Abstract Title Page
Not included in page count.

Title:
Predicting student achievement with the education production-function and per-pupil expenditure: Synthesizing regression models from 1968-1994

Authors and Affiliations:
Therese D. Pigott
Ryan T. Williams
Joshua R. Polanin
Meng-Jia Wu-Bohanon

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

Education production functions are commonly used to study the relationship between school inputs (predictors) such as per-pupil expenditure (PPE) and student outputs (outcomes) such as academic achievement. The most-cited study examining education production functions is the *Equality of Educational Opportunity* (Coleman et al., 1966), also known as the Coleman report, that followed over 600,000 K-12 students in more than 3,000 schools. The Coleman Report was controversial because it indicated that school spending was as important as student background variables to determine student achievement, which caught the attention of policy makers. As a result, a number of researchers re-analyzed the Coleman Report's findings and gathered new data to investigate the input-output model.

Not surprisingly, these studies used diverse models to study the effects of PPE, many leading to divergent findings. For instance, Hanushek (1981; 1986) used PPE, teacher experience, teacher education, teacher-pupil ratio, administrative inputs, and facilities to predict student achievement. Sanders (1993), however, used school and district level characteristics to model student achievement. Other researchers have used various teacher demographic characteristics (Murnane & Phillips, 1981), parent education or assets (Ritzen & Winkler, 1977), PPE proxies, such as the number of books in the library (Jencks, 1972), and other school-level characteristics such as minority enrollment (Harnisch, 1987) as the inputs for education production functions.

Hanushek (1996) was the first attempt to synthesize these diverse models. Hanushek collected and analyzed studies by determining the direction (positive or negative) and significance (statistically significant or nonsignificant) of each model's predictor, and then counting the number of PPE coefficients that were positively, negatively or not associated with academic achievement. Using a vote-counting procedure, Hanushek found little evidence that PPE predicts student achievement. In response to Hanushek's work, Greenwald, Hedges, and Laine (1996) synthesized education production-function regression studies using meta-analytic techniques instead of relying solely on vote-counting. The results of this analysis concluded that PPE may have a small, but statistically significant, relationship with student achievement. These two sets of syntheses differ not only in their findings, but also in the methods used to synthesize the results. In addition, debate continues over the substantive question about the relationship between school inputs such as PPE and achievement (Lips, Watkins, & Fleming, 2008) and the methodological one about appropriate synthesis methods for regression models.

One clear problem with both methodological approaches to synthesizing the literature is that both fail to account for diverse model specification across studies. Both the Hanushek (1996), and the Greenwald et al. (1996a) reviews disregard the effect of collinearity between the target slope (PPE) and other predictors. The distribution of the target slope will depend on the degree of collinearity of the predictors. For example, Greenwald et al. (1996) presented a PPE combined significance test table that explored empirically the model's positive and negative consistency (i.e., more positive or negative results than expected by chance alone). The resulting

analysis produced effect size estimates; however, the analyses utilized to produce such estimates failed to examine the model's other predictors.

Purpose / Objective / Research Question / Focus of Study:
Description of the focus of the research.

Thus, the purpose of this research to investigate the heterogeneity of PPE slope estimates in predicting student achievement. The research question guiding this project is: how does the measured relationship between per-pupil expenditure vary across studies that use different models? In concert with SREE's 2012 conference mission of "Understanding Variation in Treatment Effects", this project will inform policy-makers and practitioners on how the distribution of PPE slope estimates varies across different model specifications.

A secondary purpose driving this research is to explore methods for synthesizing regression results when the models differ across studies. To date, few methods exist to synthesize regression models when those models use different sets of predictors (Becker & Wu, 2007). Our data come from an NSF grant to explore methods for synthesizing regression coefficients, where we gathered and coded the education production function studies from Greenwald et al. (1996).

Setting:
Description of the research location.
(May not be applicable for Methods submissions)

N/A

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.
(May not be applicable for Methods submissions)

We have coded all of the 60 studies from the meta-analysis of Greenwald et al. (1996), extracting every model from these studies. The data set consists of 566 models; all outcomes are academic achievement with any number of other predictors. A total of 6,520 variables are included in these models.

Study characteristics have also been coded. We recorded information about the participant's age, race and gender. Also included was information about the school's location, type, and size. Methodological characteristics represented in the coding were the regression form, level of analysis, type of regression, and sample size.

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.
(May not be applicable for Methods submissions)

N/A

Significance / Novelty of study:
Description of what is missing in previous work and the contribution the study makes.

Previous syntheses have failed to describe the variation in school PPE estimates (e.g., Hanushek, 1996; Greenwald et al., 1996). This study contributes a deeper understanding of the education production-function literature by descriptively presenting the distribution of PPE slope estimates across a diverse pool of studies and models. Highlighting the diversity in the target slope parameter estimate, this work will demonstrate how these estimates cannot be treated homogenously.

Statistical, Measurement, or Econometric Model:
Description of the proposed new methods or novel applications of existing methods.

For this presentation, we will examine how the coefficient for per-pupil expenditure varies across a set of models. We are compiling the list of models included in our data set, and will group these models according to the types of predictors included. We will describe how the coefficient of per-pupil expenditure varies depending on the type of predictors in the model. Our goal is to understand how much variation exists in the per-pupil expenditure coefficient across our studies.

This descriptive step will inform our future efforts to develop methods for synthesizing regression coefficients when models differ across studies. We are exploring methods that use robust standard errors for average regression coefficients within studies and that use multivariate techniques to synthesize these regression coefficients across studies. We are also exploring model-based missing data methods for synthesizing these coefficients.

Usefulness / Applicability of Method:
Demonstration of the usefulness of the proposed methods using hypothetical or real data.

Few methods exist for synthesizing regression coefficients. Current methods either require the same model across studies, or require the full covariance matrix for all variables in the model. This presentation will illustrate the problems with synthesizing slope estimates across studies.

Research Design:
Description of the research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).
(May not be applicable for Methods submissions)

The research design represents a systematic review and quantitative synthesis framework. All primary studies will be observational in nature and utilize a regression analysis.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.
(May not be applicable for Methods submissions)

This work uses a meta-analytic dataset of 60 studies, assembled as part of an ongoing study of methods for combining multiple regression estimates. The dataset includes the primary studies used by Greenwald et al. (1996) for their initial meta-analysis of education production-function regression studies. These studies date from 1968-1994.

Analysis will be mainly descriptive. Distributions of conditional regression coefficients will be constructed. These will be conditioned based on model predictors (i.e., the distribution will depend on the other predictors in the model). For instance, the coefficients' distribution may change dependent on the other variables within the model. An analogous procedure is moderator analyses in meta-analysis, where the meta-analyst measures a treatment effect (or correlation) given study characteristics (Raudenbush, 2009).

Findings / Results:

Description of the main findings with specific details.
(May not be applicable for Methods submissions)

Sixty studies are included in the analysis. The dataset consists of published studies from 1968-1994. We plan to organize and describe descriptively the between-study variation of the PPE slopes. In addition, groups of predictors will be created to identify general predictor variables and analyze the PPE slope change given the groupings. For instance, grouping variables by school, teacher, parent, or student could help generalize the relationship of PPE to student achievement. It may be found that PPE shows a strong relationship controlling for teaching experience, but not, on the other hand, when controlling for parental education.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

The education production-functions provide a framework to estimate the relationship of per-pupil expenditure and student achievement. The problem remains that this relationship is hard to evaluate because, in multiple regression analysis, it is conditional on additional predictors. Although these primary studies represent an important investigation, the variance inherent between the model creates synthesis issues that have not been fully addressed to date. It is thus the purpose of this research to investigate and describe the variation of PPE slope estimates in predicting student achievement. This research will bolster the understanding of this relationship, while furthering the knowledge of regression synthesis.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Becker, B.J., & Wu, M.J. (2007). The synthesis of regression slopes in meta-analysis. *Statistical Science, 3*, 414-429.

Coleman, J.S., et al. (1966). Equality of educational opportunity. U.S. Department of Health, Education, and Welfare: Washington, DC.

Greenwald R., Hedges, L.V., & Laine, R.D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), p. 361-396.

Hanushek, E.A. (1981). Throwing money at schools. *Journal of Policy Analysis and Management, 1*, 19-41.

Hanushek, E.A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141-1177.

Hanushek, E.A. (1996). A more complete picture of school resource policies. *Review of educational research, 3,* 397-409.

Harnisch, D.L. (1987). Characteristics associated with effective public high schools. *Journal of Educational Research, 80,* 233-241.

Jencks, C.S. (1972). The quality of data collected by the equality of educational opportunity survey. In Mosteller, F., & Moynihan, D.P. (eds.) *On equality of educational opportunity*. New York, NY: Random House.

Murnane, R.J., & Phillips, B.R. (1981). What do effective teachers of inner-city children have in common? *Social science research, 10*, 83-100.

Ritzen, J.M., Winkler, D.R. (1977). The revealed preferences of local government. *Journal of Urban Economics, 4,* 310-323.

Sander, W. (1993). Expenditures and student achievement in Illinois. *Journal of Public Economics, 52,* 403-416.

Raudenbush, S.W. (2009). Analyzing effect sizes: Random-effects models. In Cooper, H., Hedges, L.V., & Valentine, J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*, 295-314, New York, NY: Sage Publication.