Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-05524-1 OPEN

The effect of funding for disadvantaged students on academic underachievement

Oleg Badunenko [1], Deni Mazrekaj [2,3,4], Subal C. Kumbhakar [5,6] & Kristof De Witte [7,8]✉

We estimate the effect of school funding for disadvantaged students on academic underachievement in primary education. For this purpose, we exploit a policy in Belgium that granted additional funding for schools with a share of disadvantaged students of at least 10%. Combining stochastic frontier analysis (a method to measure efficiency) with regression discontinuity design, we find that while additional funding does not affect achievement, it does help reduce underachievement. As a result, the policy has helped bringing students impacted by it closer to realizing their full potential compared to their peers who were not included in the policy. From a policy perspective, our findings indicate that maintaining targeted remedial measures is essential for enhancing effectiveness. Additionally, we emphasize the importance of analyzing both underachievement and achievement.

[1] Department of Economics and Finance, College of Business, Arts and Social Sciences, Brunel University of London, Uxbridge, UK. [2] Utrecht University, Department of Sociology, Utrecht, The Netherlands. [3] University of Oxford, Nuffield College, Oxford, UK. [4] KU Leuven, Leuven Economics of Education Research, Leuven, Belgium. [5] Department of Economics, State University of New York, Binghamton, NY, USA. [6] Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences, Prague, Czech Republic. [7] Leuven Economics of Education Research (LEER), KU Leuven, Leuven, Belgium. [8] United Nations University—Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT), Maastricht, The Netherlands. ✉email: k.dewitte@maastrichtuniversity.nl

Introduction

The highest performing education systems provide all children opportunities for a high-quality education (OECD 2012). To ensure equity and quality across education systems, funding strategies have been developed that allocate various levels of resources to student groups with different needs. Of particular attention are funding strategies that aim to reduce the educational gap between students from different social backgrounds. Students from a disadvantaged background are at a high risk of school dropout and typically take higher instructional costs than other students (OECD 2017). At the same time, schools with more disadvantaged students are likely to have fewer resources available to meet student needs. To account for this discrepancy and to provide equal educational opportunities, countries have devised policies that provide additional funding to schools with a larger share of disadvantaged students.

In this paper, we evaluate a funding policy in the Belgian community of Flanders that granted additional funding to schools with a share of disadvantaged students of more than 10%. Disadvantaged students are identified by the Flemish Ministry of Education based on weighted socioeconomic status characteristics (e.g., the student's mother is a high school dropout). Using unique longitudinal survey data for 1439 Flemish students in 159 schools observed over 6 years of primary education, we estimate whether this policy has been efficient in reducing students' underachievement and, as such, reducing inequality in education. For this purpose, we combine techniques from the causal inference literature (regression discontinuity design—RDD) on the one hand, and the efficiency analysis from production economics on the other to provide a causal interpretation of our findings. In particular, we focus on a narrow bandwidth around the 10% threshold and estimate underachievement using a Stochastic Frontier (SF) model. Stochastic frontier analysis with regression discontinuity design is a statistical method that combines two approaches to analyze data. The RDD estimates the causal effect of a treatment by comparing observations around a cutoff point, in this case the share of disadvantaged students of 10%. The SFA is an econometric method that estimates the maximum possible output (i.e., achievement) given inputs (e.g., ability) and measures the efficiency of achieving that output, accounting for random variation. By combining the two methods, we obtain insights into how efficiency close to the frontier is changing.

Our analysis builds on the educational production function (Hanushek, The Economics of Schooling: Production and Efficiency in Public Schools, 1986) and the human capital theory that suggests that early investment in education, especially for disadvantaged children, has the highest rate of return (Cunha and Heckman 2007). As such, we also include factors other than ability in the stochastic frontier analysis to estimate potential achievement. A key factor here is socioeconomic status. In education-focused stochastic frontier analysis models, socioeconomic status is commonly included as an input because it is a crucial factor that strongly correlates with academic achievement (De Witte and López-Torres 2017). Socioeconomic status can impact educational outcomes through various channels, such as access to educational resources, parental involvement, and neighborhood characteristics. As a result, stemming from a low socioeconomic background adversely affects students' cognitive development and educational attainment. Thus, low socioeconomic status students may not perform at their full potential in school, leading to underachievement; or similarly, high socioeconomic status students are expected to obtain higher test scores. By accounting for socioeconomic status as an input in stochastic frontier analysis models that measure underachievement, it is possible to obtain a more accurate assessment of students' underachievement.

Previous literature has exclusively focused on low achievement, not on underachievement, producing mixed results of additional funding (Bénabou et al. 2009; Chay et al. 2005; Hanushek 2003; Henry et al. 2010; Jackson et al. 2016; Leuven et al. 2007; Ooghe 2011; van der Klaauw 2008). Low achievers are students who are achieving to the full extent of their abilities but have lower achievement compared to their peers. By contrast, underachievers are students who exhibit a severe discrepancy between expected achievement and actual achievement, with this discrepancy not occurring due to a learning disability (McCoach and Siegle 2003). Thus, although it is useful to know whether additional funding has raised overall achievement, it is also useful to know whether students are achieving according to their potential. For instance, it is possible that the overall achievement rises after schools receive additional funding, but that students are still not achieving according to their potential and that further gains could be obtained. We therefore contribute to the literature by providing causal estimates of whether school funding for disadvantaged students is effective in combatting underachievement in primary education.

Whereas identifying low achievers is straightforward by observing achievement outcomes such as test scores, the identification of underachievers is difficult as potential outcomes are unobserved (Mazrekaj et al. 2022). For instance, gifted underachievers typically have average or even high outcomes, but still perform below their potential. Therefore, underachievement needs to be explicitly modeled. We use the insights from production economics and model underachievement using a stochastic frontier model. Intuitively, we estimate an achievement frontier to obtain students' potential test scores (i.e., given their IQ and socioeconomic background) and compare this frontier with students' actual test scores (Mazrekaj et al. 2022). If students deviate from this achievement frontier, they do not reach their full potential, and we consider them as underachieving. We contribute to the previous literature on underachievement by distinguishing between underachievement that is persistent (i.e., long run) and difficult to tackle over time, and underachievement that is transient (i.e., short run) based on the four-components stochastic frontier model (Badunenko et al. 2021). Note that our model explicitly tackles a potentially mechanical relationship between achievement and underachievement, namely if student achievement increases as a result of the intervention, then the gap between predicted and actual scores will go down. To distinguish between both concepts—and avoid this mechanical relationship—we first estimate the gap between the predicted and the actual scores, and second, relate this to a performance benchmark which compares a pupil's performance to the potential of other similar pupils. Given that the potential of these other students can be greater or smaller than the pupil's own potential, we avoid the mechanical relationship between the effect of the policy on achievement and underachievement.

Having obtained estimates of underachievement, we extend the stochastic frontier analysis with a Regression Discontinuity Design (RDD) and provide a causal interpretation of additional funding for disadvantaged students. By focusing on a narrow bandwidth on each side of the 10% threshold, we can compare underachievement of students in schools that are similar on observed and unobserved characteristics, with one school receiving the funding and the other schools not receiving the funding. At the threshold, these schools should be very similar apart from the funding received, enabling a causal interpretation of the findings (Leuven et al. 2007).

Institutional setting

Flemish education system. The Flemish education system provides compulsory education beginning at the age of six and continuing either until the age of 18 or until a younger age if a student has already obtained a high school diploma. Most children enter primary education at age 6. However, parents may decide to enroll their child into primary education already at the age of five or they can decide that their child is not ready to enroll at the age of six, and subsequently wait one extra year. Primary education lasts for 6 years until the age of 12. If students complete all 6 years of primary education, they receive a certificate of primary education.

Primary education in Flanders is a suitable setting to study underachievement, because parents may choose any elementary school for their child, there are no catchment areas, ability grouping nor standardized tests (Palmaccio et al. 2022). The enrollment takes place on a first-come-first-serve basis until the capacity of the school is reached. The school decides autonomously how students are distributed among classes. Given that teachers are likely to teach for the average student in heterogeneous classes (Van Klaveren and De Witte 2014), underachievement may occur for the entire distribution of students. One teacher teaches all the subjects and each school year, a new teacher is assigned to the class. Typically, about 440,000 students are enrolled into primary education each year, of which about 10% are of non-Belgian origin—mostly students with Moroccan and Turkish parents—and about 20% does not speak Dutch at home (Mazrekaj et al. 2022).

"Equal Educational Opportunities" policy. To reduce the educational gap between students from different social backgrounds, the Flemish Ministry of Education introduced the "Equal Educational Opportunities" (EEO) policy in 2002. This policy grants additional funding for primary schools with a weighted share of disadvantaged students of at least 10%. To decide whether a student is from a disadvantaged background, the Flemish Ministry of Education employs four binary demographic indicators (D'Inverno et al. 2021): (1) the student is not living with one of the biological parents (weight: 0.8), (2) the student's family belongs to a traveling population e.g., Roma or circus family (weight: 0.8), (3) the income of the student's household consists only of replacement income (weight: 0.4), and (4) the student's mother is a high school dropout (weight: 0.6). If the weighted share of disadvantaged students exceeds 10%, the school receives additional funding. The additional funding is fixed for a period of 3 years and is meant as personnel funding: it can be used for hiring additional teachers and for teacher support.

The Flemish education budget is "open", indicating that the student body (both in terms of size and constitution) determines the budget. Consequently, there is no predetermined budget for the EEO resources. Nevertheless, as an indication, thanks to the EEO policy, there was an increase in the total number of full-time equivalent teachers by 4.6% in the period 2002–2005. For an individual school with many disadvantaged students, the additional resources result in an increase of up to 6 FTE teachers per 256 students. This allows schools with a very disadvantaged student population to halve class size. For schools close to, but above, the 10% threshold, the additional resources result in the equivalent of 0.25 FTE teacher (per 256 students). Given the financing implications, there is a centralized database with administrative information on all relevant student characteristics. The data are carefully checked by the administration. With a few exceptions, schools receive the funding automatically, based on the student population at the beginning of the 3 years cycle.

In practice, using its detailed administrative data, the Ministry is simply looking at the percentage of disadvantaged students and funds the schools accordingly. In the present paper, we use the information available in our dataset to mimic this funding decision. Although most schools that are far away from the 10% threshold received already extra funding from the first wave, this is not the case for schools close to the threshold. Moreover, also the level of the funding varies, depending on the student population. Despite schools having full autonomy to use the resources, the education inspectorate examines whether the resources are used to assist the disadvantaged target group. Most schools are using the resources to hire additional teachers, who help in reducing class size or who provide additional help to students who struggle academically.

Methods

Data. We use Flemish data from the first until the last grade of primary education (grades 1 to 6). We draw on the longitudinal SiBO (*Schoolloopbanen in het Basisonderwijs*) database in which approximately 6138 students were followed from age six until age twelve in the Flemish region of Belgium. As the present study works with secondary data that were already collected, an ethical approval was not required. Moreover, the data do not contain any sensitive or personal information that could potentially identify individuals. A stratified sampling approach was used at the school entity level (an administrative entity that may include several schools) based on two criteria: type of education (community education, official subsidized education, and free subsidized education), and school entity size (the smallest 25%, the middle 50%, and the largest 25%). Within each stratum, school entities were ranked based on the percentage of disadvantaged students and the province. The goal was to oversample disadvantaged students as they were the focal point of the study. Then, school entities were selected per stratum according to a set interval. In total, 120 school entities were contacted to participate in 2002 and 77.5% agreed. The rest of the school entities were replaced by other school entities using the same approach. The overall response rate was rather high at 76%. The dataset comprises 196 schools for the period 2002–2009. We focus on a subsample of the data for which information on the socioeconomic status is available to identify disadvantaged students. Namely, the parental survey from which the socioeconomic status measure was constructed was administered to a subsample of 3534 pupils only, and not everyone filled in the survey. The final sample includes 2282 students in 165 schools observed over 6 grades of primary education, totaling 12,738 observations. It should be noted that this is an unbalanced panel dataset.

We conduct the analysis at the student level. Our educational production function (Hanushek 1986) consists of the log of math test scores as the output variable. These tests were taken at the end of each school year and have been specifically tailored for each grade by the survey administrators. The tests have a Cronbach's alpha of 0.89, indicating high reliability. The dataset also includes language tests (reading and writing), but these tests are subdivided into five different tests, each consisting of two different versions with varying reliability. This makes it very difficult to interpret and compare these tests. We therefore solely focus on the mathematics tests. We do not consider raw test scores nor standardized test scores because we need to take the logarithm of the math score to construct a measure of underachievement when using a stochastic frontier model (see next section). Taking the logarithm of standardized test scores would remove all the observations with negative test scores.

We follow the previous literature on underachievement and efficiency in education for selecting various student-level inputs (De Witte and López-Torres 2017; Mazrekaj et al. 2022). We consider the IQ test score that is inspired by the CIT-3-4 verbal cognitive test (Stinissen et al. 1975) and the non-verbal Raven's Standard Progressive Matrices Test (Raven 2000). This IQ test was administered in the third grade and has a Cronbach's alpha of 0.93. We also include an indicator for gender (1 is male and 0 is female), origin (1 if at least one of the parents was born abroad and 0 otherwise), and several measures of socioeconomic status. The latter include language at home (1 is foreign, 0 is Dutch), at least one of the parents is of foreign origin (2 is both parents of foreign origin, 1 is one parent of foreign origin, 0 is both parents Belgian), diploma of the mother (1 is maximum upper secondary education, 0 is otherwise), diploma of the father (1 is maximum upper secondary education, 0 is otherwise), net income of the household in euros in six categories, and the worker status of the father (1 is fulltime, 0 is other), and indicators for school grade. These variables were collected by surveying students' parents in the first grade. Finally, to determine the share of disadvantaged students, we follow the definition set by the Ministry of Education as outlined in the "Institutional setting" section.

Regression discontinuity design. To estimate the effect of additional funding for disadvantaged students on achievement, we use a sharp regression discontinuity design approach (for detailed information, see Calonico et al. 2014, 2017). Recent developments in the RDD methodology can be found in Cattaneo et al. (2019, 2023). This approach is based on a running variable with a well-defined threshold that determines treatment assignment. In our setting, the running variable is the share of disadvantaged students in a school, the threshold is 10%, and this threshold determines whether a school received funding or not. The RDD is sharp rather than fuzzy because all schools that were above the threshold received the funding, and all schools that were below the threshold did not receive the funding. Conceptually, RDD compares the math test scores of students in schools who have received the funding and scores of students who have not received the funding in a locality around the threshold of 10% disadvantaged students that was defined exogenously by the policy. Then, the outcomes are generalized to judge the effectiveness of the policy.

We denote the outcome variable as Y and the running variable as c. If $c \geq 10$, the policy is applied and schools received additional funding. We call the students in these schools *treated* students. If $c < 10$, the policy is not applied and schools did not receive additional funding. We call the students in these schools *control* students. To determine whether policy had an effect on achievement, or to find the treatment effect, the following steps are performed. First, the sample of all students is split into two subsamples: a subsample of treated students and a subsample of control students. We consider three bandwidths: 2% and 4% around the 10% threshold that we chose intuitively, and 2.5% that is obtained by the data-driven "optimal bandwidths" as in Calonico et al. (2015). Then, the subsample of control students is used to perform a local regression of Y on c and to obtain the slope coefficient β^C. Analogously, the subsample of treated students is used to perform a local regression of Y on c and to obtain the slope coefficient β^T. Third, we calculate the treatment effect β^{TC} as the difference between β^T and β^C. If it is positive, the policy has improved the outcome variable Y.

To gauge the statistical significance of the treatment effect, i.e., obtain the standard error of the difference between β^T and β^C, researchers use another approach. A dummy variable D is created which is equal to 1 for the treated sample and 0 for the controlled sample. Then the performance equation is estimated:

$$Y_{it} = \alpha + \tau D + \beta^C(c_{it} - 10) + \beta^{TC} D(c_{it} - 10) + \varepsilon_{it} \qquad (1)$$

whereby β^{TC} is obtained directly as a coefficient of the interaction term between the dummy variable and the difference between the running variable and the threshold. The attractiveness of the latter approach is that the statistical significance of the treatment effect is obtained directly. Note that it is not possible to obtain the standard error of the treatment effect when the two equations are estimated separately. The conclusion about the effectiveness of the policy is made based not only on the sign of τ but also whether it is statistically different from 0. Note that we use Maximum Likelihood Estimation in which we make the error components depend on individual characteristics. Clustering at school level is thus not needed in this case.

Under the assumption of continuity—all observed and unobserved factors besides the treatment and the outcome should be continuous at the threshold—RDD can be seen as a local randomized experiment, namely a randomized experiment at the threshold (Lee and Lemieux 2010). We assess the continuity assumption using the manipulation test discussed by Cattaneo et al. (2018). More specifically, the test is conducted on 893 schools. We observe 147, 145, 148, 148, 149, and 156 schools in the years from 2003 to 2008. Sixty-two schools are below the cut-off of 10%, and 832 are above the cut-off. Using different combinations of the local approximation or order of the local polynomial used to construct the point estimator (quadratic and cubic) and kernel (triangular and uniform), we obtain p values of 0.0672, 0.0871, 0.1298 and 0.3037 thus not rejecting the null hypothesis of no manipulation. Therefore, the continuity assumption appears to hold in our setting. We also show the descriptive statistics for treated and control subgroups in Table 1 when using our preferred bandwidth of 2.5% around the 10% threshold. It appears that the two groups are similar on observed characteristics around the threshold, with the exception of gender and the share of students whose mothers have at most upper secondary education.

Finally, to check the balance on the covariates to test the validity of the RD design, we perform statistical analysis to test presence of discontinuities at the cutoff, the results of which are shown in Table 2. Note that the number of effective observations is different since the procedure chooses bandwidth for each covariate separately. Except for the three income groups (2000–4999 EUR), all point estimates are small, and the 95% robust confidence intervals include zero, with large p values. The empirical evidence suggests that the covariates are not discontinuous at the cut-off.

There are several limitations of the current framework. First, the identification includes both an increase in funding and an increase in the running variable (i.e., presence of disadvantaged students). However, close to the threshold, the percentage of disadvantaged students is very similar just below and above the threshold. We acknowledge that potentially even one student might make a difference, but the marginal difference decreases as the percentage of disadvantaged students increases in the school. For example, in an average school of 256 students, it is unlikely that a teacher of an individual class will notice that there are 23 rather than 26 disadvantaged students in the school. However, it is very likely that this individual teacher will observe the difference in funding (e.g., because of hiring of new teachers, such that there is more time to help the disadvantaged students in a tutoring program). Second, as we do not have information on the exact funding that schools received, we estimate intention-to-treat (ITT) regressions, basing estimates on eligibility rather than on remedial funds actually received. Future research should aim to estimate the effect of the funding actually received on achievement and underachievement, and might also explore heterogenous effects by gender.

Stochastic frontier analysis. Different students can obtain a certain math score Y differently. Some students are using limited

Table 1 The effect of funding for disadvantaged students on log math test scores.

	Sample: to the left and to the right of the 10% threshold				
	2%	4%	2.5%		
			All	Girls	Boys
Received funding	0.113 (0.37)	0.058 (0.60)	0.110 (0.58)	0.119 (0.41)	0.114 (0.47)
Additional controls	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.289	0.283	0.281	0.266	0.266
Number of students (below and above the threshold)	377 (69, 308)	620 (94, 526)	408 (77, 331)	228 (46, 182)	180 (31, 149)
Number of schools (below and above the threshold)	36 (9, 27)	69 (18, 51)	44 (12, 32)	43 (12, 31)	42 (12, 30)
Observations	1172	2280	1404	809	595

t-statistics based on the robust standard errors are shown in parentheses. Control variables include gender (1 is male and 0 is female), origin (1 if one of the parents was born abroad and 0 otherwise), language at home (1 is foreign, 0 is Dutch), diploma of the mother (1 is maximum upper secondary education, 0 is otherwise), diploma of the father (1 is maximum upper secondary education, 0 is otherwise), net income of the household in euros in six categories, and the worker status of the father (1 is fulltime, 0 is other), and indicators for school grade.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$ (two-tailed *t*-tests).

Table 2 The effect of funding for disadvantaged students on underachievement.

	Sample: to the left and to the right of the 10% threshold				
	2%	4%	2.5%		
			All	Girls	Boys
Received funding	−6.5043*** (−2.86)	−1.3022* (−1.81)	−4.3065*** (−2.62)	−5.1427 (−0.01)	−4.1034 (−0.01)
Additional controls	Yes	Yes	Yes	Yes	Yes
Students	377	620	408	228	180
Observations	1172	2280	1404	809	595

z-statistics are shown in parentheses. Control variables include gender (1 is male and 0 is female), origin (1 if one of the parents was born abroad and 0 otherwise), language at home (1 is foreign, 0 is Dutch), diploma of the mother (1 is maximum upper secondary education, 0 is otherwise), diploma of the father (1 is maximum upper secondary education, 0 is otherwise), net income of the household in euros in six categories, and the worker status of the father (1 is fulltime, 0 is other), and indicators for school grade.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$ (two-tailed *t*-tests).

resources efficiently, whereas other students underachieve given available resources. To allow for such a possibility, the error term ε in Eq. (1) is broken down into a usual statistical noise v and underachievement u—in stochastic frontier literature called inefficiency term (see Mazrekaj et al. 2022 for details). Hence the following performance equation is estimated:

$$Y_{it} = \alpha + \tau D + \beta^C(c_{it} - 10) + \beta^{TC}D(c_{it} - 10) + \nu_{it} - u_{it} \quad (2)$$

The underachievement term u shows by how much the outcome variable Y is lower in contrast to the best possible outcome if a student has used their resources efficiently. Said otherwise, we estimate underachievement as the difference between the potential test scores students could achieve given observable characteristics and the actual test scores students achieved. The potential test scores are determined based not on student's own potential (e.g., predicted vs. awarded test score) but on the other students in the sample with similar observed characteristics. The stochastic frontier is $\alpha + \tau D + \beta^C(c_{it} - 10) + \beta^{TC}D(c_{it} - 10) + \nu_{it}$, and each student is then benchmarked against this frontier. Given that the difference between the actual performance and the benchmark is given by u, we call it underachievement because u measures the shortfall between the benchmark (maximum possible score) with the actual score. Availability of panel data allows us to further decompose the underachievement term u into a persistent and a transient underachievement component, i.e., $u_{it} = u_i^P + u_{it}^T$. Persistent underachievement can be thought of as a long-term underachievement that is very difficult if at all possible to eliminate. Transient underachievement is a short-term underachievement, which can be eradicated easier by for example providing remedial teaching or varying the class size. To account for the heterogeneity of students, the error term ν_{it} can be decomposed into the pure noise η_{it} and individual effect μ_i, which stands for unaccounted heterogeneity. We therefore extend Eq. (3) so that the error structure comprises four components:

$$Y_{it} = \alpha + \tau D + \beta^C(c_{it} - 10) + \beta^{TC}D(c_{it} - 10) + \eta_{it} + \mu_i - u_i^P - u_{it}^T \quad (3)$$

The technical difference between η_{it} and u_{it}^T is that the former can be both positive and negative, while the latter is only positive. The same technical difference exists between μ_i and u_i^P. Conceptually, individual effects can account for aspects not controlled for in a regression such as for example family circumstances or health conditions, which may have positive or negative effects on the outcome.

Both persistent and transient underachievement can be modeled to include determinants. Due to nonlinear nature of the specification, the coefficient τ is not informative and additional calculations are required. More specifically, if the log of variance of transient inefficiency $\log \sigma_T^2$ is given by $z\gamma$, where z would contain D, $(c_{it} - 10)$, the marginal effect of $D(c_{it} - 10) \times D$ is given by $\sqrt{1/2\pi} \times \tau \times e^{0/5z\gamma}$.

We note that the parameters of the frontier in Eq. (3) as well as parameters of underachievement are estimated in a single step. Thus, the econometric issues associated with a two-step procedure where first u_{it}^{T} and u_{i}^{P} are estimated and then regressed on the determinants are avoided.

We further note that the achievement variable math score appears in its logarithmic transformation in the analyses. This comes from a definition of efficiency as a ratio of the observed outcome variable to the potential: $\textit{efficiency} = Y/Y^{*}$, where Y^{*} denotes potential outcome and Y denotes the observed outcome. To ensure efficiency is smaller or equal to 1 and is larger than 0, it is specified as $\exp(-u)$, where u is inefficiency (underachievement). Taking logs of the equation $\exp(-u) = Y/Y^{*}$, we obtain $\log Y = \log Y^{*} - u$. Here $\log Y^{*}$ represents the frontier, which is specified by the researcher. Because the dependent variable is logged, we resort to the percent interpretation. More precisely, we provide a scale independent interpretation, i.e., an increase of "input" by 1% (each school would know how much it is) results in so many percent increase in achievement or underachievement.

Results

The effect of school funding on achievement. Given that the previous literature has exclusively focused on low achievement rather than underachievement, we first evaluate the impact of additional funding for disadvantaged students on achievement. As a measure of achievement, we consider the mathematics test score as this measure has a high reliability in our dataset (Cronbach's alpha of 0.89). Specifically, we take the logarithm of the math score as this is necessary to later construct a measure of underachievement when using a stochastic frontier model (see "Methods" section). Figure 1 visualizes the relationship between achievement and the share of disadvantaged students. Said otherwise, we plot two regression lines of the log of mathematics test score on the share of disadvantaged students in a school—one on each side of the threshold—while allowing for different slopes.

We observe in Fig. 1 that the math scores decrease as the share of disadvantaged students goes up to the threshold of 10%. At the threshold, we observe a drop in achievement, after which the math test scores decrease even further. The figure also shows that the association between achievement and the share of disadvantaged students is different between students in schools below and above the 10% threshold. This 10% threshold is where the policy is applied.

An important caveat of the figure is that it shows the relationship between achievement and the share of disadvantaged students for all schools, whether these schools have no disadvantaged students or only disadvantaged students. However, it is likely that students in these schools have very different characteristics. As we are interested in the causal effect of the funding policy on achievement, we now focus only on the schools near the threshold of 10%. By focusing on a narrow bandwidth on each side of the 10% threshold, we can compare test scores in schools that are similar on observed and unobserved characteristics, with one school receiving the funding and the other schools not receiving the funding (Leuven et al. 2007). We show that observed characteristics are very similar for students in schools who did and did not receive the funding in Table 1 in the "Methods" section.

The way "closeness" near the threshold is defined may influence the analysis. To avoid such ambiguity, we perform four regression discontinuity design models considering schools within 2%, 4%, and 2.5% of the threshold. For instance, considering schools within 2% of the threshold implies that we compare the achievement of students in schools in which the percentage of disadvantaged students is between 8% and 12%. This restriction reduces the sample to 377 students. The samples that consider schools within 4% and 2.5% of the threshold comprise 620, and 408 students, respectively. While the choice of 2%, and 4% is dictated by intuitive closeness to the 10% threshold, the 2.5% is obtained by the data-driven "optimal bandwidths" discussed in Calonico et al. (2020).

Table 2 presents the results of the RDD models in which we also control for students' gender, IQ, language that is spoken at home, migration status of parents, education of both mother and father, each parent's working status, and the household income. For the 2.5% sample, we also conducted the analysis where we do not include any covariates. Because the results are very similar, we proceed with the analysis that includes covariates as is common in the previous literature. For policy evaluation purposes, the interest lies in the "received funding" variable that can be interpreted as the causal effect of receiving additional funding for disadvantaged students on mathematics test scores at various thresholds.

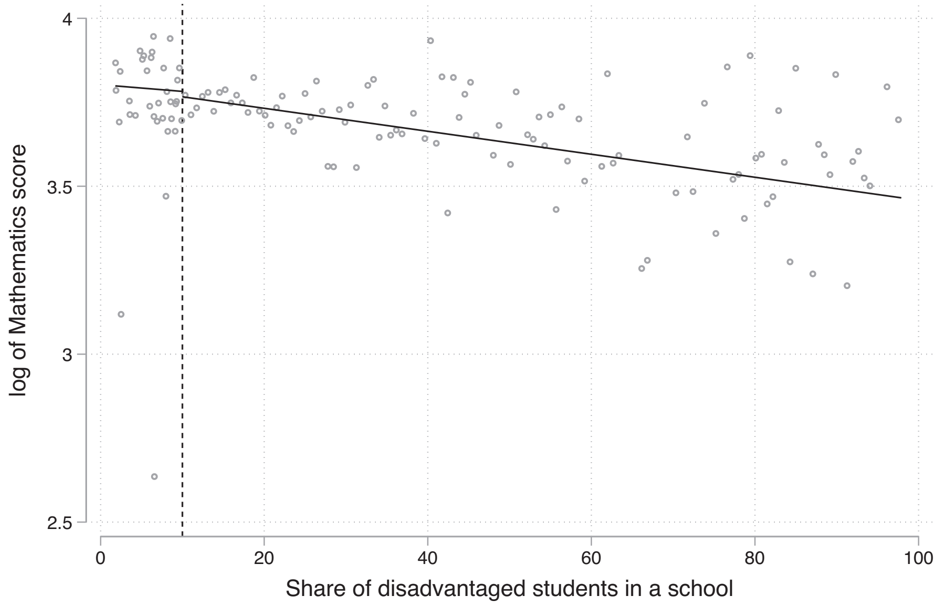

Fig. 1 The relationship between math test scores and the share of disadvantaged students. The figure shows two regression lines of the log of mathematics test score on the share of disadvantaged students in a school—one on each side of the threshold of 10%—while allowing for different slopes.

Table 3 Descriptive statistics (2.5% bandwidth).

	Full sample ($N = 1404$)	Treated ($N = 291$)	Control ($N = 1113$)	*p* value of the equality of means test
IQ	0.140	0.140	0.141	0.342
Gender	0.424	0.439	0.364	0.021
Origin (1 is foreign)	0.080	0.084	0.066	0.275
Language at home (1 is foreign)	0.152	0.158	0.127	0.190
Diploma mother (1 is max upper sec. edu.)	0.082	0.090	0.079	0.034
Diploma father (1 is max upper sec. edu.)	0.110	0.115	0.089	0.213
Worker status father (1 is fulltime)	0.080	0.084	0.065	0.065
Net income household				
Less than 1000 EUR	0.002	0.001	0.002	0.113
1000–1999 EUR	0.167	0.158	0.169	0.659
2000–2999 EUR	0.422	0.406	0.427	0.513
3000–3999 EUR	0.260	0.289	0.253	0.210
4000–4999 EUR	0.081	0.069	0.084	0.408
5000 EUR or more	0.053	0.072	0.048	0.095

Table 3 presents the results for the 2, 4, and 2.5% samples. By varying the bandwidth, we examine how the estimated treatment effect changes as we include more or fewer observations in the analysis. This can help us to understand the robustness of the findings and assess the sensitivity of the results to different choices of bandwidth. In those samples, the “received funding” dummy variable is not statistically significant. Overall, it appears from Table 3 that additional funding for disadvantaged students did not influence student achievement and this conclusion holds for both boys and girls.

The effect of school funding on underachievement. Until now, we have followed the previous literature by focusing on achievement: whether students score high or low on mathematics test scores. However, we are at least as equally interested in underachievement: whether students can reach their potential (Mazrekaj et al. 2022). We therefore estimate students' underachievement using a four-components stochastic frontier model while focusing on the narrow bandwidth around the 10% threshold. As such, we obtain causal estimates of additional funding for disadvantaged students on underachievement, rather than achievement.

Our model allows us to further separate underachievement into two interrelated parts: long- and short-term underachievement. Long-term underachievement should be understood as an intrinsic underperformance of a student. It may be due to distinct types of skills that a student has or the student's long-term goals that do not focus on math. Long-term underachievement cannot be easily remedied. By contrast, dealing with short-term underachievement is less challenging. For example, short-term underachievement may be remedied by additional tutoring, or by varying the class size. Figure 2 indicates that the median overall underachievement is about 30%, implying that the math score of an average underperforming student could have been 30% higher. The results suggest that the underachievement of 30% is almost evenly split between short- and long-term underachievement. Specifically, the mean overall, persistent and transient underachievement amount to 30%, 15%, and 18%, respectively.

Table 4 shows the effect of funding for disadvantaged students on overall underachievement. In contrast to overall achievement, it appears that the policy *did* significantly affect overall underachievement. Regardless of the bandwidth, the coefficients are significantly different from zero. The coefficients are not significant when we consider samples that include only boys or girls. However, it appears the sample is very small and even though the bias of the potential effect of the policy is arguably small given the narrow bandwidth, the variation is large.

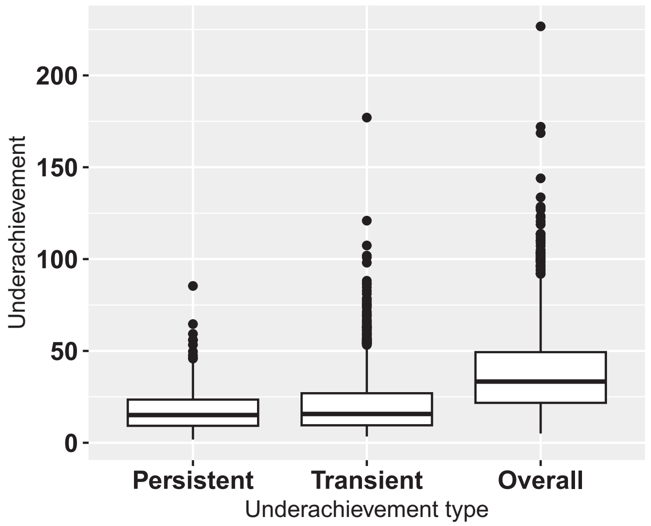

Fig. 2 Underachievement decomposed into a long term (persistent) and a short term (transient) component. The median is represented by the thick horizontal line.

Conclusion

To reduce the achievement gap between disadvantaged and other students, many education systems have developed equal educational opportunity programs. A common element in these programs is additional funding for schools with disadvantaged students. Exploiting a policy in the Flemish region of Belgium that grants additional funding for schools with a share of disadvantaged students of at least 10%, we measure how the additional funding impacts achievement and underachievement in mathematics in primary education. The latter is defined as students who do not reach their full potential for a given IQ level and socioeconomic status.

We did *not* observe an overall effect of the policy on achievement of students. With respect to reaching the full potential of students, we observed that underachievement (i.e., students performing below their potential) was significantly reduced as a result of the policy. Thus, although the policy has not led to higher student achievement, many of the students were reaching their full potential and could catch up to their peers. Our

Table 4 Formal balance analysis for the covariates.

	CER-optimal bandwidth	RD estimator	Robust inference			Effective number of observations
			p value	95% confidence interval		
				Lower bound	Upper bound	
IQ	1.46	−0.54	0.1196	−1.28	0.15	1389
Gender	0.95	−0.02	0.7707	−0.24	0.18	876
Origin (1 is foreign)	1.61	−0.03	0.6439	−0.16	0.10	1507
Language at home (1 is foreign)	0.69	0.03	0.4459	−0.03	0.07	718
Diploma mother (1 is max upper sec. edu.)	1.67	0.03	0.3749	−0.04	0.11	1350
Diploma father (1 is max upper sec. edu.)	1.09	0.07	0.3984	−0.10	0.25	906
Worker status father (1 is fulltime)	1.27	−0.06	0.4613	−0.23	0.11	1075
Net income household						
Less than 1000 EUR	1.68	0.02	0.1507	−0.01	0.03	1147
1000–1999 EUR	1.67	−0.07	0.2544	−0.22	0.06	1147
2000–2999 EUR	1.02	0.33	0.0010	0.14	0.54	742
3000−3999 EUR	0.70	−0.50	0.0002	−0.79	−0.24	518
4000–4999 EUR	0.88	0.09	0.0029	0.03	0.16	600
5000 EUR or more	0.99	0.06	0.0468	0.00	0.14	650

analysis therefore highlights the importance of considering underachievement in addition to achievement. We show that stochastic frontier models are appropriate to measure and account for underachievement.

From a policy perspective, the results indicate that equal educational opportunities are possible, but require a significant and continued investment. Our results indicate that providing additional funding to schools with disadvantaged students can lead to an improvement in mathematics underachievement. This suggests that policies targeting the allocation of resources to schools with high concentrations of disadvantaged students can be an effective strategy for reducing the gap in reaching the potential. Many education systems worldwide suffer from declining achievement in international tests (e.g., PISA, TIMSS) and high inequality in test scores, especially as a result of the school closures during the COVID-19 pandemic (Mazrekaj and De Witte 2023). Our results can inform the design and implementation of similar programs in other education systems, helping to ensure that resources are allocated efficiently to improve student outcomes and reduce underachievement. Additionally, the study highlights the importance of considering underachievement, which can provide a more nuanced understanding of the impact of policies. By using stochastic frontier models to measure and account for underachievement, education systems can gain insights into the efficiency of their policies and adjust them as needed. Thus, our findings show that students are more likely to reach their full potential and additional funding is responsible for this accomplishment. Given their IQ and background, many students underperform already in primary education. In our knowledge-intensive societies that should leave no one behind, more attention to underachieving students is necessary to reduce inequality in education.

Data availability

The data for this study are protected by a confidentiality agreement and we are precluded from sharing the data with others. Interested readers can contact the corresponding author for information on how to access the data. The R codes to replicate the results are available upon request.

Received: 26 August 2024; Accepted: 4 July 2025;
Published online: 25 August 2025

References

Badunenko O, Mazrekaj D, Kumbhakar S, De Witte K (2021) Persistent and transient efficiency in adult education. Empir Econ 60(6):2925–2942

Bénabou R, Kramarz F, Prost C (2009) The French zones d'éducation prioritaire: much ado about nothing? Econ Educ Rev 28(3):345–356

Calonico S, Cattaneo MD, Farrell MH (2020) Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. Econ J 23(2):192–210

Calonico S, Cattaneo MD, Titiunik R (2014) Robust data-driven inference in the regression-discontinuity design. Stata J 14(4):909–946

Calonico S, Cattaneo MD, Titiunik R (2015) Optimal data-driven regression discontinuity plots. J Am Stat Assoc 110(512):1753–1769

Calonico S, Cattaneo MD, Farrell MH, Titiunik R (2017) rdrobust: software for regression-discontinuity designs. Stata J 17(2):372–404

Cattaneo MD, Idrobo N, Titiunik R (2019) A practical introduction to regression discontinuity designs: foundations. Cambridge University Press

Cattaneo MD, Idrobo N, Titiunik R (2023) A practical introduction to regression discontinuity designs: extensions. Cambridge University Press

Cattaneo M, Jansson M, Ma X (2018) Manipulation testing based on density discontinuity. Stata J 18(1):234–261

Chay KY, McEwan PJ, Urquiola M (2005) The central role of noise in evaluating interventions that use test score to rank schools. Am Econ Rev 95(4):1237–1258

Cunha F, Heckman J (2007) The technology of skill formation. AEA Pap Proc 97(2):31–47

De Witte K, López-Torres L (2017) Efficiency in education: a review of literature and a way forward. J Oper Res Soc 68(4):339–363

D'Inverno G, Smet M, De Witte K (2021) Impact evaluation in a multi-input multi-output setting: evidence on the effect of additional resources for schools. Eur J Oper Res 290(3):1111–1124

Hanushek EA (1986) The economics of schooling: production and efficiency in public schools. J Econ Lit 24(3):1141–1177

Hanushek EA (2003) The failure of input-based schooling policies. Econ J 113:F64–F98

Henry GT, Fortner KC, Thompson CL (2010) Targeted funding for educationally disadvantaged students: a regression discontinuity estimate of the impact on high school student achievement. Educ Eval Policy Anal 32(2):183–204

Jackson KC, Johnson RC, Persico C (2016) The effects of school spending on educational and economic outcomes: evidence from school finance reforms. Q J Econ 131(1):157–218

Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. J Econ Lit 48:281–355

Leuven E, Lindahl M, Oosterbeek H, Webbink D (2007) The effect of extra funding for disadvantaged pupils on achievement. Rev Econ Stat 89(4):721–736
Mazrekaj D, De Witte K (2023) The impact of school closures on learning and mental health of children: lessons from the COVID-19 pandemic. Perspect Psychol Sci. https://doi.org/10.1177/17456916231181108
Mazrekaj D, De Witte K, Triebs TT (2022) Mind the gap: measuring academic underachievement using stochastic frontier analysis. Except Child 88(4):442–459
McCoach BD, Siegle D (2003) The school attitude assessment survey-revised: a new instrument to identify academically able students who underachieve. Educ Psychol Meas 63(3):414–429
OECD (2012) Equity and quality in education: supporting disadvantaged students and schools. OECD Publishing, Paris
OECD (2017) The funding of school education: connecting resources and learning. OECD Publishing, Paris
Ooghe E (2011) The impact of 'Equal Educational Opportunity' funds: a regression discontinuity design. IZA Discuss Pap 5667:1–22
Palmaccio S, Schiltz F, De Witte K (2022) Information shocks and parental response in education. A case study of an open government initiative. Gov Inf Q 39(3):101702
Raven J (2000) The Raven's progressive matrices: change and stability over culture and time. Cogn Psychol 41(1):1–48
Stinissen J, Smolders M, Coppens-Declerck L (1975) Handleiding bij de Collectieve Verbale Intelligentietest voor derde en vierde leerjaar (CIT-3-4) [Manual for the collective verbal intelligence test for grades 3 and 4 (CIT-3-4]. Centrum voor Studie en Beroepsoriëntering, Brussels, Belgium
van der Klaauw W (2008) Breaking the link between poverty and low student achievement: an evaluation of Title I. J Econ 142(2):731–756
Van Klaveren C, De Witte K (2014) How are teachers teaching? A nonparametric approach. Educ Econ 22(1):3–23

Acknowledgements
KDW and OB acknowledge financial support from the Horizon Europe project BRIDGE (Grant Number 101177154). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the granting authority. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions
Oleg Badunenko: formal analysis, data curation, investigation, methodology, visualization, validation, writing—original draft, writing—review & editing. Deni Mazrekaj: conceptualization, data curation, validation, writing—original draft, writing—review & editing. Subal C. Kumbhakar: methodology, writing—original draft, writing—review & editing, supervision. KDW: conceptualization, funding acquisition, writing—original draft, writing—review & editing, supervision.

Competing interests
The authors declare no competing interests.

Ethical approval
Ethical approval was not required as the study did not involve human participants. The article uses anonymized secondary data SiBO (*Schoolloopbanen in het Basisonderwijs*) and does not contain any studies with human participants performed by any of the authors. The data do not contain any identifiable sensitive or personal information.

Informed consent
Ethical approval was not required as the study did not involve human participants. The article uses anonymized secondary data SiBO (*Schoolloopbanen in het Basisonderwijs*) and does not contain any studies with human participants performed by any of the authors. The data do not contain any identifiable sensitive or personal information.

Additional information
Correspondence and requests for materials should be addressed to Kristof De Witte.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.