Economics of Education Review 40 (2014) 55–75

Contents lists available at ScienceDirect

Economics of Education Review

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

Understanding the role of time-varying unobserved ability heterogeneity in education production

Weili Ding [a], Steven F. Lehrer [a,b,1,*]

[a] Queen's University, Canada
[b] NBER, USA

ARTICLE INFO

Article history:
Received 6 August 2012
Received in revised form 19 January 2014
Accepted 28 January 2014
Available online 12 February 2014

JEL classification:
I20
J24
C33
C81

Keywords:
Education production function
Unobserved ability
Student learning heterogeneity
Triangular systems

ABSTRACT

Unobserved ability heterogeneity has long been postulated to play a key role in human capital development. Traditional strategies to estimate education production functions do not allow for varying role or development of unobserved ability as a child ages. Such restrictions are highly inconsistent with a growing body of scientific evidence; moreover, in order to obtain unbiased parameter estimates of observed educational inputs, researchers must properly account for unobserved skills that may be correlated with other inputs to the production process. To illustrate our empirical strategy we use experimental data from Tennessee's Student/Teacher Achievement Ratio experiment, known as Project STAR. We find that unobserved ability is endogenously developed over time and its impact on cognitive achievement varies significantly between grades in all subject areas. Moreover, we present evidence that accounting for time-varying unobserved ability across individuals and a more general depreciating pattern of observed inputs are both important when estimating education production functions.

1. Introduction

Since the landmark publication of the 1966 U.S. Department of Education study titled Equality of Educational Opportunity (aka Report Coleman et al. (1966)), hundreds of studies in the economics and education literatures have estimated education production functions to examine whether educational "inputs" correlate with cognitive achievement. Perhaps the major obstacle in production function estimation is that the decisions that a parent makes depend on their child's characteristics. Because many of the child's characteristics that affect these investment decisions are unobserved to the analyst, this gives rise to an endogeneity problem. Intuitively, if a parent adjusts to a change in unobserved innate characteristics by increasing or decreasing their investments depending on whether the change is favorable or not, then these unobserved characteristics and inputs are correlated and biased estimates result. Many researchers interpret these unobserved factors to be either innate ability or unobserved ability heterogeneity.

Many classic studies in the economics literature, including Ben-Porath (1967) and Griliches (1977), emphasize that unobserved ability is an input into the production of human capital, but are ambiguous about how they influence human capital accumulation. As a result, within the economics of education literature researchers often use imperfect proxies for unobserved ability or assume their impacts are constant over time or between siblings including twins. These strategies allow the researcher to either (partially) control for this factor or difference it out in the analysis. However, a

* Corresponding author. Tel.: +1 6135336692.
E-mail address: lehrers@queensu.ca (S.F. Lehrer).

[1] We would like to thank Richard Murnane, two anonymous reviewers and seminar participants at Queen's University, Simon Fraser University, 2008 CSWEP/CEMENT workshop, 2008 CEA Annual meetings and the AEA annual meetings session on Education for the Disadvantaged for helpful comments and suggestions. We are grateful to Alan Krueger for generously providing a subset of the data used in the study. Lehrer wishes to thank SSHRC for research support. We are responsible for all errors.

0272-7757/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.econedurev.2014.01.004

large and growing multi-disciplinary literature summarized within Knudsen, Heckman, Cameron, and Shonkoff (2006) and Cunha, Heckman, Lochner, and Masterov (2006) has demonstrated the malleability of cognitive (and non-cognitive) ability during childhood.[2] These skills are not fixed following conception but rather are related to development of specific brain structures that emerge from both epigenetic and genetic processes. Since unobserved ability heterogeneity is potentially an important contributor to the development of human capital,[3] it would be advantageous to account for its impacts when estimating education production functions in a more flexible manner than existing methods. More generally, to obtain unbiased parameter estimates of educational inputs researchers must properly control for unobserved ability when estimating education production functions.[4]

Since human capital accumulation is a dynamic processes, it is important to understand how the role of heterogeneous ability evolves over the lifecycle, particularly during periods in which it is most adaptable to policy intervention. To estimate the changing importance of heterogeneous ability differences on academic performance, we introduce a straightforward empirical approach that permits estimation of the time-varying effect of unobserved ability heterogeneity within the standard framework of education production functions.[5] Our empirical strategy exploits the triangular structure implied by the underlying model of human capital production and it is important to state explicitly that this empirical approach does not require measures that either proxy for unobserved ability or make assumptions regarding the process by which unobserved ability develops over the lifecycle.[6] Most importantly, the estimates provide guidance on not only the changing impacts of unobserved ability heterogeneity at both different ages and in different subject areas, but also shed light on how researchers should treat this factor in their analyses.

To improve our understanding of the importance of unobserved ability heterogeneity in the production of achievement at different ages we use experimental data from Tennessee's Student/Teacher Achievement Ratio experiment, known as Project STAR. We make use of the feature that teachers were randomly assigned within schools to classrooms in each year of the experiment to overcome important sources of bias in estimating education production functions, including student–teacher sorting bias (Rothstein, 2010). We empirically demonstrate that it is important to account for the time-varying effects of unobserved individual ability heterogeneity, particularly in reading, listening skills and word recognition. Further, specification tests suggest that this factor should be treated as endogenous in the empirical analysis. While our empirical application is within the economics of education, this empirical strategy could be used in other contexts where unobserved unit-specific heterogeneity is believed to play an important role and may have time-varying impacts. For example, this strategy could be used to estimate whether this source of unobserved heterogeneity accounts for much of the gaps that develop among individuals, groups, countries on outcomes such as health and wealth accumulation.

Similar to Andrabi, Das, Khwaja, and Zajonc (2011) dynamic panel methods are used to estimate an education production function. The approach introduced differs from commonly used strategies by exploiting the triangular structure of the underlying economic model of human capital accumulation, allowing us to (i) provide a structural interpretation of what is often termed the persistence parameter,[7] (ii) relax some of the assumptions implicitly made when using a traditional value added estimator, and (iii) easily employ semiparametric estimators to explore the extent of student heterogeneity in their endogenous learning rates.[8] We present evidence of substantial heterogeneity in learning rates across students, particularly in mathematics.

[2] Evidence that gaps in unobserved (cognitive) ability between individuals develop at early ages has been documented within economics (Carneiro & Heckman, 2003)) as well as the child development literature (e.g. Shonkoff & Phillips, 2000).

[3] Within the labor economics literature the empirical importance of unobserved ability heterogeneity to lifetime welfare has been clearly demonstrated. Keane and Wolpin (1997) report that age 16 measures of unobserved ability endowments account for 90% of the total variance in lifetime earnings. Murnane, Willett, and Levy (1995) find that a substantial fraction of the rise in the return to educations between 1978 and 1986 for young workers is attributable to a rise in the return to ability. Heckman and Vytlacil (2001) find this result robust only for a portion of the sample with high scores (in the fourth quartile) on the Armed Services Vocational Aptitude Battery achievement test.

[4] The idea that similar inputs affect the development of ability and test scores and there are feedbacks has been documented empirically. For example, using scores from the Armed Services Vocational Aptitude Battery achievement test to proxy for unobserved ability, Hansen, Heckman and Mullen (2004) present evidence that ability measured at early ages influences the level of education completed (i.e. high school dropouts, high school graduates, etc.) heterogeneously which affects measures of ability taken at later ages. In this study, we do not jointly estimate a unobserved ability production function using methods such as dynamic factor analyis (e.g. Cunha and Heckman (2006) and Cunha, Heckman, and Schennach (2006)) since one must ex-ante assume they know the domains of these skills otherwise they risk introducing measurement error when estimating education production functions. As we detail in the next section, our approach requires the use of instrumental variables and reduces problems of measurement error that arise from using proxies for unobserved ability heterogeneity, which is fundamentally a variable that cannot be measured (yet).

[5] The relationship between empirical specifications of education production functions and the underlying theory is examined in Boardman and Murnane (1979), Todd and Wolpin (2003) and Hanushek (1979). Researchers have also studied the appropriateness of different specifications of an education production function by considering the functional form (Figlio, 1999), levels of aggregation (Hanushek, Rivkin, & Taylor, 1996), relevant control variables (Haveman & Wolfe, 1995) and what constitute the appropriate measures of school output (Card & Krueger, 1992)).

[6] The empirical strategy allows the observed education inputs to both have impacts that vary at different ages and where these inputs could be potentially correlated with the time varying unobserved ability heterogeneity. We discuss the conditions to achieve consistent estimates with both exogenous and endogenous inputs.

[7] Andrabi et al. (2011) conclude their investigation by stating that the economic interpretation of the persistence parameter remains an area open for enquiry. This paper is able to provide a clear economic interpretation by exploiting the triangular structure of the underlying economic model.

[8] As we discuss in further detail in the next section, feasible approaches to estimate conditional quantiles with panel data and endogenous regressors are difficult to develop since standard demeaning (or differencing) techniques do not generally remove the time-invariant unobserved heterogeneity. Our approach involves first solving for the unobserved heterogeneity so that estimators based on the L_1-norm penalty can be utilized.