



Are teachers accurate in predicting their students' performance on high stakes' exams? The case of Russia



Andrey Zakharov^a, Martin Carnoy^{*,b}

^a National Research University Higher School of Economics, Ul. Myasnitskaya 20, Moscow, Russia

^b Stanford University, School of Education, 485 Lasuen Mall, Stanford, CA 94305, USA

ARTICLE INFO

Article history:

Received 21 September 2014

Received in revised form 23 March 2015

Accepted 27 April 2015

Keywords:

Educational policy
International education
Teacher expectations
High stakes exams

ABSTRACT

The paper focuses on how accurate teachers may or may not be in gauging their class' academic abilities. We use a sample of classrooms in three Russian regions to identify sources of mathematics and Russian teachers' inaccuracies in predicting their high school classes' scores on Russian and mathematics high stakes college entrance tests (the Unified State Exam, or USE). We test the hypothesis that teachers' perceptions of their relationship with their classes are good predictors of such inaccuracies. This is important because teachers often focus on their relationship with the class as an end in itself or as a means to engaging students. Good teacher–student relations may indeed result in more students' learning, but perhaps not nearly as much as teachers' believe. We find that both Russian and mathematics teachers make inaccurate predictions of their class' high stakes examination results based on how they perceive their relationship with their class. Teachers who believe they have a very good relationship with the class significantly overestimate their class' performance on the USE, and those who perceive a poor relationship, underestimate their class' performance, although this underestimate is generally not statistically significant.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Much of the literature on effective teaching emphasizes the ideal of teachers as reflective professionals capable of individualized approaches to student learning (for example, Cohen, 1993; Darling Hammond, 1996). An implicit assumption underpinning this ideal is that teachers are accurate, fair judges of their students' abilities, and that they can (and should) individualize broad curricular guidelines to fit each student's capacity and learning style.

Two strands of research have questioned this assumption. One strand, going back to the 1960s, argues that teachers may not be neutral observers of students' abilities, that teachers' expectations may vary among students, and that teachers' expectations (positive or negative) can affect students' performance (Rosenthal and Jacobson, 1968). More recently, this research has turned to finding students' (and teachers') characteristics that may affect teacher expectations in particular subjects, and, hence, their students' performance (Brophy, 1983; Rosenthal and Rubin, 1978;

Rosenthal, 1994). A number of studies find teacher gender bias—teachers viewing boys as having greater math and science skills and girls as having greater literary skills (Qing, 1999; Ready and Wright, 2011; Riegle–Crumb and Humphries, 2012; Shepardson and Pizzini, 1992), but others find no evidence of teacher gender bias (Dusek and Joseph, 1983; Madon et al., 1998). Similarly, many studies have found teacher race/ethnic bias (Ready and Wright, 2011; Rubie–Davies et al., 2006; Tenenbaum and Ruck, 2007), and social class bias (Auwarter and Aruguete, 2008; Ready and Wright, 2011). This may also relate to how teachers view students in different academic tracks (Kelly and Carbonaro, 2012; Oakes, 1985; Page, 1987; Tach and Farkas, 2006). Teacher gender and ethnicity have also been shown to play a role in affecting the performance of students' of particular gender and ethnicity (Dee, 2005; McKown and Weinstein, 2008; Ready and Wright, 2011; Van den Bergh et al., 2010). Some of these studies estimate causal effects and show that teachers' subjective judgment—consciously or unconsciously—can and does affect students' academic outcomes.

Although much less studied, the second strand of research argues that teachers' expectations for students' performance compared to actual results may differ not because of conscious or unconscious “biases,” but because of what Jussim and Harber (2005) called “predictive validity without self-fulfilling influence.” Ferguson (2003) argued that both teacher inaccuracy and bias

* Corresponding author. Tel.: +1 6508567722.

E-mail addresses: ab.zakharov@gmail.com (A. Zakharov), carnoy@stanford.edu (M. Carnoy).

regarding student performance are deviations from a “true” benchmark, namely how much students have actually learned and their performance on measures of their learning. Teachers may be poor predictors of student performance because teachers may misestimate how much certain teaching practices or classroom conditions positively or negatively affect student learning and test performance. Teachers’ expectations based on these views of practices and classroom conditions can be inaccurate but not necessarily biased if the difference of predicted and actual class average test scores does not vary systematically according to classroom (students’) characteristics.

In this paper, we focus on this second strand of “predictive validity,” namely how accurate teachers may or may not be in gauging their class’ academic abilities. We use a sample of classrooms in three Russian regions to identify sources of mathematics and Russian teachers’ inaccuracies in predicting their high school classes’ scores on Russian and mathematics high stakes college entrance tests (the Unified State Exam, or USE). We test the hypothesis that teachers’ perceptions of their relationship with their classes are good predictors of such inaccuracies. This is important because teachers often focus on their relationship with the class as an end in itself or as a means to engaging students. Good teacher–student relations may indeed result in more students’ learning, but perhaps not nearly as much as teachers’ believe.

Teachers should be able to predict how their class will score on either the mathematics or Russian section of the USE, since in the 11th grade of Russian schools great emphasis is placed on preparing for this examination, including homework assignments and practice tests. Teachers in Russia usually teach the same students for two years in either mathematics or Russian. We control for the characteristics of a class that could “bias” teachers’ expectations of students’ performance. We also control for two variables that should help teachers make better predictions of their class’ performance on the USE: (a) students’ grades in the first semester of the 11th grade (information about students’ previous performance) and (b) teachers’ years of teaching experience (accumulated expertise).

We find that both Russian and mathematics teachers make inaccurate predictions of their class’ high stakes examination results based on how they perceive their relationship with their class. Teachers who believe they have a very good relationship with the class significantly overestimate their class’ performance on the USE, and those who perceive a poor relationship, underestimate their class’ performance, although this underestimate is generally not statistically significant. Teachers’ view of their relationship with the students in their class is not significantly related to the gender composition of their class in either math or Russian, but Russian teachers are significantly more likely to report a very good relation with their class when the cultural capital (Bourdieu and Passeron, 1976; Bourdieu, 1986)¹ of their students is higher, even controlling for students’ academic performance. This may suggest some “cultural capital bias” on the part of Russian teachers.

The paper is structured as follows: in the next section, we describe our study’s Russian secondary education context; Section 3 describes our methodology and data; Section 4, the results, and the final section discusses the results and concludes.

¹ Bourdieu and Passeron (1976) defined cultural capital rather broadly on the one hand as parents’ education and the intellectual climate at home, and, on the other, rather narrowly as the quality of language and verbal interaction at home. We used a proxy for cultural capital—more than 100 books in the home as reported by the student. We also collected data on students’ parents’ education, but a large fraction of students do not answer this question, and for those students that did, parents’ education was highly correlated with books in the home. For a discussion of books in the home as a proxy for family academic resources, see Carnoy and Rothstein (2013).

2. The advantages of studying the Russian case

Russian high school is a particularly interesting context in which to study teachers’ accuracy in assessing how well their students have learned an academic subject. High school teachers generally know their students well in Russia. Almost all students have the same mathematics teacher and the same Russian language teacher for both high school grades (10–11th). In many instances, they even have the same teacher for each subject since the 5th grade. Thus 11th grade teachers have known their students for at least two and often more years. Further, for demographic reasons and student attrition (around 40% of students go to vocational school after the 9th grade) there is usually one math and one Russian class in the 11th grade of high school and it is usually of a smaller size than classes in middle or primary schools. All these factors help teachers to become more familiar with their students’ attitudes and abilities. Teacher expectations for their students should therefore be rather accurate in high school.

As some analysts have noted, studies of teacher perceptions rarely use objective measures of students’ outcomes; therefore bias in teacher expectations may be reported even if it didn’t take place (Ready and Wright, 2011). The data we use for outcomes are the students’ Unified State Examination (USE) scores. The USE is a high school exit/college entrance test used throughout Russia since 2009. It has some distinct advantages as a measure of students’ actual performance. It is a standardized test graded by agencies external to the school and thus provides an “objective” measure of students’ achievement. It is curriculum based: it measures students’ performance on the subjects they had studied. Finally it is high stakes test. It is required for graduation from upper secondary school and also serves as an entry exam to all universities in the country. Russian language and mathematics are of special importance as they are mandatory subjects to sit in USE. Many university departments require USE results in mathematics, and all departments in all universities require the results in Russian. Almost all 11th grade students want to attend university, and the USE results determine their choices. High stakes tests such as the USE are therefore likely to serve as better measures of students’ actual performance because of their high motivation to perform well (Bishop, 1997).

3. Data

3.1. Survey timing and sampling

We use data from a sociological survey conducted in May 2010 in three Russia regions: Pskovskaya and Yaroslavlskaya *oblasts* and Krasnoyarsky *krai*. These regions were selected because they provide significantly different demographic, social and economic contexts for high school education. Yaroslavlskaya *oblast* is a small region in the center of Russia, rather average in regional ratings of social and economic development. Pskovskaya *oblast* is located in the northwest part of the country. It is also small in terms of population and relatively poor economically. Finally Krasnoyarsky *krai* is a large region in the east of Russia (in Siberia) that is one of the nation’s most highly developed regional economies.

The survey was applied toward the end of a school year, two to four weeks before the USE examination, in May. By that time, teachers had a great deal of information about students’ performance, behavior and family background, considerably increasing the likelihood of accurate teacher predictions of student performance on the USE.

The data were collected based on a stratified random sample that represented final year students in each selected region. Using a list of all schools, we grouped them into strata. Among the parameters for stratification were the type of settlement (rural, urban, regional center), school type (regular school, school with advanced study of some subjects, *gymnasiums*, and *lyceums*), and

Download English Version:

<https://daneshyari.com/en/article/356021>

Download Persian Version:

<https://daneshyari.com/article/356021>

[Daneshyari.com](https://daneshyari.com)