Economics of Education Review 40 (2014) 140–151

Contents lists available at ScienceDirect

Economics of Education Review

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

Analysis of group performance with categorical data when agents are heterogeneous: The evaluation of scholastic performance in the OECD through PISA

Carmen Herrero [a], Ildefonso Mendez [b,*], Antonio Villar [c]

[a] University of Alicante & IVIE, Spain
[b] University of Murcia, Spain
[c] Pablo de Olavide University & Ivie, Spain

ARTICLE INFO

Article history:
Received 24 April 2013
Received in revised form 5 February 2014
Accepted 6 February 2014
Available online 5 March 2014

JEL classification:
I24
C14

Keywords:
Group performance
Fifteen-year-old students
Scholastic performance
Heterogeneity
Categorical data
Inverse probability weighting

ABSTRACT

This paper analyzes the evaluation of the relative performance of a set of groups when their outcomes are defined in terms of categorical data and the groups' members are heterogeneous. This type of problem has been dealt with in Herrero and Villar (2013) for the case of a homogeneous population. Here we expand their model controlling for heterogeneity by means of inverse probability weighting techniques. We apply this extended model to the analysis of the scholastic performance of fifteen-year-old students in the OECD countries, using the data in the PISA. We evaluate the relative performance of the different countries out of the distribution of the students' achievements across the different levels of competence, controlling by the students' characteristics (explanatory variables regarding schooling and family environment). We find that differences in mathematical and reading abilities across OECD countries would lower by between 40% and 50% if the students' characteristics would be those for the OECD average.

1. Introduction

We consider here an evaluation problem in which we have to compare the relative performance of several groups, out of the distribution of the achievements of their members in a set of ordered categories. Think for instance of the comparison of the health situation of different countries out of the distribution of the population in four or five health statuses (e.g. from "excellent" to "very bad"). The key elements of the problem are, therefore, the presence of several groups, the qualitative nature of the outcome variable, which resolves into a given set of ordered categories, and the focus on relative performance.

This type of problem has been addressed recently by Herrero and Villar (2013). They start by considering pairwise comparisons between groups in terms of the probability that an agent picked at random in one group belongs to a higher category than an agent randomly chosen in the other. Then they extend those comparisons to all groups involved, by taking into account both direct and indirect relations. As a result they obtain an evaluation function that corresponds to the dominant eigenvector of a matrix that describes all those comparisons (see below). This evaluation function has interesting ethical and operational properties.

* Corresponding author at: Departamento de Economía Aplicada, Facultad de Economía y Empresa, 30100 Espinardo, Murcia, Spain. Tel.: +34 868883732; fax: +34 868883745.
E-mail addresses: ildefonso.mendez@um.es, ildefonso.mendez@gmail.com (I. Mendez).

http://dx.doi.org/10.1016/j.econedurev.2014.02.001
0272-7757/

An implicit assumption in their model is that groups are homogeneous so that the distribution of the outcome variable is the sole relevant information. Yet one might be interested in evaluating not only the observed outcomes but also the extent to which those outcomes reflect diverse structural characteristics of the population that affect the agents' performance. This may well be the case in the example mentioned above (comparing the health situation of different countries), regarding the influence of aspects such as age or wealth in the final outcomes. To deal with this type of evaluation we need a methodology that permits making comparisons in terms of a common set of characteristics. This is the key point of this paper. More specifically, we combine here the original model in Herrero and Villar (2013) with inverse probability weighting (IPW) techniques that permits one controlling for differences in the distribution of the determinants of the outcome variable.

Using this methodological approach we obtain a covariate-adjusted evaluation that allows isolating the impact of the selected explanatory variables, by comparing this evaluation with the unadjusted one. In that way we can separate the part of the observed differences that is explained by the covariates and the part which cannot be accounted for. The covariate-adjusted eigenvector tells us about the relative performance of the groups once their conditioning variables have been equalized. Comparing the covariate-adjusted and the unadjusted evaluations permits one to estimate the impact of the latent variables on the relative performance.

The interpretation of the differences between both evaluations depends on the problem at hand and, in particular, on the choice of the explanatory variables. In this respect our analysis is reminiscent of the "equality of opportunity" literature, as covariate-adjusted values might be interpreted as an expression of the differential "effort", whereas the unadjusted values would reflect the interplay of both effort and "opportunity".[1] Yet, this model does not provide a "measure" of equality of opportunity, as our comparison deals with relative performance both in the adjusted and unadjusted evaluations.

We apply this extended model to the evaluation of the scholastic performance of fifteen-year-old students in the OECD using the data provided by the Program for International Students Assessment (PISA). We evaluate the performance of schoolchildren regarding mathematical ability, out of the 2012 dataset (the last one available). Our evaluation involves the estimation of the impact of the students' environment (parental and school characteristics) on the final scores. Comparing the adjusted and unadjusted evaluations allows concluding that the set of explanatory variables accounts for almost 50% of the differences in the relative performance. We also consider how those differences have evolved in the first decade of the 21st century, by comparing the results in 2012 and those in 2003. We find that differences in students' mathematical ability across OECD countries have decreased during that period, particularly so for European OECD countries.[2]

We also analyze the evolution of country differences in reading competence using the 2009 and 2000 PISA reports, that focused on that subject. We find that the differences in students' reading ability across OECD countries lowered during the first decade of the 21st century to a greater extent that the differences in mathematical competence did. The countries that improve (lower) their relative position in one of the two competences are likely to be those that improve (lower) their relative position in the other competence. Finally, we find that students' environmental factors also account for between 40% and 50% of country differences in relative performance in reading ability.

The paper is organized as follows. Section 2 presents the formal model whereas Section 3 applies it to the results on mathematical competence out of the data in the PISA (2003 and 2012). Section 4 gathers a few final comments.

2. The model

Consider a set of g groups or societies, $G = \{1, 2, \ldots, g\}$, each of which consists of n_i agents, $i \in G$. We want to compare the relative performance of those groups with respect to a given aspect, when their achievements are given in categorical terms. More precisely, we assume that there is a set of categorical positions, $H = \{h_1, \ldots, h_s\}$, ordered from best to worst, $h_1 \succ \cdots \succ h_s$ (health statuses, educational levels, age intervals, professional positions, etc.). Each group presents a given distribution of achievements across those categories. Our goal is comparing their relative performance, taking into account the role of the differences in the structural characteristics that may influence the outcome variable. To do so we divide the evaluation problem into two parts. First, we assume that all groups are homogeneous regarding those characteristics, so that the evaluation only takes into account their relative achievements. The key point here is how to make systematic comparisons out of qualitative data. Second, we consider that groups are heterogeneous and provide a method to control for such heterogeneity.

2.1. The evaluation formula when groups are homogeneous

Let $G = \{1, 2, \ldots, g\}$ stand for a set of g groups under the assumption that they are homogeneous with respect to the aspect under evaluation. Let a_{im}, for $i = 1, \ldots, g$, $m = 1, \ldots, s$, be

[1] Equality of Opportunity (EOp) is one of the most prominent concepts of distributive justice. The key idea behind this concept is that the concern about inequality should not focus on the equality of outcomes but rather on the existence of a common *playing field* for all people. From this perspective agents' outcomes can be regarded as deriving from two different sources: *effort and opportunity*. Effort refers to people's decisions whereas opportunity refers to the agents' external circumstances. A fair society is one in which final outcomes do not depend much on the agents' external circumstances, that is, a society in which all people share similar opportunities. In that society outcome differences are basically determined by the agents' preferences and effort and not by aspects that are beyond their control and responsibility (see Arneson, 1989; Cohen, 1989; Fleurbaey, 2008; Roemer, 1993, 1998).

[2] Here again our analysis is very close to that of equality of opportunity in education. See on this respect Peragine and Serlenga (2008), Lefranc, Pistolesi, and Trannoy (2008), Checchi and Peragine (2010), Calo-Blanco and Villar (2010), OECD (2010b), Villar (2012), and Calo-Blanco and García-Pérez (2013).