_{ij} \quad (6)$$

Equivalently, Equation (6) can also be estimated by FE estimator and RE estimator as in Equation (2) and Equation (4).

Table 6 reports the results of FAT-PET. All the four columns reject the null hypothesis $H0$: $\alpha 1$ = 0 at the 5 % significance level, suggesting the presence of publication bias. The positive FAT coefficients suggest upward publication bias, indicating that the current literature favors the publication of positive impacts of education on health. In three of the four columns, I also reject the null hypothesis: $\alpha 0$ = 0, with the estimates of $\alpha 0$ at least significant at the 10% level. The only exception is the "Fixed Effects (Weight2)" regression. In general, the PET coefficients suggest that education is positively correlated with health. However, bias-adjusted estimates of the mean true effect of education on health range from 0.008 to 0.015, far below the value that Doucouliagos (2011) identifies as being "small". Therefore, the overall effect of education on health is not economically meaningful.

Table 7 further presents the PEESE results. Compared to Table 6, all the four columns reject the null hypothesis $\alpha 1$ = 0, $\alpha 0$ = 0. The effects of bias become larger and significant at 1% level, further confirming the persistent publication bias in the literature. The precision effect become more significant and slightly larger, ranging from 0.011 to 0.024. Again, these values are still far below the threshold that Doucouliagos (2011) identifies as being "small". In sum, Table 6 and Table 7 consistently support the point that the effect of education on health is very small and the current literature suffers from substantial publication bias.

6. Modelling Heterogeneity

In this section, I employ meta-regression analysis to explore why the estimates vary systematically across different studies. The differences in the reported estimates may stem from model specification, research design or from differences across countries and over time. To do that, I estimate the following regression s:

$$PCC_{ij} = \alpha_0 + \alpha_1 se\left(PCC_{ij}\right) + \sum_{k=1}^{K} \alpha_{k+1} X_{ki} + \varepsilon_i \qquad (7)$$

Where X_{ki} is the vector of moderator variables, α_{k+1} is the coefficients, $se\left(PCC_{ij}\right)$ is the standard error of PCC_{ij} , ε_i is the error term. Vector X_{ki} contains the following variables:

Measures of Health: Seven measures of health are included in the regression analysis. They are: general health, ADL, disease, mortality, obesity, mental health and whether the health is self-reported. Dummy variables are created to represent each measure of health. Mortality, Obesity and Disease are three most popular measures of health in the dataset, with 25.37%, 21.43% and 19.72% respectively. General health is also a popular measure, accounting for 17.81% of the estimates. However, mental health is less commonly used, with only 7.6% in the dataset. The majority of health measures is self-reported (63%).

Measures of Education: Both the continuous and categorical education measures are used in the literature. The continuous education measure is years of schooling. Categorical education measures are coded into three levels: primary, secondary and tertiary. 46% of the estimates use continuous years of schooling. The remaining estimates (54%) are categorical education levels. The mean number of education variables in the dataset is 1.866.

Regions: The main countries/regions in the dataset cover North America, Europe, East Asia Pacific and some other countries. 35.4% of the estimates are based on North American and 51.5 % from Europe. Asian countries account for 12.1% of the estimates.

Income level: A country's income level is coded on the basis of UN classifications[1]. Most estimates (91.5%) use data from high-income counties. Only a small fraction (8.5%) use data from middle-low income countries.

Data characteristics: most studies use individual-level data (97.8%) and panel data (72.2%). I will explore whether data characteristics influence the reported results.

Sample characteristics: There are five categories of samples: whole population, male sample, female sample, sample aged 25 to 50, sample aged 50 or above, and sample mixed with gender and age. About half of the estimates (48.8%) is based on the whole population. I will investigate whether the effects differ across different samples.

Control variables: Most studies control a set of variables. I include six common variables: *Age, Gender, Race, Marital Status, Income and Occupation*. The variables are coded as 1 if they are included in the regression equation as explanatory variables and 0 if otherwise. Most studies control gender and age in regression specification.

Endogeneity: While many studies use OLS or non-linear models, some studies control for the endogeneity by employing Instrumental Variable (IV), Regression Discontinuity (RD), Fixed Effects and Experiment/Quasi-experiment methods. To be concise, I code the studies controlling for endogeneity as 1 and 0 if otherwise. 30.2 % of the estimates are based on the studies addressing endogeneity. I will investigate whether the effect sizes differ once the endogeneity is controlled for.

Calculation of T-statistics and Standard Error. I also code the methods to calculate the t-statistics. There are three set of dummy variables: *tNormal*, *tCalculatedBypValue*, *tCalculatedByCI*. Most of the T-statistics is calculated by normal ways (68%), followed by confidence interval (19%) and P-value (13%). Moreover, 49.2% of the estimates assume non-

[1] The details can be accessed at: https://www.un.org/en/development/desa/policy/wesp/wesp_current/2014wesp_country_classification.pdf

spherical standard errors.

Publications type: The final set of variables relate to different dimensions of publication process. A dummy variable *Unpublished* is created to account for the difference between published and unpublished studies. Four sets of dummy variables (*Economics Journal, Sociology Journal, Population & development journal, Public health & Medicine Journal, Science Journal*) are created to indicate the different journal types I also explore the effect of quality of journal through *JournalRank,* which is based on the *Scopus Citescore Metrics.* The continuous *Pubyear* is set to quantify the time trend of the effect sizes.

Table 8 lists all explanatory variables, and their mean values, minimum and maximum. Equation (7) is estimated by FE and RE with *weight 1* and *weight 2*.

Given the large number of variables included in the dataset, multicollinearity may arise and confound significant relationships. Moreover, there is model uncertainty on which variables should be included in the regression specification. To overcome these problems, I choose to use variable selection procedure developed by Lindsey & Sheather (2010) to select the best set of control variables. [1]The advantage of this procedure is that it avoids the arbitrary selection of variables. On the basis of the information criteria such as Akaike's Information Criterion (AIC), Akaike's corrected Information Criterion (AIC) and Bayesian Information Criterion (BIC), it helps researchers obtain the optimal model by dropping redundant variables iteratively (Amin, 2016). This selection procedure is performed by invoking the command "*vselect*" in Stata. I employ the backward elimination procedure to perform the selection process to determine the best model specification with the smallest BIC value. I also lock the publication bias variable (*SE)* and a variable that represent attempts to address endogeneity (*endogeneity*) in each round of the selection process.

[1] Another popular approach for model selection is Bayesian Model Averaging (BMA), which has been widely used in the meta-analysis literature. However, BMA is very susceptible to weighting schemes (Gunby, Jin & Reed, 2017). This paper uses variable selection procedure instead.

In the initial round, all the 39 moderator variables (excluding reference variables) are included in the regression specification. At each subsequent round, the variable selection procedure eliminates the variable that results in the largest reduction in BIC. It keeps on doing that until the BIC can no longer be minimized. After obtaining the best set of moderator variables, I re-estimate the model by FE and RE estimators with weight 1 and weight 2.

The main findings are presented in Table 9. The *SE* term is positive and statistically significant at least at the 5 percent level in four estimations. The positive, significant coefficients of *SE* after controlling for a set of moderator variables suggests that the FAT results from TABLE 7 are not a spurious outcome caused by omitted variables. The education-health literature suffers from substantial publication bias.

Not all the health variables are consistently significant across all four estimations. The coefficients on disease are negative and always significant, indicating the effects of education on disease are weak compared to the effects of education on general health. This finding underlines the heterogeneity in the effects of education on health. Moreover, the coefficients on ADL, mortality and depression are sensitive to weight option and not consistently significant, implying no significant difference between the effects on general health, ADL, mortality and depression. The positive coefficients of *SelfReported* in three of the four columns suggest that studies with self-reported health measures tend to overstate the impact of education on health.

With regard to education measures, compared to primary education, continuous years of schooling appears to have a larger and significant effect on health. In contrast, secondary education and tertiary education do not seem to make a clear difference compared with primary education.

Whether the studies correct for endogeneity or not systematically affects estimates. Holding other variables constant, controlling for endogeneity will reduce the PCC values by

0.026~0.037, which are all significant at 1% level. The further analysis of the sample that addresses endogeneity shows that none of the mean effects in FAT-PET-PEESE estimations exceeds 0.01.[1]

The last variable that is consistently significant across all four estimations is *pubYear*. The negative coefficients suggest that the estimated effects of education on health decrease with time.

The results for the other moderator variables in Table 9 are more inconclusive. In the following discussion, I focus on those variables that are significant at 10% level in three of the four regressions. The coefficients for *sample_male and sample_other* are negative and significant, suggesting that men and other samples reap less health returns from education compared to the whole population. This finding is not surprising because men are more likely to have unhealthy behaviors, which may reduce the health benefits of education.

The negative coefficient of *tCalculatedbyCI* indicates that the estimated effect of education on health tend to be smaller if the t-statistics is calculated by confidence intervals. One possible concern is that the mean effect may be downward biased by including estimated effects using these t-statistics. Table 9 shows that the coefficient ranges from -0.001 to -0.044. However, after accounting for this effect, the mean effect size is still below the threshold 0.07 value which Doucouliagos (2011) identifies as “small”.

Regarding publication characteristics, compared to non-journal publications, economics journal and sociology journal seem to publish more negative findings. The positive coefficient estimates of *Journalrank* suggest that high-quality journals report more positive effects of education on health. Neither of other moderator variables show consistent significance across the four estimations.

[1] However, Galama et al. (2018) and Grossman (2015) argue that the smaller effect after the endogeneity is addressed may be due to the fact that IV estimation produces local average treatment effect (LATE) rather than average treatment effect (ATE). That is, the affected population is different.

7. Robustness Checks

I carry out robustness checks on previous findings in two ways. First, following Stanley & Doucouliagos (2012), I use an alternative effect size calculation and test for the robustness of the main results by implementing z transformations of the PCCs and the standard errors. The results of the FAT-PET-PEESE based on Z transformation are shown in TABLE 10. It can be seen that, compared to TABLE 6 and TABLE 7, the main results in TABLE 10 remain unchanged. Second, I re-estimate the FAT-PET-PEESE using the full sample. The results in TABLE 11 remain quantitatively similar to the results with truncated sample. Therefore, across all the robustness checks, the main qualitative conclusions from above are still valid.

8. Conclusion

The education-health nexus has been an intriguing topic in the field of economics and other disciplines for many years. A large body of literature has been emerging over the past decades to examine the education effects on health. However, no consensus has been reached on how large the effect is and whether the effect is causal.

This paper provides a meta-analysis on the extensive literature that examine the impact of education on health. The final sample consists of 4,671 estimates from 105 studies. The main findings indicate that the overall effect of education on health is very small, although statistically significant. This finding echoes Clark & Royer (2013) and Meghir et al., (2018) in that education plays no big role in the process of health production. I also find that there is severe publication bias favoring a positive impact of education on health. The meta-regression analysis indicates that studies that do not control for endogeneity are prone to exaggerate the effects of education on health. The effects become smaller for more recent studies. Thus, this study implies that the theory of demand for health capital that assume a positive role of education in health deserves further investigation.

In terms of policy implications, education has been proposed as one of important health policy initiatives in countries including US and UK (Clark & Royer, 2013). The similar initiative also appears regularly in international organizations such as OECD and WHO (OECD, 2010; WHO, 2015). However, the findings of this paper cast doubt on the feasibility of policies designed to improve health through educational interventions.

To sum up, this study represents the first comprehensive meta-analysis to identify the overall effect of education on health. However, there are still many unknowns regarding the complex relationship between education and health. One important area for future work is to better understand the mechanisms underlying the discrepancies among the estimates across different studies. Another fruitful direction is to examine the effects of education quality on health. The last direction, as cautioned by Grossman (2015), is that the validity of instruments in the current literature needs to be assessed. It is anticipated that this paper will stimulate more research in the future.

References

Albouy, V., Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics*, 28, 155–68.

Amin, A. (2016). Exploring the role of economic incentives and spillover effects in biodiversity conservation policies in sub-Saharan Africa. *Ecological Economics*, 127, 185-191.

Amin, V., Behrman, J.R., Spector, T.D. (2013). Does more schooling improve health outcomes and health related behaviors? Evidence from U.K. twins. *Economics of Education Review* ,35,134-148.

Anderson, E., d'Drey, M.A.J., Duvendack, M., Esposito, L. (2018). Does government spending affect income poverty? A Meta-regression Analysis. World Development. 103, 60-71.

Arendt, J. N. (2005) Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review*, 24, 149–60.

Becker, G. S. (1964). Human capital. New York, USA: NBER.

Behrman, J.R., Rosenzweig, M.R. (2004). Returns to birthweight. *Review of Economics and Statistics*, 86 (2), 586–601.

Bijwaard, G.E., Kippersluis, H., Veenman, J. (2015). Education and Health: The role of cognitive ability. *Journal of Health Economics*, 42:29-43.

Black, D., Hsu, Y.C., Taylor, L.J. (2015). The effect of early-life education on later-life mortality. *Journal of Health Economics*,44,1-9.

Braakmann, N. (2011). The causal relationship between education, health and health related behavior: evidence from a natural experiment in England. *Journal of Health Economics*, 30, 753–63.

Buckles, K, Hagemann, A., Malamud, O., Morrill, M., Wozniak, A. (2016). The effect of college education on mortality. *Journal of Health Economics*, 50,99-114.

Card, David E. 1999. The Causal Effect of Education on Earnings. In *Handbook of Labor Economics*,Vol. 3A, edited by Orley Ashenfelter and David Card, 1801-1863. Amsterdam: Elsevier.

Clark, D., & Royer, H. (2013). The effects of education on adult mortality and health: evidence from Britain. *American Economic Review*, 103 (6), 2087–2120.

Cutler, D. M., & Lleras-Muney, A. (2006). Education and Health: Evaluating Theories and Evidence. NBER Working Paper 12352.

Doucouliagos, H. (2011) How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics. Economics Series 2011/5, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Doucouliagos, H., Paldam, M. (2013) The robust result in meta-analysis of aid effectiveness: A response to Mekasha and Tarp. *The Journal of Development Studies*, 49(4): 584-587.

Egger, M., Davey Smith, G., Schneider, M. and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315: 629–634.

Fletcher, J. M. (2015). New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. *Social Science & Medicine*, 127, 101-107.

Fuchs, V.R. (1982). Time preference and health: an exploratory study. In: Fuchs, V. (Ed.), Economic Aspects of Health. The University of Chicago Press, Chicago.

Furnee, C. A, Groot, w., van den Brink, H. (2008). The health effects of education: a meta-analysis. *European Journal of Public Health*. 18(4), 417–421.

Galama, T. J., Lleras-Muney, A., van Kippersluis, H. (2018). The effects of education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. NBER Working Paper No. 24225

Gallet, C.A., Doucouliagos, H. (2017). The impact of healthcare spending on health outcomes: A meta-regression analysis. *Social Science & Medicine*, 179, 9-17.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy* ,80, 223–255.

Grossman, M., Kaestner, R. (1997). Effects of Education on Health, in J. R. Behrman and N. Stacey (eds.) The Social Benefits of Education (Ann Arbor: University of Michigan Press).

Grossman, M. (2006). Education and Nonmarket Outcomes. In Handbook of the Economics of Education, Vol. 1, edited by Eric Hanushek and Finis Welch, 577 – 633. Amsterdam: North-Holland.

Grossman, M. (2015). The relationship between health and schooling: what is new? *Nordic Journal of Health Economics*, 3(1), 7-17.

Groot, W., Van den Brink, H. (2007). The health effects of education. *Economics of Education Review*, 26, 186–200.

Gunby, P., Jin, Y. and Reed, W.R. (2017). Did FDI really cause Chinese economic growth? A meta-analysis. *World Development*, 90, 242-255.

Hamad,R., Elserb,H., Tranc, D.C., Rehkopfc, D.H., Goodman, S.N. (2018).How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws. *Social Science & Medicine*, 212,168-178.

James, J. (2015). Health and education expansion. *Economics of Education Review*, 49, 193-215.

Kemptner, D., Jürges, H., Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: evidence from Germany. *Journal of Health Economics*, 30,340–54.

Lager, A. C. J., & Torssander, J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proceedings of the National Academy of Sciences*, 109 (22), 8461-8466.

Lindsey, C., Sheather, S. (2010). Variable selection in linear regression. *Stata Journal*. 10, 650-669.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies*,72 (1): 189-221.

Meghir, C., Palme, M., & Simeonova, M. (2018). Education and Mortality: Evidence from a Social Experiment. *American Economic Journal: Applied Economics*, 10(2): 234-256.

Minasyan, A., Zenker, J., Klasen, S., Vollmer, S. (2019). Educational gender gaps and economic growth: A systematic review and meta-regression analysis. *World Development*. 122,199-217.

Neves, P.C., Afonso, O., Silva, S.T. (2016). A Meta-analytic Reassessment of the Effects of Inequality on Growth. *World Development*. 78, 386-400.

OECD. (2010). Improve health and social cohesion through education. Paris, OECD.

Oreopoulos, P. (2006) Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, 96, 152–75.

Oreopoulos, P. (2007) Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91, 2213–29.

Reed, W.R. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. Economics: *The Open-Access, Open-Assessment E-Journal*,9 (2015-30): 1-40.

Rosenzweig, M.R., Schultz, T.P. (1989). Schooling, information and nonmarket productivity: contraceptive use and its effectiveness. *International Economic Review*, 30, 457–477.

Schultz, T. W. (1961). Investment in human capital. *American Economic Review*, 51(1), 1-17.

Stanley, T. D. (2005). Beyond publication bias. *Journal of Economic Surveys*, 19(3),309–345.

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1), 103–127.

Stanley, T.D., Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. Abingdon: Routledge.

Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., Rost, K. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys*, 27(2), 390–394.

Stanley, T.D., Doucouliagos, H. (2014) Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, 5, 60–78.

United Nations. (2015). Transforming our world: the 2030 Agenda for Sustainable Development.

Van Kippersluis, H., O'Donnell, O. & Van Doorslaer, E. (2011). Long-run returns to education: Does schooling lead to an extended old age? *Journal of Human Resources*, 46 4 695-721.

Webbink, D., Martin, N. G. and Visscher, P. M. (2010). Does education reduce the probability of being overweight? *Journal of Health Economics*, 29, 29–38.

WHO. (2015). Health 2020: Education and health through the life-course. Geneva, WHO.

Xue, X.D., Reed, W. R. (2019). Social Capital and Health: A Meta-Analysis.Working Papers in Economics 19/01, University of Canterbury, Department of Economics and Finance.

TABLE 1
Source of Selected Studies

Source	***Estimates***	***Percentage***
Journal Articles	3434	73.52
Working Papers	1150	24.62
Books	15	0.32
Conference Papers	72	1.54
Total	4,671	100

NOTE: Author's calculations.

TABLE 2
Common Journal Outlets

Journal	*Estimates*	*Percentage*
Economics Journal		
Journal of Health Economics	372	10.83
Economics & Human Biology	309	9
Economics of Education Review	216	6.29
OECD Journal: Economic Studies	170	4.95
Health Economics	161	4.69
China Economic Review	129	3.76
Economics Perspectives	116	3.38
Journal of Population Economics	103	3
Applied Economics	84	2.45
Sociology Journal		
Social Science & Medicine	635	18.49
Social Science Research	264	7.69
Population Journal		
Demography	154	4.48
Public Health Journal		
American Journal of Epidemiology	93	2.71
International Journal of Epidemiology	83	2.42

NOTE: Authors' calculations.

TABLE 3

Health Measurements

Measure	Estimates	Percentage
Physical Health	**3484**	
Mortality	1185	25.37
Obesity	1001	21.43
Disease	921	19.72
ADL/IADL	377	8.07
Mental Health	**355**	
Depression	319	6.83
Cognition	36	0.77
General Health	**832**	17.81
Total	**4,671**	

NOTE: Authors' calculations.

TABLE 4
Categorization of Education

Measure	Estimates	Percentage
Continuous (years of schooling)	2146	45.94
Categorical (education levels)	2525	54.06
Primary education	493	10.6
Secondary education	1261	27
Tertiary education	771	17

NOTE: Author’s calculations.

TABLE 5
Mean Effects without Correcting for Publication Bias

Statistics	***Fixed Effects (Weight1)***	***Fixed Effects (Weight2)***	***Random Effects (Weight1)***	***Random Effects (Weight2)***
Mean	0.012*** (3.43)	0.014** (2.33)	0.025*** (9.73)	0.03*** (10.98)
Observations	4,671	4,671	4,671	4,671

NOTE: Columns (1)-(4) report the mean value for PCC. T-statistics is reported in parenthesis. All of the results are estimated by Weighted Least Squares (WLS) with cluster robust standard errors.

* p<0.1, ** p<0.05, *** p<0.01.

TABLE 6

The Funnel Asymmetry Test (FAT) and Precision Effect Test (PET)

Statistics	***Fixed Effects (Weight1)***	***Fixed Effects (Weight2)***	***Random Effects (Weight1)***	***Random Effects (Weight2)***
PET **(α_0)**	0.008*	0.011	0.012***	0.015***
	(1.82)	(1.6)	(4.02)	(4.07)
FAT **(α_1)**	1.6***	1.943**	1.131***	1.403***
	(2.75)	(2.31)	(4.28)	(4.5)
Observations	4,671	4,671	4,671	4,671

NOTE: All the results are estimated by Weighted Least Squares (WLS) with cluster robust standard errors as described in equation (5). T-statistics is reported in parenthesis.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

TABLE 7
Precision Effect Estimate with Standard Error (PEESE)

Statistics	*Fixed Effects (Weight1)*	*Fixed Effects (Weight2)*	*Random Effects (Weight1)*	*Random Effects (Weight2)*
Precision (α_0)	0.011*** (3.16)	0.013** (2.25)	0.021*** (8.01)	0.024*** (8.77)
Bias (α_1)	48.29*** (3.51)	52.18*** (3.43)	19.6*** (2.77)	26.08*** (3.8)
Observations	4,671	4,671	4,671	4,671

NOTE: All the results are estimated by Weighted Least Squares (WLS) with cluster robust standard errors as described in equation (6). T-statistics is reported in parenthesis.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

TABLE 8
Description of Variables

Variable	*Description*	*Mean*	*Min*	*Max*
	EFFECT SIZE			
PCC	Partial Correlation Coefficient	0.027	-0.145	0.275
S.E.	The standard error of the PCC	0.014	0.001	0.123
	HEALT MEASURES			
GeneralHealth	=1, if health measure is general health (Reference)	0.178	0	1
ADL	=1, if health measure is activities of daily living or health limitations	0.081	0	1
Disease	=1, if health measure is specific diseases	0.198	0	1
Mortality	=1, if health measure is mortality	0.254	0	1
Obesity	=1, if health measure is obesity	0.214	0	1
MentalHealth	=1, if health measure is mental health	0.076	0	1
SelfReported	=1, if self-reported health	0.63	0	1
	EDUCATION MEASURES			
Primary	=1, if education measure is primary education (Reference)	0.106	0	1
Secondary	=1, if education measure is secondary education	0.27	0	1
Tertiary	=1, if education measure is tertiary education	0.165	0	1
Education attainment	=1, if education measure is years of education	0.46	0	1
NumberofEducvariables	Number of Education variables	1.866	1	8
	COUNTRIES			
NorthAmerica	=1, if countries in North America included (Reference)	0.354	0	1
Europe	=1, if countries in North Europe included	0.515	0	1
AsiaPacific	=1, if countries in Asia and Pacific included	0.121	0	1
OtherCountry	=1, if countries studied is OCED or transnational	0.01	0	1

Variable	Description	Mean	Min	Max
	INCOME LEVEL			
Middle-Low_Income	=1, if country studied is a middle or low-income country (Reference)	0.085	0	1
HighIncome	=1, if country studied is a high-income country	0.915	0	1
	DATA CHARACTERISTICS			
Aggregate	=1, if estimate is from aggregate-level data (Reference)	0.022	0	1
Individual	=1, if estimate is from individual-level data	0.978	0	1
Cross-sectional	=1, if estimate is from cross-sectional data (Reference)	0.278	0	1
Panel	=1, if estimate is from panel data	0.722	0	1
	SAMPLE CHARACTERISTICS			
WholeSample	=1, if estimate is from whole population (Reference)	0.488	0	1
FemaleSample	=1, if estimate is from female population	0.22	0	1
MaleSample	=1, if estimate is from male population	0.241	0	1
Age25to50Sample	=1, if estimate is from populations aged 25 to 50	0.022	0	1
Age50aboveSample	=1, if estimate is from populations aged 50 and above	0.022	0	1
OtherSample	=1, if estimate is from none of the above	0.007	0	1
	CONTROL VARIABLES			
Age	=1, if age is a control variable in the regression	0.681	0	1
Gender	=1, if gender is a control variable in the regression	0.806	0	1
Race	=1, if race is a control variable in the regression	0.286	0	1
MaritalStatus	=1, if marital status is a control variable in the regression	0.228	0	1
Income	=1, if income is a control variable in the regression	0.133	0	1
Occupation	=1, if occupation is a control variable in the regression	0.1	0	1
	ESTIMATION METHOD			
Endogeneity	=1, if endogeneity is addressed, including Instrumental Variable, Regression	0.302	0	1

Variable	Description	Mean	Min	Max
	Discontinuity Design, Fixed Effects, Experiment/Quasi-Experiment.			
	CALCULATION OF T STATISTIC			
tNormal	=1, if t-statistic=coefficient/standard error (Reference)	0.68	0	1
tCalculatedBypValue	=1, if t-statistic is derived from p-values	0.13	0	1
tCalculatedByCI	=1, if t-statistic is derived from confidence intervals	0.19	0	1
SENonspherical	=1, if standard error is non-spherical	0.492	0	1
	PUBLICATION TYPE			
Unpublished	=1, if unpublished (Reference)	0.265	0	1
Economics Journal	=1, if published in economics journal	0.41	0	1
Sociology Journal	=1, if published in sociology journal	0.195	0	1
Development & Population Journal	=1, if published in development and population journal	0.06	0	1
Public health & Medicine Journal	=1, if published in public health and medicine journal	0.057	0	1
Science Journal	=1, if published in science journal	0.013	0	1
JournalRank	Scopus Citescore Metrics, 0 for non-journal publications	2.136	0	8.58
PubYear	Year study was published/appeared	2011.5	1987	2018

TABLE 9

Meta-Regression Analysis

Variables	*Fixed Effects (Weight1) (1)*	*Fixed Effects (Weight2) (2)*	*Random Effects (Weight1) (3)*	*Random Effects (Weight2) (4)*
SE	**1.533*** (4.682)**	**1.044** (2.161)**	**1.093*** (4.046)**	**1.246*** (3.738)**
ADL	-0.003 (-0.924)	0.005 (0.666)	-0.008* (-1.848)	-0.007 (-0.819)
Disease	**-0.010*** (-4.312)**	**-0.009*** (-3.756)**	**-0.020*** (-5.064)**	**-0.025*** (-4.461)**
Mortality	----	----	----	-0.021*** (-3.728)
Obesity	----	----	-0.006 (-1.267)	-0.014** (-2.033)
Depression	----	----	----	-0.012 (-1.421)
SelfReported	**0.012*** (4.484)**	**0.019*** (3.307)**	**0.012*** (3.318)**	----
Secondary	----	1.365 (1.362)	0.010 (1.280)	----
Tertiary	----	----	0.419* (1.869)	----
Education attainment	**0.007** (2.011)**	**0.018*** (2.698)**	**0.023*** (2.839)**	**0.012** (2.155)**
NumberofEducvariables	-0.002*** (-2.952)	-0.002* (-1.892)	----	-0.002 (-0.686)
EastAsiaPacific	-0.006* (-1.727)	-0.015** (-2.604)	----	0.009 (-1.656)
Europe	0.005 (1.338)	----	-----	-0.008 (-1.439)
Other country	0.012** (2.588)	**----**	0.018*** (3.679)	0.018 (1.318)
Individual data	-0.008 (-1.130)	0.014 (1.445)	-0.012 (-1.074)	-0.012 (-0.998)
Female sample	-0.007* (-1.914)	----	-0.005 (-1.167)	----
Male sample	**-0.010*** (-3.568)**	**----**	**-0.014*** (-3.968)**	**-0.010** (-2.299)**
Other sample	**-0.024*** (-5.714)**	**-0.034*** (-7.756)**	**----**	**-0.021*** (-3.220)**
Age	-0.007** (-2.527)	-0.008** (-2.046)	-0.004 (-0.990)	----

gender	0.012*** (2.918)	----	0.005 (1.365)	----
Race	-0.003 (-0.645)	-0.011* (-1.736)	----	----
Marital Status	-0.003 (-0.999)	----	-0.004 (-0.943)	----
Occupation	----	0.010* (1.807)	-0.011*** (-2.821)	-0.013** (-2.035)
Endogeneity	**-0.026*** (-7.421)**	**-0.037*** (-2.888)**	**-0.028*** (-6.267)**	**-0.032*** (-6.827)**
tCalculatedbyCI	**-0.020*** (-4.023)**	**-0.044*** (-4.241)**	**----**	**-0.001** (-2.280)**
SENonspherical	----	----	0.259** (2.137)	0.183 (1.189)
Economics Journal	**-0.014*** (-3.189)**	**-0.036*** (-3.883)**	**----**	**-0.000 (-1.124)**
Sociology Journal	**-0.017*** (-2.729)**	**-0.033*** (-3.771)**	**----**	**-0.001*** (-2.829)**
Development & Population Journal	----	-0.018 (-1.511)	0.003** (2.078)	0.002 (0.996)
Science Journal	-0.028*** (-3.672)	----	-0.004* (-1.807)	----
Journal rank	**0.005*** (5.161)**	**0.007*** (6.727)**	**0.002 (1.486)**	**0.003** (2.491)**
Pub Year	**-0.001** (-2.551)**	**-0.002** (-2.120)**	**-0.001*** (-2.650)**	**-0.001* (-1.802)**
Constant	**1.615** (2.590)**	**3.123** (2.136)**	**2.393*** (2.679)**	**1.622* (1.859)**
Observations	4,671	4,671	4,671	4,671
R-squared	0.430	0.555	0.459	0.506

NOTE: The results are estimated by the Weighted Least Squares (WLS) with cluster robust standard errors on the basis of equation (7). The coefficient is reported in the top and the associated t-statistics is reported in parentheses below. Two variables *(SE, endogeneity)* were locked into each regression specification. Backwards stepwise procedure was employed to choose the best set of control variables that minimize the Bayes Information Criterion.

* p<0.1, ** p<0.05, *** p<0.01.

TABLE 10

Robustness Checks

(Z-transformed PCCs and Standard Errors)

Statistics	***Fixed Effects (Weight1)***	***Fixed Effects (Weight2)***	***Random Effects (Weight1)***	***Random Effects (Weight2)***
		FAT-PET		
PET (α_0)	0.008* (1.81)	0.011 (1.6)	0.012*** (4.0)	0.015*** (3.99)
FAT (α_1)	1.61*** (2.76)	1.953** (2.32)	1.136*** (4.27)	1.418*** (4.46)
Observations	4,671	4,671	4,671	4,671
		PEESE		
Precision (α_0)	0.011*** (3.16)	0.013** (2.25)	0.021*** (7.98)	0.024*** (8.69)
Bias (α_1)	48.48*** (3.51)	53.51*** (3.44)	19.69*** (2.77)	26.48*** (3.76)
Observations	4,671	4,671	4,671	4,671

NOTE: All the results are estimated by Weighted Least Squares (WLS) with cluster robust standard errors as described in equation (6) and equation (7). T-statistics is reported in parenthesis.

* p<0.1, ** p<0.05, *** p<0.01.

TABLE 11

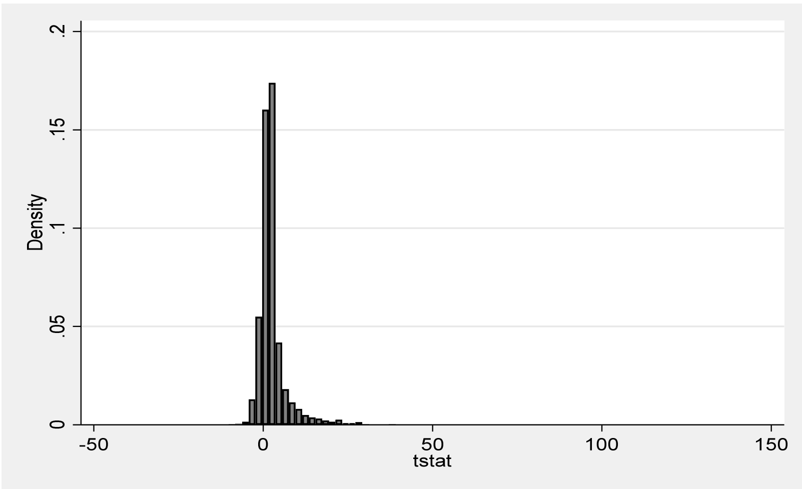
Robustness Checks (Full Sample)

Statistics	***Fixed Effects (Weight1)***	***Fixed Effects (Weight2)***	***Random Effects (Weight1)***	***Random Effects (Weight2)***
		FAT-PET		
PET (α_0)	0.008*	0.01	0.013***	0.005
	(1.76)	(1.6)	(3.18)	(0.69)
FAT (α_1)	1.857***	2.362***	1.326***	2.645***
	(3.13)	(2.82)	(4.02)	(3.61)
Observations	4,718	4,718	4,718	4,718
		PEESE		
Precision (α_0)	0.011***	0.013**	0.024***	0.026***
	(3.23)	(2.27)	(7.62)	(8.54)
Bias (α_1)	52.14***	61.46***	18.2***	38.8***
	(3.81)	(4.21)	(2.53)	(4.04)
Observations	4,718	4,718	4,718	4,718

NOTE: All the results are estimated by Weighted Least Squares (WLS) with cluster robust standard errors as described in equation (6) and equation (7). T-statistics is reported in parenthesis.

* *p*<0.1, ** *p*<0.05, *** *p*<0.01.

FIGURE 1
Histogram Distribution of *t*-Statistics



Distribution of t-statistics	***Obs.***	***Percent***
t < -2.00	169	3.62
-2.00 ≤ t ≤ 2.00	2,319	49.65
t > 2.00	2,183	46.74
Total	4,671	100

FIGURE 2
Distribution of Partial Correlation Coefficient (PCC)

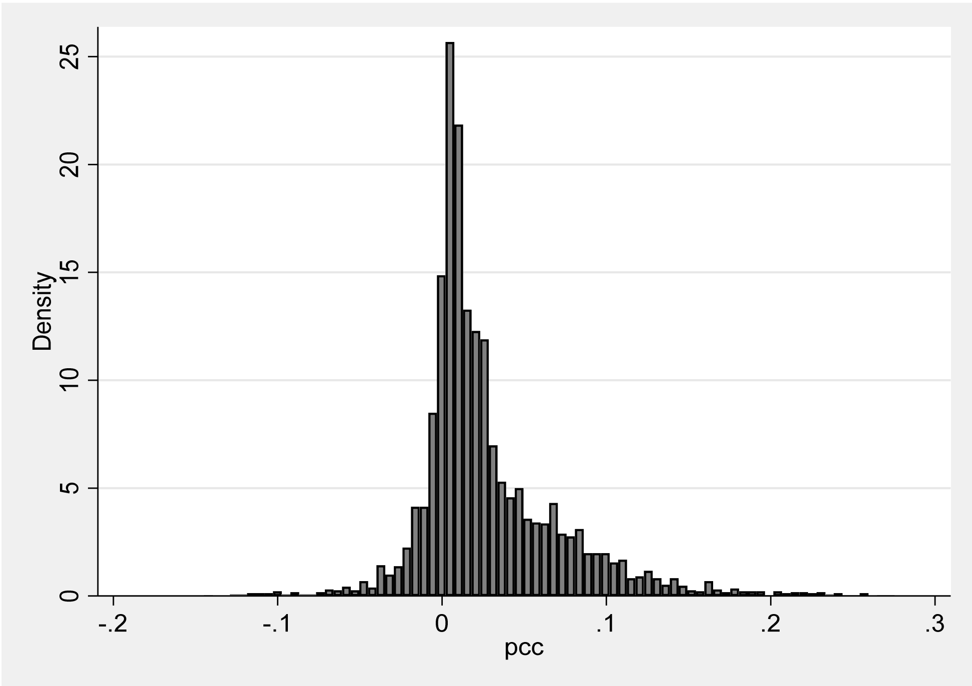
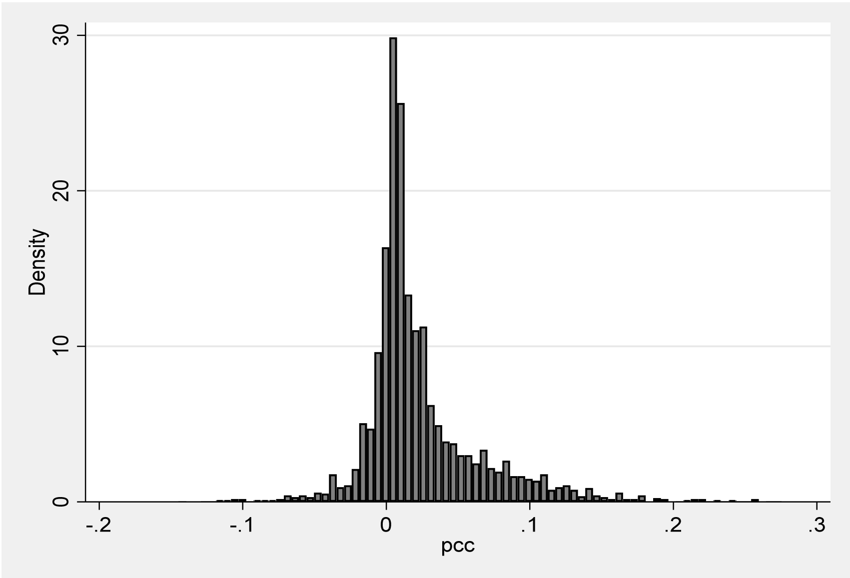
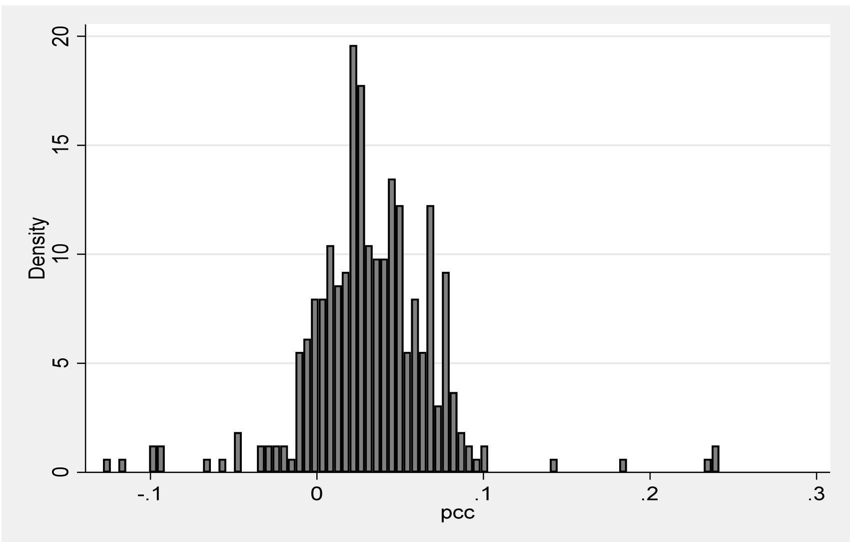


FIGURE 3
Distribution of PCC Values by Health Measures

A. Physical Health



B. Mental Health



C. General Health

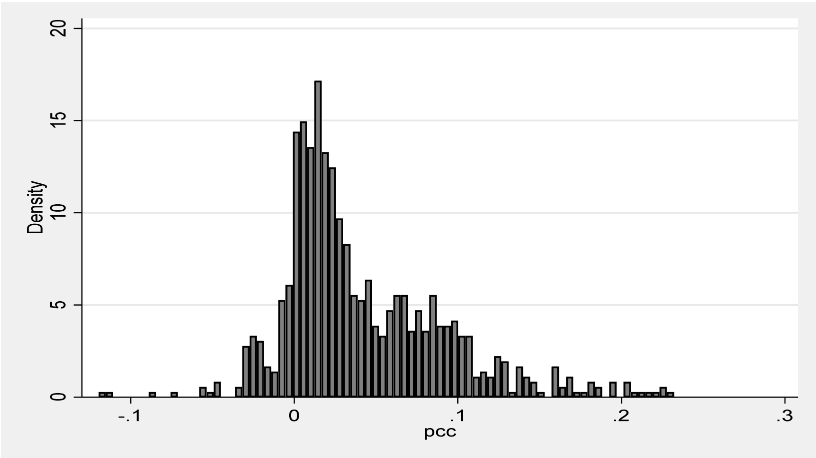
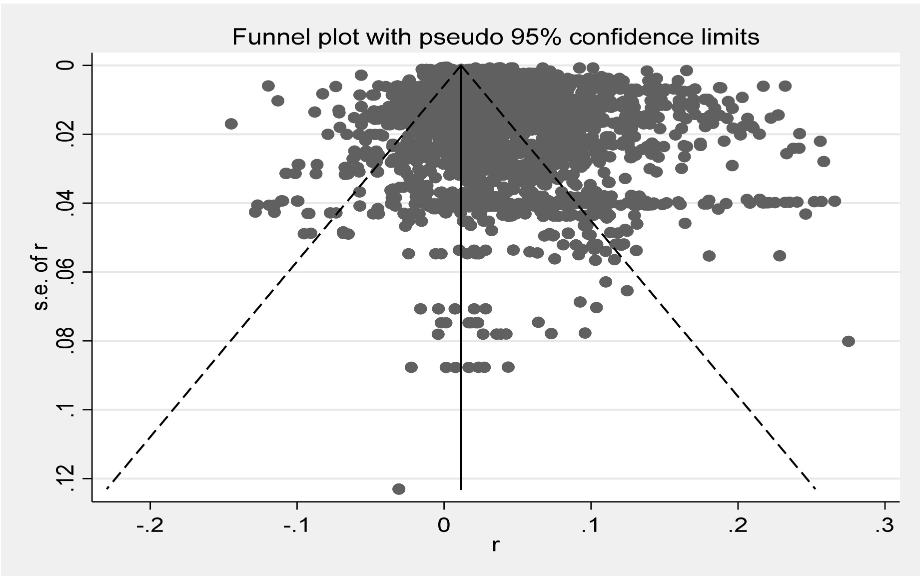


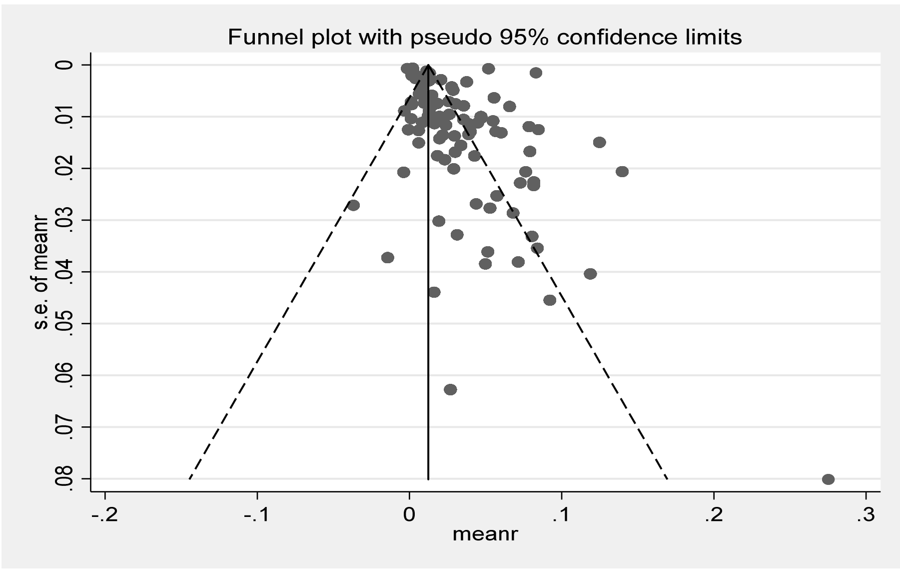
FIGURE 4

Funnel Plots

A. Individual Estimates

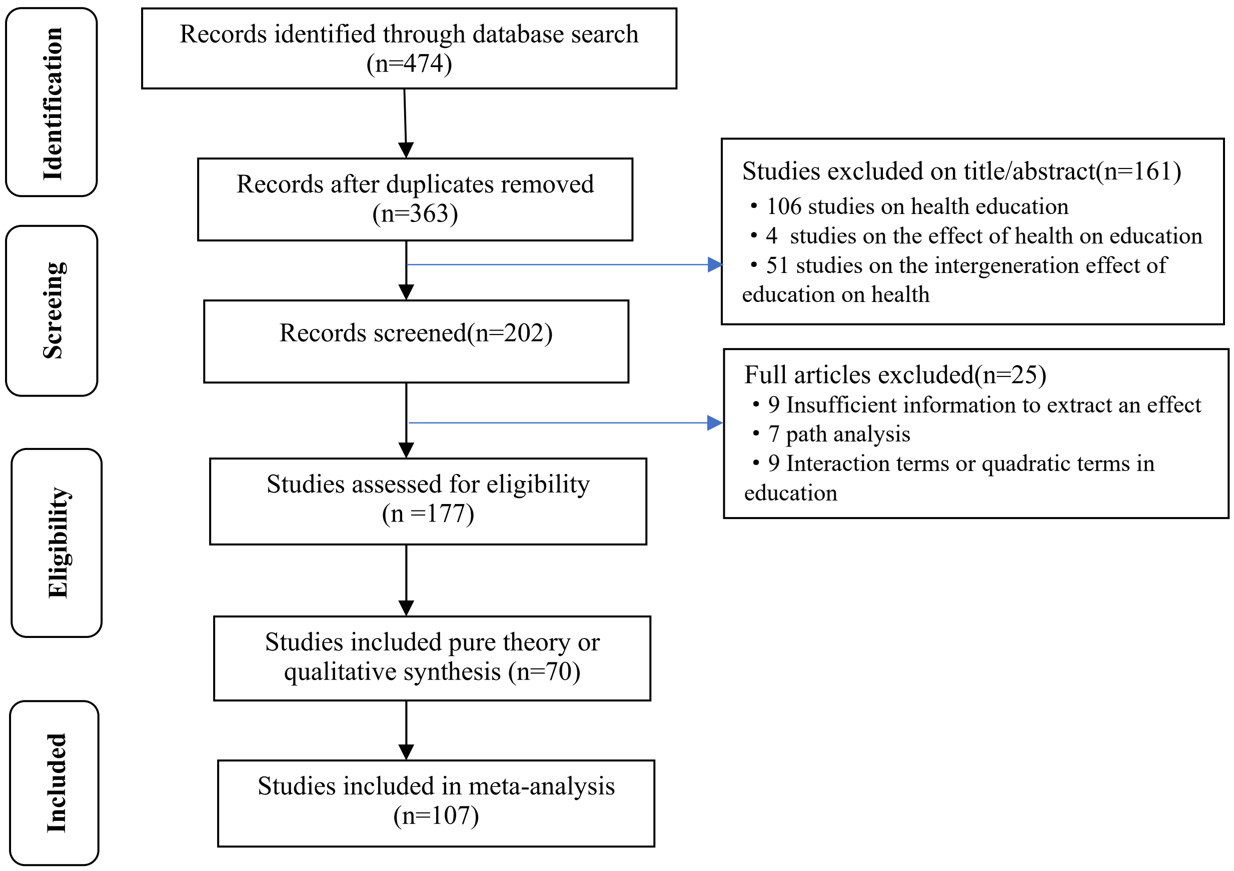


B. Mean Study Estimates



APPENDIX A

PRISMA Flow Diagram for Inclusion in the Meta-Analysis



APPENDIX B

List of Studies

ID	Authors	PubYear	Journal Name	Estimates
1	Meghir et al.	2018	American Economic Journal: Applied Economics	10
2	Clark & Royer	2013	American Economic Review	18
3	Oreopoulos	2006	American Economic Review	8
4	Madsen et al.	2010	American Journal of Epidemiology	93
5	Ross & Wu	1995	American Sociological Review	13
6	Becchetti et al.	2017	Applied Economics	72
7	Zhong	2015	Applied Economics	12
8	Cook	2018	Applied economics letters	4
9	Zhong	2016	China Economic Review	50
10	Xie & Mo	2014	China Economic Review	79
11	Castro	2012	Demographic Research	42
12	Ross & Mirowsky	1999	Demography	24
13	Lundborg et al.	2016	Demography	50
14	Atella & Kopinska	2014	Demography	60
15	Behrman et al.	2011	Demography	20
16	Grossman	2008	Eastern Economic Journal	2
17	Mazumder	2008	Economics Perspectives	116
18	Ma et al.	2018	Economics and Human Biology	18
19	Amin et al.	2018	Economics and Human Biology	12
20	Dursun et al.	2018	Economics and Human Biology	238
21	Kim	2016	Economics and Human Biology	41
22	Arendt	2005	Economics of Education Review	14
23	James	2015	Economics of Education Review	39
24	Groot & Brink	2007	Economics of Education Review	74
25	Jamison et al.	2007	Economics of Education Review	6
26	Silles	2009	Economics of Education Review	32
27	Amin	2013	Economics of Education Review	57
28	Adams	2002	Education Economics	25
29	Edwards	2015	Education Economics	23
30	Hirokawa et al.	2006	European journal of epidemiology	60
31	Brunello et al	2015	Health Economics	16
32	Dee & Sievertsen	2018	Health Economics	10
33	Kiuila & Mieszkowski	2007	Health Economics	60

34	Seo & Senauer	2011	Health Economics	10
35	Van der pol	2011	Health Economics	53
36	Auld & Sidhu	2005	Health Economics	12
37	Fujiwara & Kawachi	2009	International Journal of Epidemiology	53
38	Bann et al.	2016	International Journal of Epidemiology	30
39	Ayyagari et al.	2011	International Journal of Health Economics and Management	16
40	Bai & Li	2018	Journal of Economic analysis & policy	20
41	Braakmann	2011	Journal of Health Economics	45
42	Albouy	2009	Journal of Health Economics	11
43	Kemptner et al.	2011	Journal of Health Economics	80
44	Kemna	1987	Journal of Health Economics	24
45	Cutler et al.	2011	Journal of Health Economics	20
46	Grimard et al.	2010	Journal of Health Economics	24
47	Buckles et al	2016	Journal of Health Economics	25
48	Black et al.	2015	Journal of Health Economics	45
49	Palme & Simeonova	2015	Journal of Health Economics	16
50	Webbink et al.	2010	Journal of Health Economics	82
51	Powdthavee	2010	Journal of Human Capital	22
52	Hardarson et al.	2001	Journal of Internal Medicine	24
53	Brunello et al.	2013	Journal of Labor Economics	14
54	Oreopoulos	2007	Journal of Public Economics	12
55	Loundborg	2013	Journal of population economics	63
56	Jurges et al.	2013	Journal of population economics	40
57	Leuven et al.	2016	Labour Economics	20
58	Devaux et al.	2011	OECD Journal:Economic Studies	170
59	Lager &Torssander	2012	PNAS	59
60	Lutz & Kebede	2018	Population and Development Review	1
61	Böckerman et al.	2017	Preventive Medicine	6
62	Lleras-muney	2005	Review of Economics Studies	26
63	Gerdtham et al.	2016	Scand. J. of Economics	35
64	Lynch & Hippel	2016	Social Science & Medicine	12
65	Bijwaard et al.	2017	Social Science & Medicine	88
66	Bockerman & Maczulskij	2016	Social Science & Medicine	48
67	Benson et al.	2018	Social Science & Medicine	40
68	Zhang	2010	Social Science & Medicine	30
69	Zajzcova et al.	2012	Social Science & Medicine	88
70	Huang & Zhou	2013	Social Science & Medicine	23
71	Knesebeck & Dragano	2006	Social Science & Medicine	140
72	Luo et al.	2015	Social Science & Medicine	84
73	Montez et al.	2018	Social Science & Medicine	20

74	Fletcher	2015	Social Science & Medicine	44
75	Mazzonna	2014	Social Science & Medicine	18
76	Denney et al.	2010	Social Science Research	96
77	Rogers et al.	2013	Social Science Research	168
78	Parinduri et al.	2016	The Journal of Development studies	63
79	Berger & Leigh	1989	The Journal of Human Resources	4
80	Asghar et al.	2009	The Pakistan Development Review	20
81	Arkes	2003	RAND Working Paper	2
82	Barcellos et al.	2018	CESR-SCHAEFFER working paper	8
83	Bijwaard et al.	2016	IZA working paper	36
84	Braakmann	2010	University of Lüneburg working paper	6
85	Braakmann	2010	University of Lüneburg working paper	132
86	Brunello et al	2009	working paper	37
87	Buckles et al	2012	IZA working paper	32
88	Cawley & Choi	2015	NBER working paper	54
89	Cesur et al	2014	NBER working paper	89
90	Chevalier&Feinstein	2006	IZA working paper	200
91	Cipollone & Rosolia	2011	Temi di discussione working paper	32
92	Clark & Royer	2010	NBER working paper	28
93	Cutler & Lleras-Muney	2006	NBER working paper	51
94	Fonseca et al.	2018	working paper	24
95	Fonseca et al.	2011	RAND working paper	24
96	Grabner	2009	working paper	176
97	Huang	2015	IZA working paper	24
98	Janke et al.	2018	IZA working paper	36
99	Lacroix	2018	working paper	19
100	Lillard & Molloy	2007	working paper	5
101	Lundborg	2008	Tinbergen Institute Discussion Paper	30
102	Meghir et al.	2012	IZA working paper	14
103	Albarran et al.	2017	working paper	15
104	Quis & Reif	2017	conference paper	72
105	Schneider & Schneider	2006	working paper	3
106	Spasojevic	2010	books	15
107	Deaton & Paxson	1999	NBER working paper	78