Economics and Human Biology 21 (2016) 137–146

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: http://www.elsevier.com/locate/ehb

On the genetic bias of the quarter of birth instrument

Cornelius A. Rietveld [a,b,c,*], Dinand Webbink [a,c,d]

[a] Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands
[b] Erasmus University Rotterdam Institute for Behavior and Biology, Erasmus University Rotterdam, Rotterdam, The Netherlands
[c] Tinbergen Institute, The Netherlands
[d] Institute for the Study of Labor (IZA), Bonn, Germany

ARTICLE INFO

Article history:
Received 17 August 2015
Received in revised form 8 January 2016
Accepted 14 January 2016
Available online 27 January 2016

Keywords:
Quarter of birth
Instrumental variable
Genetic variation
SNPs

ABSTRACT

Many studies in economics use quarter of birth as an instrument for identifying the causal effect of schooling on outcomes such as earnings and health. The key assumption in these studies is that people born in different quarters of the year do not differ systematically in their unobserved abilities. This study uses genetic data from the US Health and Retirement Study to analyze the validity of the quarter of birth instrument. We find some evidence that genetic factors influencing education are not randomly distributed over the year. However, these factors only slightly change the effect of quarter of birth on schooling.

1. Introduction

In one of the most influential papers in modern empirical economics Angrist and Krueger (1991) (A&K) use quarter of birth as an instrumental variable for estimating the wage returns to schooling[1]. They exploit the fact that, due to differential exposure to compulsory schooling, males born later in the year, especially those born in the third and fourth quarter, attain on average more schooling than those born earlier in the year. Thevalidity of this approach has been questioned as quarter of birth is only weakly correlated with schooling. Even a weak correlation between quarter of birth and unobserved ability, for instance due to seasonal effects, might yield a large inconsistency in the IV estimates (Bound et al., 1995)[2]. These validity concerns appear to be unsettled. For instance, a recent study shows that 'winter babies' are more likely to have unmarried mothers, teenage mothers or less educated mothers, and that maternal schooling peaks for mothers who give birth in the second quarter (Buckles and Hungerman, 2013). However, these patterns in family background "cannot account for the seasonal pattern in schooling and wages

* Corresponding author at: Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Tel.: +31 10 4081401; fax: +31 10 4089141.
E-mail address: nrietveld@ese.eur.nl (C.A. Rietveld).

[1] Quarter of birth is used in many other studies, see for instance Angrist and Krueger (1992, 1995, 2001), Levin and Plug (1999), Plug (2001), Adams (2002), Gelbach (2002), Lemke and Rischall (2003), Hansen et al., 2004, Skirbekk et al., 2004, Lefgren and McIntyre (2006), Andini (2008), Leigh and Ryan (2008), Angrist and Pischke (2009) and Maurin and Moschion, 2009. For a survey of the broader literature on returns to schooling see Heckman et al. (2006).

[2] This concern also raised interest for the weak instrument problem, see for instance Staiger and Stock (1997), Chamberlain and Imbens (2004), Cruz and Moreira (2005), Dufour and Taamouti (2007), Honoré and Hu (2004), Imbens and Rosenbaum (2005), Chesher (2005), Chernozhukov and Hansen (2006), Small (2007), Berkowitz et al. (2008), Conley et al. (2012). Recent studies have also raised concerns about the so-called 'persistence bias'. Estimates of schooling returns might be biased due to the persistence of individual wages (e.g. Andini, 2013, 2014).

http://dx.doi.org/10.1016/j.ehb.2016.01.002
1570-677X/

both of which exhibit third and fourth quarter peaks" (Angrist and Pischke, 2015, p. 233).

Educational attainment also has a genetic component and heritability studies show that around 40% of the differences in schooling can be explained by genetic differences across individuals in a population (Branigan et al., 2013). This implies that the quarter of birth instrument can also be biased from a genetic perspective, because genetic factors influencing educational attainment may not be evenly spread over children born in different quarters of the year. Recent advances in genetic research, in particular the increasing availability of individual-level genetic data, provide new opportunities for investigating this potential bias and may provide new insight into the validity of the key assumption mentioned above.

In the first part of this paper we investigate whether genetic differences between individuals born in different quarters of the year are important for the relationship between quarter of birth, schooling and wages, as analyzed by A&K. We exploit recent findings from a Genome-Wide Association Study (GWAS) on educational attainment of 126,559 individuals that found statistically significant and replicable associations between three genetic variants and education (Rietveld et al., 2013, 2014). We investigate the potential bias of the instrumental variable (IV) estimates of the effect of schooling on wages due to these three genetic variants and due to a linear combination of all analyzed genetic variants in Rietveld et al. (2013) that are available in our data. The explanatory power of these genetic variants is relatively small, but a weak correlation between quarter of birth and unobserved ability might generate large bias for the IV estimates.

In the second part of this paper, we investigate whether the genetic factor for educational attainment is associated with the genetic factor for being born in a specific quarter of the year. The fact that genetic variants identified in GWAS's explain only a relatively small proportion of the heritability[3] of a trait – the so-called 'missing heritability' puzzle (Manolio et al., 2009) – led to the development of the Genomic-Relatedness-Matrix Restricted Maximum Likelihood (GREML) method (Yang et al., 2010). This method analyzes which proportion of the variation in a specific outcome can be explained by all observed genetic factors within a sample. This is a great advantage over the analyses in the first part of this paper, where the two genetic measures capture only a relatively small part of the genetic factor for educational attainment[4]. We use GREML to estimate this proportion for educational attainment and for quarter of birth. Moreover, we perform a bivariate GREML analysis (Lee et al., 2012) to analyze whether the genetic factors that explain educational attainment and quarter of birth are correlated.

We find that individuals born in the second quarter of the year are less likely to have the three specific genetic variants associated with educational attainment. For our second genetic measure, the linear combination of all genetic variants, we do not find seasonal patterns in the distribution. Controlling for the genetic measures only slightly changes the estimated effect of quarter of birth on educational attainment. For the subsample of males for which we also have data on wages we find that the genetic factors only slightly change the OLS-estimates of the effect of schooling on wages and the reduced form estimates of quarter of birth on wages. The findings provide no evidence that IV-estimates of the effect of schooling on wages are biased by omitting genetic factors. Our current knowledge about the genetic architecture of educational attainment does not warrant concerns about a genetic bias in A&K type models. In the second part of our analysis we find that genetic factors explain 26% of the variance in schooling. We also find that genetic endowments of individuals born in the second quarter of the year are slightly different from the endowments of individuals born in other quarters of the year. However, the overlap of the genetic factors underlying schooling and being born in the second quarter is statistically insignificant. Again, these findings do not point to a serious genetic bias of the quarter of birth instrument.

Our study contributes to the economic literature by using recent advances in genetic research for investigating a key issue in one of the modern 'classics' in empirical economics. This novel perspective contributes to the literature that aims to estimate causal effects of schooling on a wide range of outcomes and links labor economics to genoeconomics (Benjamin et al., 2012). Our paper is the first study in the returns to schooling literature that exploits genetic data. Moreover, our study contributes to the literature across social and natural sciences that investigates the seasonality of outcomes.

2. Empirical strategy

For estimating the causal effect of education on wages A&K use quarter of birth as an instrumental variable for schooling. They argue that the interaction between quarter of birth and compulsory school attendance laws generates exogenous variation in education. Children born in the first quarter of the year will typically start schooling when they are close to age 7 and will reach age 16 in the middle of tenth grade. Children born in the last quarter of the year will start schooling just before turning age 6 and will finish tenth grade before reaching age 16. They exploit this natural experiment by estimating a 2SLS model:

$$S_i = \alpha_0 + \alpha_1 X_i + \sum_j \alpha_j Q_{ij} + \epsilon_i \quad (1)$$

$$\ln W_i = \beta_0 + \beta_1 X_i + \beta_2 S_i + \eta_i \quad (2)$$

where S_i is years of schooling of individual i, X_i is a vector of covariates, Q_{ij} is a dummy variable indicating whether the individual was born in quarter j ($j = 1, 2, 3$) and W_i denotes wages. Estimation of the parameters in this model yields the causal effect of education on wages when quarter of

[3] Heritability is defined as the proportion of variation in the outcome variable that can be explained by genetic differences in a sample (Visscher et al., 2008). Traditionally, heritability is estimated by comparing monozygotic and dizygotic twins (Plomin et al., 2013).

[4] A disadvantage of GREML is that it analyzes the genetic factor for educational attainment and quarter of birth. It does not provide a way to associate the genetic factor for educational attainment with the observed quarter of birth.