Social Science & Medicine 212 (2018) 168–178

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Review article

How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws

Rita Hamad[a,*], Holly Elser[b,c], Duy C. Tran[c], David H. Rehkopf[c], Steven N. Goodman[c]

[a] University of California San Francisco, Philip R. Lee Institute for Health Policy Studies, Department of Family & Community Medicine, 995 Potrero Avenue, Building 80, Ward 83, San Francisco, CA, 94110, USA
[b] University of California Berkeley, School of Public Health, Division of Epidemiology, Berkeley, CA, USA
[c] Stanford University, School of Medicine, Stanford, CA, USA

ARTICLE INFO

Keywords:
Compulsory schooling laws
Educational attainment
Instrumental variables
Policy evaluation
Regression discontinuity
Systematic review
Meta-Analysis

ABSTRACT

Rich literatures across multiple disciplines document the association between increased educational attainment and improved health. While quasi-experimental studies have exploited variation in educational policies to more rigorously estimate the health effects of education, there remains disagreement about whether education and health are causally linked. The aim of this study was to conduct a systematic review and meta-analysis to characterize this literature, with a focus on quasi-experimental studies of compulsory schooling laws (CSLs). Articles from 1990 to 2015 were obtained through electronic searches and manual searches of reference lists. We searched for English-language studies and included manuscripts if: (1) they involved original data analysis; (2) outcomes were health-related; and (3) the primary predictor utilized variation in CSLs. We identified 89 articles in 25 countries examining over 25 health outcomes, with over 600 individual point estimates. We systematically characterized heterogeneity on key study design features and conducted a meta-analysis of studies with comparable health outcome and exposure variables. Within countries, studies differed in terms of birth cohorts included, the measurement of health outcomes within a given category, and the type of CSL variation examined. Over 90% of manuscripts included multiple analytic techniques, such as econometric and standard regression methods, with as many as 31 "primary" models in a single study. A qualitative synthesis of study findings indicated that educational attainment has an effect on the majority of health outcomes—most beneficial, some negative—while the meta-analysis demonstrated small beneficial effects for mortality, smoking, and obesity. Future work could focus on inconsistent findings identified by this study, or review the health effects of other types of educational policies.

Rich literatures across multiple disciplines document the association between increased educational attainment and improved health (Eide and Showalter, 2011; Braveman et al., 2011). Proposed mediating pathways include greater employment potential, augmented psychosocial resources, and reduced risky health behaviors (Ross and Wu, 1995). Given the recent increased attention to reducing health inequities, it is important to identify whether population-level policies to address socioeconomic factors like education do in fact lead to improvements in health (Low et al., 2005). In this way, societies can determine whether addressing socioeconomic determinants may reduce health inequities, or whether alternative strategies are more appropriate, such as investing in healthcare systems (Adler and Newman, 2002; Adler and Ostrove, 1999).

Yet most studies on the effects of education on health are correlational, making it difficult to establish whether observed relationships are causal, the result of reverse causation, or confounded by unobserved factors such as personality traits or family socioeconomic status (Berger and Leigh, 1989; Montez and Friedman, 2015). Randomization of educational interventions is often logistically difficult or ethically problematic, although a small number of experimental studies have demonstrated potential long-term positive impacts of early childhood education (Campbell et al., 2014; Englund et al., 2015). Nevertheless, experimental studies often cannot achieve sufficient follow-up to document long-term outcomes and typically have limited sample sizes.

Recent decades have seen increased efforts to estimate the causal effects of education on various health outcomes at a population level. Often using quasi-experimental and econometric methods, these studies exploit temporal or geographic variations in policies that lead to differences in educational attainment, and then link these to health outcomes among affected populations (Grimard and Parent, 2007; Lleras-

* Corresponding author.
E-mail address: rita.hamad@ucsf.edu (R. Hamad).

https://doi.org/10.1016/j.socscimed.2018.07.016
Received 11 January 2018; Received in revised form 6 July 2018; Accepted 10 July 2018

Muney, 2005). The most popular policies examined in this way are compulsory schooling laws (CSLs), legislation that has been passed in different countries at different times to establish a minimum number of years of educational attainment among school-aged children. Previous research has confirmed that implementation of CSLs affects educational attainment (Lleras-Muney, 2002; Schmidt, 1996), thereby creating a quasi-randomly assigned exposure whose effect on health can then be examined.

Despite a proliferation of studies on the health effects of CSLs—or perhaps because of it—there remains disagreement about whether educational attainment is in fact causally linked to improved health (Mazumder, 2008; Grossman, 2015). Because the existing evidence spans multiple disciplines, there is a need to systematically review these studies that examine CSLs and health. While one previous study conducted a systematic review and meta-analysis of CSL studies in Europe, showing small effects of education on mortality, self-reported health, and obesity (Ljungdahl and Bremberg, 2015), a significant portion of the CSL literature was not captured by this search. The present study identifies three times as many manuscripts. No systematic review to our knowledge provides such a comprehensive compilation of the literature on CSLs across a broader range of countries and health outcomes.

In this study, we conducted a systematic review of the literature on CSLs, assembling studies that span multiple disciplines and geographic settings. We selected CSL policies in particular because other types of educational policies like school funding or student-teacher ratios address fundamentally different aspects of educational attainment—e.g., quality versus quantity—and because of the large number of quasi-experimental CSL studies that have been conducted and the persistent disagreement about study conclusions. We catalogued all health outcomes that have been examined, from fertility and mortality to biomarkers, and conducted a meta-analysis for a subset of studies with comparably constructed health outcome and exposure variables. In doing so, we hope to provide a comprehensive overview of the state of the CSL literature to date. Our goal is to explain the persistent disagreement regarding the causal effects of educational attainment on health, and to guide future research targeting remaining gaps in the evidence.

1. Methods

1.1. Search strategy

We conducted a search on Google Scholar, a comprehensive resource that includes published and unpublished works. Guidelines for the conduct of systematic reviews and meta-analyses highlight the importance of including both published and unpublished studies, given the possibility for publication bias that would otherwise lead to underrepresentation of null and unpopular results (Rothstein et al., 2006). The search included English-language articles from January 1, 1990 to August 1, 2015. Studies were included if they used the terms “health” AND “compulsory schooling.” A similar search on PubMed found no additional studies. We also examined the reference lists of relevant review articles.

Manuscripts were screened by three investigators (RH, HE, and DCT) preliminarily for relevance based on study titles, abstracts, and main text. If more than one version of a manuscript was identified, only the most recent version was included. Potentially relevant manuscripts were then read in full, and those that met the following inclusion criteria were included in the analysis:

1. The study must involve original data analysis.
2. Study outcomes must be health-related.
3. The primary predictor must be related to compulsory schooling laws. This includes policy variations such as school entry age, exit age, total number of years of compulsory schooling, and quarter-of-birth.

The coding instrument—using the online database REDCap (Harris et al., 2009)—was initially piloted with double entry of a small subset of articles to ensure intercoder reliability. Data were then extracted from the final sample.

Several of the manuscripts in the sample have been published in peer-reviewed journals since the search completion, and these have been updated in our data set (e.g., Brunello et al., 2016). Our review does not, however, include manuscripts produced after August 2015 when our initial search was completed.

1.2. Data elements

For each manuscript, we documented the first author's name, the year of the most recent version, and whether it was published in a peer-reviewed journal.

We next tabulated study characteristics, focusing on features that might explain conflicting findings in the literature. First, we abstracted the health outcomes under examination. Those that only appeared in a single manuscript were categorized as “Other.” We next documented the countries in which health outcomes were examined. Another key feature that we abstracted was the birth cohorts included, since the effects of education might differ based on period effects and historical context (Clay et al., 2012; Delaruelle et al., 2015; Anonymous, 2013). In some studies that examined intergenerational effects of CSLs, we abstracted both the cohort that was affected by the legislation as well as the cohort of their children (e.g., Birgisdóttir, 2013). We also characterized the type of CSL variation that was used in each study's identification strategy. For example, those that exploit variation in school entry age might result in different findings than those that exploit school exit age, since an additional year of schooling in early childhood may have dissimilar effects compared with a year of schooling in adolescence.

We also recorded the analytic methods employed, i.e., standard techniques like ordinary least squares and Cox regression, versus econometric techniques based on quasi-experimental variation in the exposure like instrumental variables (IV) or regression discontinuity (Greenland, 2000; Lee and Lemieuxa, 2010). Different types of analyses might produce different results, which may explain inconsistent findings across studies. For example, IV estimates are a local average treatment effect representing the effect on the compliers rather than an overall average effect. We then explored whether each study restricted analyses to a given sociodemographic subset (e.g., only men, only those without a college education) or whether subgroup analyses were conducted (e.g., by gender or race). Again, heterogeneity might inform differences in study findings across manuscripts. For example, early childhood education has been shown to differentially affect health outcomes in men versus women (Conti et al., 2015). For each study, we documented the largest and smallest sample sizes that were analyzed in each manuscript (e.g., an overall model versus the smallest subgroup analysis). Larger studies might be better powered to produce statistically significant results. We also abstracted the “first stage” coefficient from each study, i.e., the primary coefficient when regressing educational attainment on the policy variable(s) of interest.

Finally, we abstracted the effect sizes and uncertainty measures (i.e., confidence intervals, standard errors, or *P*-values) for the primary models in each manuscript (N = 621 models). We prioritized abstracting confidence intervals or standard errors, although in some cases only *P*-values were reported. In most cases, no single model was highlighted as the primary model by study authors. Consequently, we followed several steps in selecting which estimates to abstract and which model(s) to consider as the primary. First, we prioritized models that included the overall sample, rather than subgroup analyses. Second, we selected the more rigorous models, e.g., those which included adjustment for additional covariates or which employed econometric rather than standard analytic methods. Third, for studies that examined multiple health outcomes, we abstracted the effect size