

Teacher Effort in the Classroom

Ofer Malamud*
Matias Martinez†
Andreea Mitrut+
Cristian Pop-Eleches±
Miguel Urquiola±

October 2026

Abstract

We examine how teacher effort varies with student characteristics using novel student-specific measures of relative teacher effort in Romanian high schools. We document four main findings. First, teachers report allocating more effort to disruptive students, consistent with the demands of classroom management. Second, teachers report greater effort towards lower-ability students, consistent with a compensatory allocation of effort. Third, girls receive less teacher effort than boys, but this gender difference disappears, or reverses, when conditioning on ability and disruptiveness. Fourth, teachers report less enjoyment when teaching disruptive and lower-ability classrooms. We discuss the implications for debates on retaining and rewarding teachers.

For useful comments, we thank participants at ERMAS 2025 conference. We thank the Center for Development Economics and Policy (CDEP) and the Faculty Grants Program at SIPA (Columbia) for financial support. We thank Naman Garg for outstanding research assistance. Local IRB approval was received through Gallup Romania, which had permission from the Romanian Government to conduct human subject research. Gallup Romania shared the de-identified data with Andreea Mitrut, the project PI, who received IRB approval from the University of Gothenburg.

* Northwestern University, † Vanderbilt University, + University of Gothenburg, ± Columbia University.

1. Introduction

Teacher effort is a scarce resource, yet little is known about how it is allocated across students who differ in ability and behavior. Teachers face the task of allocating their effort across heterogeneous students, often balancing instructional needs with the demands of classroom management. These decisions affect not only individual student learning but also the design and effectiveness of key education policies, such as teacher performance pay (Glewwe et al., 2010; Lavy, 2009), school accountability systems (Neal & Schanzenbach, 2010; Reback, 2008), and the adoption of new curricula or technology (Arias Ortiz et al., 2024). At the same time, the demands of teacher effort in the classroom can affect working conditions, influencing teacher turnover and attrition, especially if teaching some students (such as more disruptive students) requires disproportionately more effort (Hanushek et al., 2004; Lankford et al., 2002).

Although teacher effort is a key input in education production, its distribution across individual students is largely unobserved. This is partly due to the difficulty of measuring classroom interactions directly (Ahn, 2013). As a result, existing studies typically infer teacher effort from aggregate outcomes or policy responses, and rely on proxies such as absenteeism or instructional time (e.g., Duflo et al., 2012; Lavy, 2009). Even detailed classroom observations—such as those based on the CLASS framework (Pianta et al., 2008) which evaluates the quality of teacher-student interactions, and the Stallings (1986) protocols that record time spent on different tasks—miss important but difficult-to-measure activities, such as providing individualized attention or monitoring students. As a result, these measures also leave open the question of how teachers allocate their effort to students within and across classrooms, and what student-level characteristics systematically shape this allocation.

Our paper addresses this measurement challenge by studying how teachers target their scarce effort across individual students, focusing on the roles of student ability and disruptive behavior. These characteristics are the ones most closely associated with the two domains commonly used to frame teacher effort in the education literature (Carroll, 1963; Kounin, 1970): (i) providing instruction/delivering academic content, and (ii) managing the classroom environment and addressing disruptions. In economics, Lazear (2001) formalizes this tradeoff by modeling disruptive behavior as a key constraint on instructional output. Accordingly, teacher effort in providing instruction likely depends on students' academic ability, while teacher effort in managing the classroom likely depends on students' disruptiveness. Furthermore, since teacher

effort may be influenced by student gender and socioeconomic background, we also explore the role of these factors in our analyses.

We examine relative teacher effort using unique survey data we collected in 87 Romanian high schools that tracked students into classes based on a standardized admission score. Our data collection focused on teachers who taught multiple tracked classes. This allows us to study the allocation of teacher effort both within and between classes that differ in overall ability. In each class, we selected five students from the top, middle, and bottom of the class based on their admission scores. Teachers who taught these classes received the names of these ten students (five from each class) and were asked to rank them by how much time and effort they devote to each student. This measure captures the relative allocation of teacher effort in a broad sense, including individualized instruction, monitoring student progress, answering questions, and classroom management.

We also asked teachers to compare the overall time and effort they devote across their different classes. In addition, teachers ranked students (and classes) on other dimensions such as perceived academic ability and disruptiveness. Finally, we asked each of the ten sampled students to rank the same set of ten students, including themselves, along the same dimensions of academic ability and disruptiveness. To the best of our knowledge, this is the first study to directly measure relative teacher effort across students with different characteristics, such as ability and disruptiveness, both within and across classes, providing direct evidence on how teachers report allocating effort across students rather than indirect inference from aggregate outcomes or time-use proxies.

Our analysis yields four key findings about how teachers allocate effort and experience their classroom environment. First, teachers report allocating more effort towards disruptive students, consistent with the demands of classroom management. Second, teachers report targeting greater effort towards lower-ability students, consistent with a compensatory allocation of teacher effort. Third, while girls appear to receive less effort than boys, this difference disappears, or even reverses, when conditioning on ability and/or disruptiveness. We do not observe any relationship between teacher effort and students' socio-economic status. Fourth, teachers report lower enjoyment when their classrooms have more disruptive or lower-ability students.

These findings emerge from our baseline OLS estimates which compare students within the same classroom taught by the same teacher. We show that these results are essentially

unchanged when controlling for predetermined student and family characteristics such as gender, ethnicity, number of siblings, and mother's education. Furthermore, we use multiple reports of the same student to assess the robustness of our results to two potential sources of bias. The first is reverse causality: teacher effort may influence student disruptiveness or ability rather than the other way around. The second is reporting bias: teachers may give more favorable assessments to students they prefer, affecting reported effort, ability, and disruptiveness. To mitigate these concerns, we construct an instrumental variable using rankings of the same student's academic ability and disruptiveness, reported by other teachers who taught the student in different subjects at the same school.

The validity of our instrumental variable approach requires that other teachers' reports affect teacher effort only through the instrumented student characteristic. This assumption could be violated if teachers' assessments reflect common information about students beyond the instrumented characteristics, or if teachers share information about students' performance and behavior. We conduct several exercises to provide evidence about the plausibility of this exclusion restriction. First, we show that our results hold when restricting the analysis to teacher pairs who teach very different subjects (e.g., sports and math) and observe students in different instructional contexts such that their assessments may be less likely to reflect common information compared to teachers in closely related subjects (e.g., math and physics). Second, we show similar results when we use rankings of the same students by their peers, rather than by teachers. Third, our findings are robust when we use students' predetermined high-school admission scores as an alternative instrument for reports of academic ability. Fourth, since in most cases we have more than two teacher reports for the same student, we implement an overidentification test to assess the validity of our instrumental variables. While these exercises are not necessarily conclusive about the validity of the exclusion restriction, the consistent pattern between our IV and OLS estimates, together with these robustness checks, provides some reassurance that the patterns we document are not entirely driven by reverse causality or teacher-specific reporting bias.

We interpret these findings in the Romanian context, where—as in many other countries—no explicit incentives reward teacher effort. Despite this, teachers report greater effort in more disruptive and lower-ability classrooms. In a complementary class-level analysis, we find that teachers also report less enjoyment in more disruptive and lower-ability classrooms, suggesting that the greater effort associated with these environments may come at a cost to teachers'

enjoyment of teaching. We explore the potential implications of our findings to the policy questions of retaining and rewarding teachers in the conclusion.

This study intersects with several strands of the literature related to teacher effort and the economics of education more broadly. First, as discussed, we contribute to measuring teacher effort by introducing a novel approach to assessing how teachers allocate effort to specific students. In contrast, the existing research has typically relied on proxies of teacher effort at the teacher or class level. These include measures of teacher absence (Duflo et al., 2012; Gershenson, 2016; Leaver et al., 2021; Miller et al., 2008) or self-reported hours of instruction and overtime (Gibney et al., 2023; Jones, 2013; Lavy, 2009). While informative, these measures do not capture how teachers allocate effort across individual students.¹

Second, our finding that teachers allocate effort towards students with lower academic ability relates to the prior literature examining how certain incentives can lead teachers to focus on students with different levels of academic ability (Duflo et al., 2011; Goulas et al., 2025; Salvati, 2023; Zhang et al., 2025). Most of this research infers teacher effort from how school policies shift the distribution of academic achievement gains across the baseline distribution. For example, the gains induced by accountability often concentrate near thresholds associated with sanctions for low performance, suggesting that instructional effort shifts away from higher-achieving and lower-achieving students (Reback, 2008; Neal & Schanzenbach, 2010; Jennings & Sohn, 2014).² While this literature argues that teacher effort responds to incentives and can either mitigate or amplify initial differences, direct evidence on teacher effort at the student level remains limited.

Third, we add to the growing literature examining the impact of disruptive students in the classroom. Several empirical studies find that exposure to disruptive peers reduces classmates' academic achievement (Carneiro et al., 2025; Carrell & Hoekstra, 2010; Goulas et al., 2024; Zhao & Zhao, 2021). A likely mechanism for this result is that disruption diverts a teacher's attention from instruction to classroom management. These studies, along with the models by Carroll (1963) and Lazear (2001) mentioned earlier, highlight the importance of classroom management but do not show that teacher effort responds directly to disruptive students as documented in our study.

¹ A couple of intriguing studies do examine teacher-student interactions captured through classroom observations and teacher-student exchanges during virtual tutoring sessions (Allen et al., 2013; Zhang et al., 2025).

² More recently, evidence from eye-tracking and interaction logs suggests that attention is often directed toward lower-performing students (e.g., Chaudhuri et al., 2022; Zhang et al., 2025), while vignette-based experimental evidence suggests attention may shift away from low-achievers when a standout high achiever is present (Goulas et al., 2025).

Fourth, we also build on the literature examining how teacher effort and attention vary by gender. Observational studies show that boys receive more interaction and disciplinary attention (Bassi et al., 2018; Consuegra et al., 2018; Jones & Dindia, 2004). However, recent evidence from lab experiments using eye-tracking software indicates that both male and female teachers are more likely to allocate their attention towards girls (Gabel et al., 2025). Furthermore, there is an existing literature comparing blind and non-blind assessments, often finding that teachers favor girls over boys with similar externally measured achievement (Cornwell et al., 2013; Lavy, 2008). Our findings contribute to this research by analyzing how student gender affects teacher effort within classrooms.³

The paper is organized as follows. Section 2 provides background on the Romanian educational system. Section 3 describes our data collection and the relevant samples. Section 4 explains the empirical strategy. Section 5 presents the results, and Section 6 concludes.

2. Background

In Romania, students preparing to enter high school (grades 9–12) receive an *admission score*, calculated as a weighted average of their middle school GPA (grades 5–8) and their performance on a national standardized 8th-grade examination. After learning their score, students submit a ranked list of preferred specializations within high schools. A specialization is a distinct curricular track within a high school (e.g., Mathematics, Natural Sciences, Technical Studies, Social Studies, Literature). Students are assigned to specializations through a centralized allocation system that uses a serial dictatorship algorithm to match them to their top choices based on their admission scores.

Before students submit their ranked preferences for high schools, the Ministry of Education publicizes the number of available slots for each specialization in every high school. These slots come in multiples of 28, the maximum class size, so most specializations typically offer one or two self-contained classes of 28 students each. Furthermore, many schools (and all schools in our sample) track students within a specialization based on their admission scores. Thus, the top 28

³ While we do not find a significant relationship between teacher effort and student SES, research suggests that teachers' subjective judgments can vary by SES. Evidence from randomized student-profiles indicates a negative SES gradient in teachers' beliefs about students' success prospects (Brunello et al., 2025), and studies of teacher assessments and track recommendations show that lower-SES students tend to receive less favorable assessments/recommendations than more advantaged peers (Feron et al., 2016; Zumbuehl et al., 2025).

students are placed in one class, while the next group of 28 students is assigned to another. Such tracking results in classes that are well-defined in terms of ability as measured by their admission score and have a stable composition over time.

All students within a given specialization follow the same curriculum, but instruction is delivered by teachers in separate classes for each subject. Thus, students in different classes within the same school and specialization do not attend academic lessons together, though they may interact during breaks or extracurricular activities. Teacher assignments to classes are determined by the school principal, and many teachers have multiple classes within the same specialization. These class assignments generally remain fixed throughout all four years of high school. While teachers may exchange information about students, particularly with colleagues teaching similar subjects, the scope for shared subject-specific information may be more limited across teachers in very different subjects.⁴

3. Data

Our analysis draws on an unusually rich set of linked administrative and survey data. First, we use administrative data from the Romanian Ministry of Education covering four student cohorts. These data include each student's admission score and assigned school specialization. We merge this information with school-level records on class-level assignments. This allows us to identify the full list of student names and their respective admission scores for each class within a given school, grade, and specialization. We also have the names of all teachers in these schools and the subjects they taught to each class.

Second, we collected unique teacher and student surveys in the spring of 2017 across 87 high schools in 74 towns. We selected these schools because they track students into at least two classes based on their admission scores. Within each pair of tracked classes, we selected five students from each class to span the admission-score distribution: the highest-scoring student, the lowest-scoring student, the median student, and the two additional students nearest to the cutoff between the top and bottom classes.⁵ Specifically, in such tracked specializations with two classes

⁴ Prior research in education suggests that secondary school teachers operate within distinct subject-specific contexts that shape interaction patterns, consistent with collaboration being structured along disciplinary boundaries (Grossman & Stodolsky, 1995).

⁵ In a separate paper using these data (Malamud et al., 2025), we leverage the oversampling of students near the cutoff between classes to estimate the effect of tracking on student perceptions.

of 28 students each (i.e., 56 students in total), we focused on the five students ranked 1, 14, 26, 27, 28 in the top class and the five students with equivalent ranks in the bottom class: 29, 30, 31, 43, 56. We surveyed all teachers who taught the same subject to these tracked classes and were therefore familiar with students in both classes. We also surveyed the targeted students. Our final sample includes 1,843 teachers and 3,530 students across 435 classroom pairs.

Our teacher survey provided each teacher with a list of the 10 students ordered alphabetically. We construct our main dependent variable from a question asking each teacher to rank these students according to the relative time and effort they devoted to each of them (e.g., preparing materials, monitoring progress, and answering questions).⁶ Teachers were also asked to rank the same 10 students along several other dimensions that serve as our main independent variables: academic ability (e.g., ease of understanding difficult concepts, high innate aptitude), disruptiveness (e.g., harassing or disparaging classmates), student effort (e.g., completing homework, attentiveness in class), and expected performance (e.g., anticipated success on the Baccalaureate exam taken at the end of high school). Thus, our main measure captures the relative allocation of teacher effort across sampled students, rather than absolute differences in time or effort.

Our main analysis compares the five students within each class, normalizing the resulting ranks at the classroom level on a 1-to-5 scale. However, our teacher survey also elicited comparisons between the two tracked classes (as a whole) along the same dimensions used in ranking students: i.e., student ability, disruptiveness, effort, and expected performance. In addition, we asked teachers to compare the two tracked classes in terms of how much they enjoy teaching the class, the difficulty of the material covered, and the effort devoted to the class. These variables capture teacher comparisons at the class-level and supplement the individual-level data collected from the list of selected students.

A complementary student survey asked the same students, in the same format as for teachers, to rank all ten students—including themselves—along the same dimensions presented to teachers. We use these peer reports as alternative measures of student characteristics in our

⁶ The exact phrasing was: “Please take into account the time and effort you devote to your students. For example, you can think about preparing homework or additional individualized materials for certain students, individually monitoring some students to gauge their progress, the time allocated to answering extra clarifications/questions about the material taught, the time spent resolving problematic situations created by certain more disruptive students, and so on. Please rank the above 10 students in terms of how much time and effort you devote to each.”

robustness analysis. Students also completed a background questionnaire that included demographic and family characteristics (e.g., gender, number of siblings, ethnicity, parental education). Teachers likewise completed a questionnaire on their demographic background and professional experience, including years of teaching and whether they held a tenured position. Appendix Table A1 reports summary statistics for all main student- and teacher-level variables used in the analysis.

4. Empirical strategy

We examine how teachers report allocating relative effort across students with different levels of academic ability and disruptive behavior using the following estimation equation:

$$E_{ict} = \beta_0 + \beta_1 A_{ict} + \beta_2 D_{ict} + \gamma X_{ict} + \lambda_{ct} + \varepsilon_{ict} \quad (1)$$

where:

- E_{ict} is teacher *effort* toward student i in class c taught by teacher t
- A_{ict} is the academic *ability* of student i in class c taught by teacher t
- D_{ict} is the *disruptiveness* of student i in class c taught by teacher t
- X_{ict} is a vector of student characteristics—in some robustness specifications, we include a set of control variables including student gender, ethnicity, and mother’s education.
- λ_{ct} : is a class-by-teacher fixed effects that absorbs unobserved factors within each teacher-classroom combination, such as teacher style, overall classroom culture, and class-specific practices.
- ε_{ict} is an error term. Standard errors are clustered at the class-by-teacher level.⁷

The estimates exploit within-classroom variation across students, holding fixed the teacher-classroom combination. Three main concerns arise when interpreting the OLS estimates from equation (1). First is reverse causality since the effort teachers report allocating may affect disruptiveness or perceived ability, rather than the other way around. Second, reporting biases may arise because the same focal teacher reports the baseline measures of effort, ability, and disruptiveness; for example, a teacher who favors a certain student might report allocating greater

⁷ We also verify that our main findings are robust to clustering at the teacher level only, as well as clustering by class-by-teacher and student.

effort to that student and rate the student higher in disruptiveness. Third, because ability and disruptiveness may be correlated with other unobserved student characteristics also associated with teacher effort, one might worry about omitted-variable bias.

To assess the robustness of our results to reverse causality and reporting bias, we implement an instrumental variables (IV) approach using other teachers' reports of the same student's disruptiveness and ability as instruments. The key identifying assumption is that, conditional on classroom-by-teacher fixed effects and observed covariates, the non-focal teachers' reports reflect the student's underlying traits, unrelated to the focal teacher's biases or effort decisions. Because different teachers observe the same student in distinct instructional contexts, their assessments provide alternative measures of ability and disruptiveness that are less directly tied to the focal teacher's own reporting behavior and effort decisions.

For this IV strategy to be valid, the non-focal teachers' ratings must be strongly correlated with the focal teacher's ratings of disruptiveness and ability. It must also affect the focal teacher's effort only through this channel. This assumption may not hold if teacher assessments reflect common information about students beyond the characteristics being instrumented, or if teachers share information about students' performance or behavior.

We conduct several exercises to provide evidence for the plausibility of this assumption. First, we show that our results hold when restricting the analysis to reports from non-focal teachers who teach very different subjects (e.g., sports and math). These teachers observe students in different instructional contexts, and their assessments may therefore be less likely to reflect common subject-specific information than those of teachers in closely related subjects (e.g., math and physics).⁸ Second, we use rankings of the same students by their peers, rather than by teachers, as alternative measures of academic ability and disruptiveness. We consider reports from peers in both the same class and the other tracked class, thereby limiting the possibility of within-class spillovers. Third, for academic ability, we also use students' predetermined high-school admission scores as an alternative measure that does not rely on contemporaneous teacher or peer assessments. Fourth, since we have multiple teacher and peer reports per student, we can run an overidentification test to assess the validity of our exclusion restrictions. As mentioned earlier, these exercises are not necessarily conclusive about the validity of the exclusion restriction. However, the consistent pattern between our IV and OLS estimates, together with these robustness

⁸ For this purpose, we categorize subjects into STEM, humanities and languages, sports, and others.

checks, provides some reassurance that the patterns we document are not driven solely by teacher-specific reporting bias or by the particular measures of student characteristics used in our baseline specifications.

We consider another potential concern—omitted-variable bias arising from ability or disruptiveness being correlated with other characteristics—by looking at specifications that condition on observed student characteristics that may be correlated with ability, disruptiveness, and effort teachers report allocating. We first control for a set of predetermined characteristics, such as student gender, ethnicity, number of siblings, and mother’s education. We also explore the impact of a broader set of controls that include self-esteem (measured by the Rosenberg index), whether students received tutoring, and exposure to bullying. While addressing omitted variable bias requires that unobserved characteristics are also uncorrelated with our regressors, generating credible exogenous variation in student ability and behavior would be challenging outside of an experimental setting.

Finally, we run analogous specifications at the *classroom level* to complement the student-level analysis and leverage additional measures available in the teacher survey. In particular, we first aggregate the individual-level data to the classroom level and compare the two classes taught by the same teacher. In these specifications, we include teacher fixed effects for the two tracked classes taught by the same teacher from which the 10 targeted students were selected. Second, the teacher survey also includes direct comparisons of the overall effort teachers report allocating to each class, rather than rankings of individual students. These measures provide a complementary assessment of effort allocation across the two tracked classrooms that do not rely on aggregating student-level rankings. The IV strategy is applied similarly at the classroom level, using non-focal teacher reports of differences in disruptiveness and academic ability across classrooms.

5. Results

OLS estimates

Table 1 examines how teacher effort relates to students’ academic ability and disruptive behavior. It presents OLS estimates with classroom-by-teacher fixed effects, exploiting variation across students within the same classroom taught by the same teacher. Columns (1) and (2) show two

clear patterns: teachers report allocating more effort to lower-ability students and to more disruptive students. The magnitudes of these associations are similar but of opposite sign: a one standard deviation increase in academic ability (in column 1) and disruptiveness (in column 2) are associated with a decrease of 0.23 and an increase of 0.26 standard deviation units in the outcomes, respectively.⁹ Including ability and disruptiveness jointly (column 3) attenuates the results, but the magnitudes remain sizable and statistically significant. Column (4) shows that these relationships are robust to conditioning on predetermined student and family characteristics; adding controls for student gender, ethnicity, number of siblings, and mother's education level essentially does not change the coefficients on academic ability and disruptive behavior.¹⁰ We also include a control variable for the alphabetical ordering of student names to account for the order in which these names were displayed on the surveys.

Table 1: Teacher Effort, Student Academic Ability, and Student Disruption

	Dependent variable: Teacher effort			
	(1)	(2)	(3)	(4)
Academic Ability	-0.227*** [0.015]		-0.167*** [0.015]	-0.178*** [0.018]
Disruptive behavior		0.257*** [0.014]	0.210*** [0.014]	0.210*** [0.018]
Includes control variables	No	No	No	Yes
N	11,814	11,814	11,814	8,684
R-squared	0.061	0.076	0.101	0.176
Mean of DV	2.976	2.976	2.976	2.948

DV: Dependent variable; N: Number of observations

Notes: Teacher effort is measured using teachers' within-class rankings of the time and effort they devote to each sampled student, and ranges from 1 (lowest relative effort) to 5 (highest relative effort). Academic ability and disruptive behavior are teachers' within-class rankings of each sampled student's ability and disruptiveness, respectively, and range from 1 (lowest) to 5 (highest). Each column represents an independent regression. All regressions include classroom-by-teacher fixed effects. Column (4) additionally controls for the student's gender, ethnicity, number of siblings, mother's education, and alphabetical order of names. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

⁹ To derive these magnitudes, we rescale the coefficients by the standard deviations of these variables in Appendix Table A1. Because teacher effort is measured as relative ranking, these standardized magnitudes describe differences in the relative allocation of effort across students and do not correspond to changes in absolute time or effort.

¹⁰ The inclusion of these control variables reduces the estimation sample because of missing values for some student characteristics. However, re-estimating the specifications in columns (1)–(3) on the smaller sample used in column (4) yields almost identical results. Results are also similar when we include a broader set of control variables which may not be exogenous (e.g., tutoring, self-esteem, and exposure to bullying), as shown in Appendix Table A2.

IV Estimates

Our baseline OLS estimates rely on measures of teacher effort, academic ability, and disruptive behavior reported by the same teacher, raising concerns about reporting bias and reverse causality as discussed in Section 4. **Table 2** assesses the robustness of these relationships using alternative sources of information on student ability and disruptiveness. In Panel A, we instrument the focal teacher's assessments with reports of the same students by other teachers in the school. The same two patterns documented in Table 1 emerge in columns (1) and (2): reported teacher effort is negatively associated with academic ability and positively associated with disruptive behavior. Both relationships remain sizable and statistically significant when we include them jointly in column (3), with the IV estimates larger in magnitude than their OLS counterparts, particularly for academic ability.

The negative relationship between academic ability and reported teacher effort is remarkably stable across alternative instruments. In particular, it persists when we use reports from teachers in unrelated subjects (Panel B), as well as when we use peer reports from the same or adjacent tracked class (Panels C and D) and when we instrument for academic ability using the high-school admission score, which is determined prior to the teacher-student interactions in our setting (Panel E). The positive relationship between teacher effort and disruptive student behavior is also robust across these alternative instruments, except when using peer reports from the adjacent class and including academic ability jointly (Panel D). Finally, the estimates remain similar when we construct separate instruments from two groups of non-focal teachers, and the corresponding overidentification tests do not reject instrument validity (Panel F).¹¹

Taken together, these results indicate that the main patterns are robust to using different sources of information about student characteristics. The negative relationship with academic ability is consistent with a compensatory allocation of teacher effort toward students with greater academic needs, while the positive relationship with disruptiveness is consistent with classroom management demands.

¹¹ The corresponding first-stage estimates are reported in Appendix Table A3 and are strong across specifications. Additionally, our results are very similar to including the same set of observed student and family characteristics included in column (4) of Table 1.

Table 2: Teacher Effort and Alternative Measures of Student Characteristics, IV Estimates

	Dependent variable: Teacher effort		
	(1)	(2)	(3)
Panel A. Preferred IV: Reports from other teachers (N = 11,214)			
Academic Ability	-0.435*** [0.038]		-0.280*** [0.055]
Disruptive behavior		0.389*** [0.034]	0.220*** [0.050]
Panel B. Alternative IV 1: Teacher unrelated subjects (N = 10,717)			
Academic Ability	-0.421*** [0.042]		-0.258*** [0.061]
Disruptive behavior		0.385*** [0.036]	0.235*** [0.053]
Panel C. Alternative IV 2: Peer reports from same class (N = 10,494)			
Academic Ability	-0.412*** [0.040]		-0.293*** [0.060]
Disruptive behavior		0.390*** [0.043]	0.199*** [0.066]
Panel D. Alternative IV 3: Peer reports from the other class (N = 9,508)			
Academic Ability	-0.509*** [0.057]		-0.553*** [0.142]
Disruptive behavior		0.347*** [0.064]	-0.063 [0.152]
Panel E. Alternative IV 4: Admission score (N = 11,249)			
Academic Ability	-0.505*** [0.044]		-0.409*** [0.069]
Disruptive behavior		0.390*** [0.034]	0.145** [0.059]
Panel F. Alternative IV 5: Over-identified (two groups of non-focal teachers; N = 9,333)			
Academic Ability	-0.428*** [0.037]		-0.283*** [0.051]
Disruptive behavior		0.385*** [0.035]	0.216*** [0.049]
Hansen J p-value	0.323	0.959	0.740

IV: Instrumental variable; N: Number of observations

Notes: Teacher effort is measured using teachers' within-class rankings of how much time and effort they devote to each sampled student, and ranges from 1 (lowest relative effort) to 5 (highest relative effort). Academic ability and disruptive behavior are teachers' within-class rankings of each sampled student's academic ability and disruptive behavior, respectively, and range from 1 (lowest) to 5 (highest). Each column within panels represents an independent regression. All regressions include classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Finally, Appendix Table A4 confirms that our main results are robust to alternative specifications that explicitly account for the ordinal nature of the teacher effort measure. These include collapsing our rankings of teacher effort into binary indicators, estimating Plackett-Luce rank-ordered logit models (Luce, 1959; Plackett, 1975) and using Bradley-Terry models (Bradley & Terry, 1952), which break ranks into pairwise comparisons, along with their corresponding IV specifications (see Appendix Section 1 for more details).

The Role of Gender and Socioeconomic Status (SES)

Table 3 examines how reported teacher effort varies by student gender and SES, with the latter proxied by whether a student’s mother has completed college.¹² Panel A shows a sizable gender difference: teachers report allocating less effort to girls than to boys (column 1). However, this gap narrows by more than 30 percent when we condition on students’ academic ability (column 2) and disappears when we condition on disruptive behavior (columns 3 and 4). In the IV specifications that condition on disruptive behavior (columns 6 and 7), the pattern reverses entirely, suggesting that—controlling for academic ability and behavior—teachers actually report devoting more effort to girls than to boys. Taken together, this suggests that the baseline gender differences in reported teacher effort largely reflect differences in perceived disruptiveness between boys and girls, rather than a simple gender gradient in teacher effort. Consistent with this interpretation, Appendix Table A5 shows that girls are rated as significantly less disruptive and slightly more academically able than boys.

In contrast, Panel B of Table 3 provides little evidence of significant differences in reported teacher effort by SES across our various specifications. Thus, within these tracked classrooms, socioeconomic background appears to play a more limited role in the allocation of teacher effort.

¹² Results are robust to using different proxies of SES based on parental education.

Table 3: Student Gender, Socioeconomic Status, and Teacher Effort

	Dependent variable: Teacher effort						
	OLS				IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Student's gender							
Student is female	-0.229*** [0.051]	-0.155*** [0.049]	0.034 [0.051]	0.040 [0.050]	-0.096** [0.045]	0.234*** [0.068]	0.140* [0.080]
Academic ability		-0.233*** [0.019]		-0.176*** [0.018]	-0.419*** [0.048]		-0.231*** [0.078]
Disruptive behavior			0.264*** [0.018]	0.214*** [0.018]		0.465*** [0.054]	0.297*** [0.086]
N	8,245	8,245	8,245	8,245	8,245	8,245	8,245
R-squared	0.069	0.119	0.128	0.155	0.025	0.032	0.085
Mean of DV	2.948	2.948	2.948	2.948	2.948	2.948	2.948
KP F statistic	-	-	-	-	486.4	392.0	45.1
Panel B: Student's socioeconomic status							
Mom attended college	-0.046 [0.060]	-0.008 [0.058]	-0.037 [0.058]	-0.010 [0.057]	0.024 [0.052]	-0.032 [0.050]	0.004 [0.051]
Academic Ability		-0.238*** [0.019]		-0.176*** [0.018]	-0.435*** [0.047]		-0.259*** [0.070]
Disruptive behavior			0.261*** [0.017]	0.210*** [0.017]		0.381*** [0.038]	0.231*** [0.059]
N	8,245	8,245	8,245	8,245	8,245	8,245	8,245
R-squared	0.064	0.117	0.128	0.155	0.018	0.054	0.088
Mean of DV	2.948	2.948	2.948	2.948	2.948	2.948	2.948
KP F statistic	-	-	-	-	519.7	787.3	73.3

N: Number of observations; DV: Dependent variable; KP: Kleibergen-Paap rank Wald

Notes: Teacher effort is measured using teachers' within-class rankings of the time and effort they devote to each sampled student, and ranges from 1 (lowest relative effort) to 5 (highest relative effort). Academic ability and disruptive behavior are teachers' within-class rankings of each sampled student's academic ability and disruptive behavior, respectively, and range from 1 (lowest) to 5 (highest). Each column within panels represents an independent regression. All regressions include classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Class-level analysis

In addition to the individual-level data, in which teachers reported their effort toward specific students, our survey also asked teachers to compare their overall effort across different classes within the same subject. We use these data to assess whether the patterns at the individual level also extend to comparisons across classrooms taught by the same teacher. **Table 4** shows similar class-level patterns in both OLS and IV specifications when we aggregate individual-level measures of teacher effort to the class level (Panel A) and when we use class-level measures of effort based on direct comparisons of classes (Panel B).¹³ In short, Table 4 confirms that our individual-level results on teacher effort hold when teachers compare across classes.

The class-level survey also enables us to examine other dimensions of teachers' classroom experience and instruction. **Table 5** shows how classroom characteristics relate to teachers' classroom experience and instruction. Teachers report greater enjoyment in classes with higher-ability students and less disruptive behavior (column 1). They also report being able to cover more difficult material and having harder-working students in higher-ability, less-disruptive classrooms (columns 2 and 3). Higher classroom ability is further associated with more high-achieving students and higher teacher expectations (columns 4 and 5). The final two columns suggest that classroom composition is related to instructional targeting: higher ability is associated with greater targeting of the top third of students, while greater disruption is associated with greater targeting of the bottom third.

Finally, Appendix Table A6 examines heterogeneity of our main estimates across several student, teacher, and regional characteristics. We do not observe significant differences in how teachers respond to higher-ability or more disruptive students by student gender, student SES, school quality, teacher experience, or whether the teacher is teaching in a higher-achieving (top) class or a relevant subject. Apart from the marginal significance for teacher gender, the only other dimension with significant differences in teacher effort is urban status, suggesting that teachers in urban schools may allocate effort differently than those in rural schools. Overall, these findings indicate that the compensatory and classroom management mechanisms highlighted above are

¹³ These relationships are less precisely estimated when academic ability and disruptive behavior are included jointly, particularly for academic ability.

highly stable across teacher, student, and school characteristics, with urbanicity as the only exception.

Table 4: Perceived Student Ability, Student Disruption, and Self-Reported Teacher Effort, Class-level Estimates

	Dependent variable: Aggregated teacher effort					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Aggregated measures						
Students' ability	-0.237*** [0.055]		-0.104* [0.057]	-0.383*** [0.080]		-0.124 [0.168]
Students' disruption		0.357*** [0.052]	0.310*** [0.057]		0.417*** [0.079]	0.321* [0.172]
N	1,940	1,940	1,940	1,940	1,940	1,940
R-squared	0.070	0.134	0.142	0.034	0.116	0.127
Mean of DV	5.493	5.493	5.493	5.493	5.493	5.493
KP F statistic	-	-	-	232.2	185.6	17.27
Panel B: Class-level measures						
Students' ability	Dependent variable: Class-level teacher effort					
	-0.134*** [0.045]		-0.017 [0.053]	-0.149*** [0.041]		0.123 [0.091]
Students' disruption		0.225*** [0.044]	0.215*** [0.053]		0.322*** [0.061]	0.451*** [0.130]
N	1,940	1,940	1,940	1,940	1,940	1,940
R-squared	0.018	0.051	0.051	0.018	0.041	0.011
Mean of DV	0.500	0.500	0.500	0.500	0.500	0.500
KP F statistic	-	-	-	1,981	375.4	34.22

N: Number of observations; DV: Dependent variable; KP: Kleibergen-Paap rank Wald

Notes: Aggregated teacher effort is the paired-classroom-level average of teachers' student-level effort rankings, with higher values indicating greater average effort. Class-level teacher effort is a binary indicator equal to 1 for the class to which the teacher reported devoting more total energy, and 0 for the comparison class. Each column within panels represents an independent regression. All regressions include teacher fixed effects. Robust standard errors in brackets are clustered at the teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Table 5: Perceived Student Ability, Student Disruption and Other Outcomes, Class-level Estimates

	Dependent variables						
	Enjoys class- room	Cover materials	Hard working students	High achievers	Teacher expecta- tions	Targets top third	Targets bottom third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS estimates							
Students' ability	0.394*** [0.050]	0.454*** [0.050]	0.651*** [0.045]	0.878*** [0.028]	0.469*** [0.048]	0.043** [0.019]	-0.041*** [0.014]
Students' disruption	-0.394*** [0.050]	-0.265*** [0.050]	-0.282*** [0.045]	-0.049* [0.028]	0.017 [0.048]	-0.007 [0.019]	0.037** [0.014]
N	1,940	1,940	1,940	1,940	1,940	1,940	1,940
R-squared	0.480	0.407	0.703	0.821	0.212	0.690	0.790
Mean of DV	0.500	0.500	0.500	0.500	0.500	0.0964	0.127
Panel B: IV estimates							
Students' ability	0.230*** [0.073]	0.631*** [0.076]	0.659*** [0.067]	1.020*** [0.043]	0.656*** [0.082]	0.068** [0.030]	-0.025 [0.027]
Students' disruption	-0.711*** [0.104]	-0.126 [0.107]	-0.418*** [0.093]	-0.005 [0.062]	0.072 [0.120]	-0.001 [0.043]	0.082** [0.039]
N	1,940	1,940	1,940	1,940	1,940	1,940	1,940
R-squared	0.409	0.383	0.683	0.806	0.185	0.015	0.031
Mean of DV	0.500	0.500	0.500	0.500	0.500	0.0964	0.127
KP F statistic	34.22	34.22	34.22	34.22	34.22	34.22	34.22

N: Number of observations; DV: Dependent variable; KP: Kleibergen-Paap rank Wald

Notes: The dependent variables in columns (1)–(5) are binary indicators based on teachers' direct comparisons of two classrooms they taught in the same school year. Each indicator equals 1 for the class selected by the teacher and 0 for the comparison class: the class the teacher enjoyed teaching more; the class in which the teacher reported covering more difficult material; the class with more hard-working students; the class with more high-achieving students; and the class for which the teacher reported higher expectations to reach learning objectives. Targets top third and Targets bottom third are binary indicators equal to 1 if the teacher reported aiming instruction at the top or bottom third of students in that classroom, respectively, and 0 otherwise. Regressions include teacher fixed effects. Robust standard errors in brackets.

*** p<0.01, ** p<0.05, * p<0.10.

6. Conclusion

This paper examines how teachers' classroom effort relates to student characteristics in Romanian high schools. Using linked administrative and survey data that provide novel measures of relative teacher effort, we show that teachers report allocating more effort to lower-ability and more

disruptive students, consistent with compensatory instructional effort and classroom management demands. In addition, although teachers report directing less effort toward girls, this gap is largely explained, and in some specifications reversed, by controlling for differences in ability and disruptive behavior. Overall, our results are consistent with teacher effort responding to instructional and managerial needs in the classroom rather than to student gender.

At the same time, teachers report lower enjoyment when teaching more disruptive and lower-ability classrooms, highlighting the potential tradeoff between student outcomes and teacher welfare. These patterns may have important implications for teacher retention and compensation. Even without strong incentives for teacher effort (as holds in Romania), teachers may expend more effort on more disruptive and lower-achieving students if these students require greater instructional or classroom-management work. For example, if instruction is harder when students are disruptive, teachers may devote more effort to addressing these students' issues, in line with Lazear's (2001) model of student disruptiveness. However, because teachers report lower enjoyment in classes with disruptive and low-achieving students, sustaining this higher effort in the long run may require compensating mechanisms, such as intrinsic motivation, wage differentials, or institutional support (e.g., smaller class sizes and co-teaching arrangements). Absent such mechanisms, higher effort in more demanding classrooms may come at the cost of teacher satisfaction and retention, potentially exacerbating educational inequality. Whether these differences in reported effort and enjoyment ultimately translate into teacher retention or other labor-supply responses remains an important question for future research.

Our findings also suggest that variation in teacher effort across classroom environments should play a more central role in debates about teacher effectiveness. In recent years, value-added measures have become popular in research and policy as tools to evaluate teachers. The canonical value-added literature—e.g. Rockoff (2004), Rivkin et al. (2005), Chetty et al. (2014)—typically treats value-added as a fixed teacher attribute that is constant across students and classrooms. Yet, more recent work (Ahn et al., 2025; Biasi et al., 2021; Delgado, 2023) documents teacher–student match effects, highlighting heterogeneity in teacher effectiveness across students and classrooms. However, even this work focuses on student outcomes rather than on the input side of the education production process. Our results point to an important behavioral response: teachers may adjust their effort in response to student characteristics and classroom composition. A more complete

understanding of teacher effectiveness should therefore consider both inputs (e.g., effort) and outputs (e.g., student achievement), as well as the tradeoffs inherent in teaching different students.

References

- Ahn, T. (2013). The Missing Link: Estimating the Impact of Incentives on Teacher Effort and Instructional Effectiveness Using Teacher Accountability Legislation Data. *Journal of Human Capital*, 7(3), 230-273. doi:<https://doi.org/10.1086/673872>
- Ahn, T., Aucejo, E. M., & James, J. (2025). The Importance of Student-Teacher Matching: A Multidimensional Value-Added Approach. *The Review of Economics and Statistics*, 1-45. doi: <https://doi.org/10.1162/REST.a.1629>
- Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. *School Psychology Review*, 42(1), 76-98. doi:<https://doi.org/10.1080/02796015.2013.12087492>
- Arias Ortiz, E., Cristia, J. P., Malamud, O., Méndez, C., & Messina, J. (2024). Promoting Technology Adoption in Schools: Preliminary Evidence from Peru. doi:<https://doi.org/10.18235/0013123>
- Bassi, M., Mateo Díaz, M., Blumberg, R. L., & Reynoso, A. (2018). Failing to notice? Uneven teachers' attention to boys and girls in the classroom. *IZA journal of labor economics*, 7(1), 1-22. doi:<https://doi.org/10.1186/s40172-018-0069-4>
- Biasi, B., Fu, C., & Stromme, J. (2021). Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage. *NBER Working Paper* 28530. doi:<https://doi.org/10.3386/w28530>
- Bradley, R. A., & Terry, M. E. (1952). Rank Analysis of Incomplete Block Designs: The Method of Paired Comparisons. *Biometrika*, 39(3/4), 324-345. doi:<https://doi.org/10.2307/2334029>
- Brunello, G., Crocè, C., Giustinelli, P., & Rocco, L. (2025). Teacher personality and the perceived socioeconomic gap in student outcomes. *Economics Letters*, 247, 112096. doi:<https://doi.org/10.1016/j.econlet.2024.112096>
- Carneiro, P., Cruz-Aguayo, Y., Salvati, F., & Schady, N. (2025). Disruption in the Classroom: Experimental evidence from Ecuador. *Working paper Inter-American Development Bank*, N° IDB-WP-1688. doi:<http://dx.doi.org/10.18235/0013443>
- Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, 2(1), 211-228. doi:<https://doi.org/10.1257/app.2.1.211>
- Carroll, J. B. (1963). A Model of School Learning. *Teachers College Record*, 64(8), 1-9. doi:<https://doi.org/10.1177/016146816306400801>
- Chaudhuri, S., Muhonen, H., Pakarinen, E., & Lerkkanen, M.-K. (2022). Teachers' visual focus of attention in relation to students' basic academic skills and teachers' individual support for students: An eye-tracking study. *Learning and Individual Differences*, 98, 102179. doi:<https://doi.org/10.1016/j.lindif.2022.102179>
- Chetty, R., Friedman, J. N., & Rockoff, J. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9), 2593–2632. doi:<http://doi.org/10.1257/aer.104.9.2593>
- Consuegra, E., Halimi, M., & Engels, N. (2018). Gendered teacher-student classroom interactions, student sense of equity, and student achievement. *Tijdschrift voor Genderstudies*, 21(2), 131-152. doi:<https://doi.org/10.1080/13540602.2016.1158958>

- Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, 48(1), 236-264. doi:<https://doi.org/10.3368/jhr.48.1.236>
- Delgado, W. (2023). Disparate Teacher Effects, Comparative Advantage, and Match Quality. *Annenberg Institute for School Reform at Brown University, EdWorkingPaper No. 23-848*.
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5), 1739-1774. doi:<https://doi.org/10.1257/aer.101.5.1739>
- Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4), 1241-1278. doi:<https://doi.org/10.1257/aer.102.4.1241>
- Feron, E., Schils, T., & Weel, B. t. (2016). Does the teacher beat the test? The value of the teacher's assessment in predicting student ability. *De Economist*, 164(4), 391-418. doi:<https://doi.org/10.1007/s10645-016-9278-z>
- Gabel, S., Alijagic, A., Keskin, Ö., & Gegenfurtner, A. (2025). Teacher gaze and attitudes toward student gender: evidence from eye tracking and implicit association tests. *Social Psychology of Education*, 28(1), 72. doi:<https://doi.org/10.1007/s11218-025-10036-6>
- Gershenson, S. (2016). Linking Teacher Quality, Student Attendance, and Student Achievement. *Education Finance and Policy*, 11(2), 125-149. doi:https://doi.org/10.1162/EDFP_a_00180
- Gibney, V. H., West, K. L., & Gershenson, S. (2023). Blurred Boundaries: A Day in the Life of a Teacher. In D. S. Hamermesh & S. W. Polachek (Eds.), *Time Use in Economics* (Vol. 51, pp. 0): Emerald Publishing Limited.
- Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher Incentives. *American Economic Journal: Applied Economics*, 2(3), 205-227. doi:<https://doi.org/10.1257/app.2.3.205>
- Goulas, S., Griselda, S., Megalokonomou, R., & Zenou, Y. (2024). Disruptive Peers and Academic Performance: Short-and Long-Term Outcomes. *IZA Discussion Paper Series, No. 17539*.
- Goulas, S., Megalokonomou, R., & Sartori, T. (2025). *All Eyes on the Nerd? The Unequal Distribution of Teachers' Attention*. Retrieved from <https://www.econstor.eu/handle/10419/334657>
- Grossman, P. L., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(8), 5-23. doi:<https://doi.org/10.3102/0013189X024008005>
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *Journal of Human Resources*, 39(2), 326-354. doi:<https://doi.org/10.3368/jhr.39.2.326>
- Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, 87(2), 125-141. doi:<https://doi.org/10.1177/0038040714525787>
- Jones, M. D. (2013). Teacher behavior under performance pay incentives. *Economics of Education Review*, 37, 148-164. doi:<https://doi.org/10.1016/j.econedurev.2013.09.005>
- Jones, S. M., & Dindia, K. (2004). A meta-analytic perspective on sex equity in the classroom. *Review of Educational Research*, 74(4), 443-471. doi:<https://doi.org/10.3102/00346543074004443>
- Kounin, J. S. (1970). Observing and Delineating Technique of Managing Behavior in Classrooms. *Journal of Research and Development in Education*, 4(1), 62-72.

- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62. doi:<https://doi.org/10.3102/01623737024001037>
- Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 92(10-11), 2083-2105. doi:<https://doi.org/10.1016/j.jpubeco.2008.02.009>
- Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5), 1979-2011. doi:10.1257/aer.99.5.1979
- Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, 116(3), 777-803. doi:<https://doi.org/10.1162/00335530152466232>
- Leaver, C., Ozier, O., Serneels, P., & Zeitlin, A. (2021). Recruitment, Effort, and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools. *American Economic Review*, 111(7), 2213-2246. doi:<https://doi.org/10.1257/aer.20191972>
- Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- Malamud, O., Mitrut, A., Pop-Eleches, C., & Urquiola, M. (2025). Self-, Peer-, and Teacher Perceptions Under School Tracking. *The Review of Economics and Statistics*, 1-41. doi:https://doi.org/10.1162/rest_a_01555
- Miller, R. T., Murnane, R., & Willett, J. (2008). Do Teacher Absences Impact Student Achievement? *Educational Evaluation and Policy Analysis*, 30(2), 181-200. doi:<https://doi.org/10.3102/016237370708318019>
- Neal, D., & Schanzenbach, D. W. (2010). Left Behind by Design: Proficiency Counts and Test-Based Accountability. *The Review of Economics and Statistics*, 92(2), 263-283. doi:<https://doi.org/10.1162/rest.2010.12318>
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Plackett, R. L. (1975). The analysis of permutations. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 24(2), 193-202. doi:<https://doi.org/10.2307/2346567>
- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, 92(5-6), 1394-1415. doi:<https://doi.org/10.1016/j.jpubeco.2007.05.003>
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *econometrica*, 73(2), 417-458. doi:<https://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review*, 94(2), 247-252. doi:<https://doi.org/10.1257/0002828041302244>
- Salvati, A. (2023). Teacher Instruction, Classroom Composition, and Student Achievement. *Mimeo*.
- Stallings, J. A. (1986). 2. Using Time Effectively: A Self-Analytic Approach. In K. Zumwalt (Ed.), *Improving Teaching* (pp. 15-28). Alexandria, Va: Association for Supervision and Curriculum Development.
- Zhang, Q., Wang, R. E., Ribeiro, A. T., Demszky, D., & Loeb, S. (2025). Educator Attention: How computational tools can systematically identify the distribution of a key resource for students. *arXiv preprint arXiv:2502.20135*.

- Zhao, L., & Zhao, Z. (2021). Disruptive peers in the classroom and students' academic outcomes: Evidence and mechanisms. *Labour Economics*, 68, 101954.
doi:<https://doi.org/10.1016/j.labeco.2020.101954>
- Zumbuehl, M., Chehber, N., & Dillingh, R. (2025). Can skills differences explain the gap in track recommendation by socio-economic status? *Education Economics*, 33(2), 161-179.
doi:<https://doi.org/10.1080/09645292.2024.2319832>

Online Appendix for:

Teacher Effort in the Classroom

Ofer Malamud
Matias Martinez
Andreea Mitrut
Cristian Pop-Eleches
Miguel Urquiola

October 2026

Appendix Table A1: Summary Statistics	1
Appendix Table A2: OLS Estimates of Student Ability and Disruption on Teacher Effort, Including Control Variables	2
Appendix Table A3: First Stage Estimates	3
Appendix Table A4: Estimates of Student Ability, Disruption, and Teacher Effort, Alternative Specifications and Models	4
Appendix Table A5: OLS Estimates of Gender and Socioeconomic Differences in Student Ability and Disruptiveness	5
Appendix Table A6: IV Estimates of Student Ability, Disruption, and Teacher Effort, Heterogeneity Analysis	6
Appendix Section 1: Addressing the Ordinal Nature of Rankings.....	7
Appendix Section 2: Teacher Survey	8
Appendix Section 3: Targeted Students Survey	19

Appendix Table A1: Summary Statistics

	N	Mean	SD
Teacher ranked effort			
Teacher self-reports of effort	14,275	2.968	1.411
Student ranks, focal teacher reports			
Academic ability	14,137	2.968	1.410
Disruption	12,184	2.970	1.411
Student ranks, self-reports			
Academic ability	12,724	3.601	1.224
Disruption	12,577	2.368	1.331
Student's background			
Lives in an urban area	2,643	0.473	0.499
Female	2,643	0.588	0.492
Romanian	2,632	0.975	0.158
Number of siblings	2,633	1.146	1.236
Received tutoring	2,623	0.358	0.480
Rosenberg index	2,643	0.001	0.990
Bullying exposure	2,643	-0.010	0.991
Mother's education level			
Less than high school	2,379	0.261	0.439
Finished high school	2,379	0.486	0.500
College or more	2,379	0.253	0.435
Teacher characteristics/behaviors			
Female	1,843	0.659	0.474
Teaching experience	1,820	19.609	10.427
Targets the top third (of academic ability)	1,787	0.123	0.328
Targets the middle third	1,787	0.184	0.387
Targets the bottom third	1,787	0.055	0.229
Treats all students the same	1,787	0.639	0.481
Instrumental variables			
Non-focal teacher reports of ability (preferred)	13,636	2.967	0.947
Non-focal teacher reports of disruption (preferred)	11,548	2.970	0.990
Admission score (ability)	18,430	0.056	0.958
Peer reports from the same class (ability)	16,299	2.786	1.119
Peer reports from the same class (disruption)	16,183	3.074	1.065
Peer reports from the other class (ability)	14,854	2.660	1.076
Peer reports from the other class (disruption)	14,864	2.679	1.075
Teacher unrelated subjects (ability)	16,876	2.968	1.056
Teacher unrelated subjects (disruption)	16,184	2.971	1.091
Group 1 of non-focal teachers (ability)	12,997	2.968	1.071
Group 2 of non-focal teachers (ability)	12,459	2.970	1.116
Group 1 of non-focal teachers (disruption)	10,844	2.966	1.110
Group 2 of non-focal teachers (disruption)	10,335	2.973	1.154

N: Number of observations; SD: Standard deviation.

Appendix Table A2: OLS Estimates of Student Ability and Disruption on Teacher Effort, Including Control Variables

	Likely exogenous control variables			Broader control variables		
	Dependent variable: Teacher effort					
	(1)	(2)	(3)	(4)	(5)	(6)
Academic Ability	-0.231*** [0.019]		-0.175*** [0.018]	-0.229*** [0.019]		-0.173*** [0.018]
Disruptive behavior		0.260*** [0.018]	0.210*** [0.018]		0.259*** [0.018]	0.210*** [0.018]
Student is female	-0.164*** [0.050]	0.022 [0.051]	0.030 [0.051]	-0.157*** [0.050]	0.026 [0.052]	0.035 [0.051]
Student is Romanian	-0.130 [0.254]	-0.210 [0.239]	-0.165 [0.250]	-0.157 [0.254]	-0.246 [0.239]	-0.195 [0.250]
Student number of siblings	-0.010 [0.019]	-0.015 [0.019]	-0.010 [0.018]	-0.016 [0.019]	-0.021 [0.019]	-0.016 [0.018]
Mother's education level (Reference category: Less than high school)						
Completed high school	-0.091* [0.054]	-0.089* [0.054]	-0.088* [0.053]	-0.078 [0.054]	-0.072 [0.054]	-0.075 [0.053]
College or more	-0.090 [0.071]	-0.096 [0.071]	-0.066 [0.070]	-0.074 [0.072]	-0.070 [0.073]	-0.045 [0.071]
Student received tutoring				-0.116* [0.061]	-0.139** [0.060]	-0.119** [0.059]
Rosenberg index				-0.024 [0.024]	-0.039 [0.024]	-0.030 [0.024]
Bullying exposure				0.051* [0.026]	0.043* [0.026]	0.039 [0.026]
N	8,233	8,233	8,233	8,233	8,233	8,233
R-squared	0.125	0.134	0.160	0.127	0.137	0.162
Mean of dependent variable	2.950	2.950	2.950	2.950	2.950	2.950

Notes: The dependent variable is the focal teacher's within-class ranking of relative effort devoted to each sampled student, ranging from 1 (lowest) to 5 (highest). Academic ability and disruptive behavior are the corresponding within-class rankings reported by the focal teacher. Each column represents an independent regression. All regressions include binary variables indicating the alphabetical order of students' names and classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A3: First Stage Estimates

	Academic ability		Disruptive behavior	
	(1)	(2)	(3)	(4)
Original IV: Reports from other teachers (N = 11,214)				
Academic Ability IV	0.483*** [0.016]	0.425*** [0.018]		-0.153*** [0.018]
Disruptive behavior IV		-0.137*** [0.017]	0.497*** [0.016]	0.439*** [0.018]
F statistic of excluded instruments	863.5	469.3	975.5	562.4
Alternative IV 1: Admission score (N = 11,249)				
Academic Ability IV	0.497*** [0.020]	0.429*** [0.020]		-0.185*** [0.020]
Disruptive behavior IV		-0.220*** [0.016]	0.495*** [0.016]	0.462*** [0.016]
F statistic of excluded instruments	621.2	418.1	975.1	562.5
Alternative IV 2: Peer reports from same class (N = 10,494)				
Academic Ability IV	0.402*** [0.014]	0.367*** [0.015]		-0.150*** [0.015]
Disruptive behavior IV		-0.126*** [0.016]	0.370*** [0.015]	0.323*** [0.016]
F statistic of excluded instruments	830.1	456.1	580.9	343.7
Alternative IV 3: Peer reports from the other class (N = 9,508)				
Academic Ability IV	0.319*** [0.017]	0.295*** [0.017]		-0.192*** [0.017]
Disruptive behavior IV		-0.154*** [0.017]	0.271*** [0.017]	0.241*** [0.017]
F statistic of excluded instruments	341.8	219.8	246.4	195.8
Alternative IV 4: Teacher unrelated subjects (N = 10,717)				
Academic Ability IV	0.399*** [0.016]	0.347*** [0.016]		-0.122*** [0.016]
Disruptive behavior IV		-0.132*** [0.016]	0.440*** [0.015]	0.396*** [0.017]
F statistic of excluded instruments	652.7	354.3	818.7	449.7
Alternative IV 5: Over-identified (two groups of teacher colleagues; N = 9,333)				
Academic Ability IV, Group 1	0.290*** [0.018]	0.259*** [0.018]		-0.080*** [0.018]
Academic Ability IV, Group 2	0.277*** [0.017]	0.251*** [0.017]		-0.072*** [0.016]
Disruptive behavior IV, Group 1		-0.050*** [0.017]	0.296*** [0.017]	0.262*** [0.018]
Disruptive behavior IV, Group 2		-0.078*** [0.016]	0.271*** [0.016]	0.242*** [0.016]
F statistic of excluded instruments	507.3	263.2	491.9	267.4

Notes: Academic ability and disruptive behavior are measured using teachers' within-class rankings of each sampled student and range from 1 (lowest) to 5 (highest). Each column within each instrument set reports a separate first-stage regression. All regressions include classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A4: Estimates of Student Ability, Disruption, and Teacher Effort—Alternative Specifications and Models

	Dependent variable: Teacher effort			
	Binarized effort (1,2 vs 3,4,5)	Binarized effort (1,2,3 vs 4,5)	Plackett-Luce	Bradley-Terry
	(1)	(2)	(3)	(4)
Panel A: OLS/ML estimates				
Ability	-0.107*** [0.013]	-0.103*** [0.014]	-0.162*** [0.012]	-0.139*** [0.012]
Disruption	0.214*** [0.014]	0.203*** [0.014]	0.190*** [0.012]	0.177*** [0.011]
N	11,214	11,214	31,226	21,100
Mean of DV	0.594	0.392	0.286	0.459
Panel B: IV estimates				
Ability	-0.280*** [0.072]	-0.227*** [0.060]	-0.311*** [0.047]	-0.217*** [0.071]
Disruption	0.253*** [0.064]	0.232*** [0.055]	0.179*** [0.043]	0.254** [0.067]
N	11,214	11,214	31,226	21,100
Mean of DV	0.594	0.392	0.286	0.459

OLS: Ordinary least squares; ML: Maximum likelihood; DV: Dependent variable; N: Number of observations.

Notes: Teacher effort is measured using teachers' within-class rankings of how much time and effort they devote to each sampled student, and ranges from 1 (lowest relative effort) to 5 (highest relative effort).

Academic ability and disruptive behavior are teachers' within-class rankings of each sampled student's academic ability and disruptive behavior, respectively, and range from 1 (lowest) to 5 (highest). Columns (1) and (2) report linear models estimated via OLS in Panel A and IV in Panel B. Columns (3) and (4) report nonlinear models estimated by maximum likelihood, and their IV specifications in Panel B use a two-stage control-function approach. All models include classroom-by-teacher fixed effects. Robust standard errors clustered at the classroom-by-teacher level are reported in brackets. For the nonlinear IV models in Columns 3 and 4 of Panel B, standard errors are obtained from 1,000 classroom-by-teacher cluster-bootstrap replications.

*** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A5: OLS Estimates of Gender and Socioeconomic Differences in Student Ability and Disruptiveness

	Academic ability			Disruptive behavior		
	(1)	(2)	(3)	(4)	(5)	(6)
Student is female	0.320*** [0.049]		0.338*** [0.049]	-0.989*** [0.048]		-1.003*** [0.048]
Mother completed college		0.163*** [0.059]	0.206*** [0.059]		-0.034 [0.058]	-0.160*** [0.055]
N	8,821	8,821	8,821	8,821	8,821	8,821
R-squared	0.097	0.090	0.100	0.169	0.086	0.170
Mean of DV	3.026	3.026	3.026	2.944	2.944	2.944

DV: Dependent variable; N: Number of observations

Notes: Academic ability and disruptive behavior are teachers' within-class rankings of each sampled student's academic ability and disruptive behavior, respectively, and range from 1 (lowest) to 5 (highest). Each column represents an independent regression. All regressions include classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A6: IV Estimates of Student Ability, Disruption, and Teacher Effort, Heterogeneity Analysis

Interacting variable	DV: Teacher effort							
	School quality	Urban	Top class	Female teacher	Teacher experience	Relevant subject	Female student	High SES
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AA	-0.255*** [0.080]	-0.288*** [0.063]	-0.280*** [0.082]	-0.232** [0.096]	-0.266*** [0.093]	-0.300*** [0.072]	-0.214*** [0.079]	-0.277*** [0.074]
DB	0.233*** [0.065]	0.274*** [0.057]	0.260*** [0.070]	0.089 [0.094]	0.159** [0.081]	0.210*** [0.069]	0.293*** [0.070]	0.242*** [0.059]
AA x interacting variable	-0.052 [0.110]	0.126*** [0.048]	-0.011 [0.111]	-0.085 [0.117]	-0.020 [0.115]	0.058 [0.110]	-0.001 [0.050]	0.049 [0.048]
DB x interacting variable	-0.029 [0.104]	-0.080* [0.047]	-0.087 [0.101]	0.187* [0.111]	0.129 [0.104]	0.022 [0.098]	0.054 [0.052]	-0.048 [0.049]
N	11,214	9,067	11,214	11,214	11,115	11,214	9,067	8,245
R-squared	0.078	0.092	0.074	0.079	0.076	0.075	0.085	0.086
Mean of DV	2.976	2.959	2.976	2.976	2.976	2.976	2.959	2.948
KP F statistic	21.97	45.29	26.79	17.40	21.61	31.20	28.76	36.44

DV: Dependent variable; AA: Academic ability; DB: Disruptive behavior; N: Number of observations; KP: Kleibergen-Paap rank Wald

Notes: Teacher effort is measured using teachers' within-class rankings of how much effort they devote to each sampled student, and ranges from 1 (lowest relative effort) to 5 (highest relative effort). Interacting variables are binary indicators for schools above the sample median admission score, students living in urban areas, top-track classes, female teachers, teachers with at least 18 years of experience, STEM subjects, female students, and students whose mothers attended college. All regressions include classroom-by-teacher fixed effects. Robust standard errors in brackets are clustered at the classroom-by-teacher level.

*** p<0.01, ** p<0.05, * p<0.10.

Appendix Section 1: Addressing the Ordinal Nature of Rankings

This appendix section describes alternative specifications to address the ordinal nature of our ranking outcomes, which are presented in Appendix Table A4.

Panel A estimates non-IV versions of each alternative specification. Columns (1) and (2) use binary outcomes based on effort ranks and estimate linear probability models with teacher-by-classroom fixed effects. Column (3) uses the Plackett-Luce rank-ordered logit model, which treats each teacher's full-effort ranking as a sequence of choices: first, the highest-effort student is chosen from all sampled students in the teacher-classroom; then the highest-effort student is chosen among those remaining; and so on. Column (4) uses a Bradley-Terry type pairwise model. We reshape each ranking into all unique student pairs within a teacher-classroom and define the outcome as 1 if the first student in the pair is ranked higher in effort than the second student, and 0 otherwise. Ability and disruptiveness are similarly transformed into binary indicators equal to 1 if the first student is ranked higher than the second student on that dimension.

Panel B estimates the corresponding IV specifications. Columns (1) and (2) use the same binary effort outcomes and estimate IV linear probability models, instrumenting the focal teacher's ratings of student academic ability and disruptive behavior with non-focal teachers' ratings, while including teacher-by-classroom fixed effects. For columns (3) and (4), we use a control-function, or two-stage residual inclusion, approach. In the first stage, we regress the focal teacher's academic ability and disruptive behavior rankings, or the corresponding pairwise indicators in the Bradley-Terry model, on non-focal-teacher instruments while absorbing teacher-by-classroom fixed effects. We then include the resulting residuals in the nonlinear second-stage models, alongside the original ability and disruptiveness measures. For the Plackett-Luce model, the second stage is estimated using the manually expanded conditional logit representation of the ranking; for the Bradley-Terry model, the second stage logit includes teacher-by-classroom fixed effects directly. This approach preserves the ranking or pairwise-comparison structure of the data while addressing endogeneity through a control-function correction.

Appendix Section 2: Teacher Survey

TEACHER SURVEY PART I

THE GALLUP ORGANIZATION
ROMANIA
London str. 6, sector 1 Bucharest
Tel: 021-318 1497

PART A
Project **159970**

Locality:
Code City:
Code interviewer

Name
op. interview
Name of the school:

Teacher code:

Interview date: (day, month)

The supervisor's name:

Duration of the interview (minutes)

Hello, I'm Gallup interview operator Romania. Gallup Romania is conducting a study of teachers in Romania. The survey is conducted in partnership with the Ministry of Education and Research. I would like to ask you some questions related to the beginning. and then your students. Your answers shall be confidential.

Demographic info

A0: Place of birth:

98 . Urban:

99 . Rural:

County:-----

Code county (according to the list D):

A1: Is the respondent male or female?

1 . Male 2. Female

A2: What is your nationality?

1 . Romanian 2. Hungarian 3. Roma 4. German 5. Another

A3: Date of birth

Year: Month: Day: 0. No answer

A4: Current Domicile:

Locality.... County

A5: Marital Status

1 . Married 2. Live together 3. Divorced 4. Widowed 5. Never married

A6: How many children do you have?

A7: What is the highest level of (graduated) education?

- 01 Primary school (grades 1-4)
- 02 Secondary school (grades 5-8)
- 03 High school (grades 9-12)
- 04 Vocational school, 1-4 years
- 05 Complementary school, 1-2 years,
- 06 School of craftsmen (foreman), 1-3 years
- 07 Post-secondary school, 1-3 years
- 08 Short-term higher education (SSD) (3 years of College)
- 09 Long-term higher education
- 10 Master's degree
- 11 PhD

A8: What was your Baccalaureate grade?

A9: What college did you graduate from?

College name

Profile/Specialization

1. State 2. Private

Locality: Code county (*according to the list D*):

0. I did not graduate from any university

Questions about current occupation and education

A10: What is your “teaching degree”?

1: Teaching Degree 1; 2: Teaching Degree 2; 3: Completed; 4: Beginner

In what year did you take the exam for “Teaching Degree 1”? And what was your grade?

In what year did you take the exam for “Teaching Degree 2”? And what was your grade?

A11: Your current position is:

- 1 Professor with higher education
- 2 Professor SSD (with short-term studies)
- 3 Teacher 1 (lower than professor) with higher education
- 4 Teacher 2 (lower than professor) with SSD
- 5 Teacher, educator, foreman, with secondary education
- 6 Professor, teacher, instructor, with high school level education, without specialized training

A12: What subjects are you currently teaching?

- 1 . Biology 2. Chemistry 3. English 4. Physics
- 5 . Philosophy 6. Geography 7. German 8. Informatics
- 9 . History 10. Math 11. Music 12. Religion
- 13 . Romanian 14. Sports
- 99 . Other subjects:

A13: How much have you worked so far in the field of education:

A14: Are you a permanent (tenured) or substitute employee in the position you hold?

- 1. Tenure 2. Not-tenure/Substitute

C1. I have to read you a list of intervals, could you please tell me in which of these fits your total household income from last month, i.e., the total amount of money obtained by all household members from salaries, pensions, business on their own, and rent allowances, etc.? (in RON)

- | | |
|--------------------|--------------------|
| 1. under 2,500 RON | 6. 6,501-7,500 RON |
| 2. 2,501-3,500 RON | 7. over 7,500 RON |
| 3. 3,501-4,500 RON | 8. NS/NR |
| 4. 4,501-5,500 RON | |
| 5. 5,501-6,500 RON | |

C2. To what extent do you agree or disagree with the following statements about the hours for the subject(s) you are teaching are allocated in your high school:

<i>Statements</i> <i>/One possible answer on each line/</i>	Strongly agree	Agree	Disagree	Strongly disagree
1. Teachers with seniority have priority in choosing the class/generation that they will teach	4	3	2	1
2. Teachers with the best training in their area have priority in choosing the class/generation that they will teach	4	3	2	1
3. In general, teachers with the highest seniority teach classes with students who have the best performance	4	3	2	1
4. In general, teachers with the best training in their field teach classes with students who have the best performance	4	3	2	1
5. All or most of the teachers teach at least one class with students who have the best performances	4	3	2	1
6. As a teacher, you prefer to teach in classes that have the best students in the school	4	3	2	1

C3. In your class, do you work in small groups of students?

1. Yes
2. Not at the moment, but I am planning for next year
3. No

C4. (Before your class) how much time do you expect a student to prepare for your class?

1 15 minutes **2** 30 minutes **3.** 45 minutes
4 1 hour **5.** more than 1 hour **0.** Not at all

C5. Students in the same class often have different levels. When you teach a class, on which group of students you concentrate/focus more?

1. the lower 1/3
2. middle 1/3
3. the best 1/3

C6. In your experience, what percentage of your students at this school do you think should take private tutoring in the subject you teach?

... ... (answer in%)

C7. In your experience, what percentage of your students at this school take private tutoring in the subject you teach?

... ... (answer in%)

TEACHER SURVEY PART II

Name of School:

County:

Grade:

Name of Teacher:

Subject

Teacher ID

1.) You have been teaching the following 10 students this year [CONFIRM WITH TEACHER]:

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

2.) Please take into account the time and effort you devote to your students. For example, you can think about preparing homework or additional individualized materials for certain students, individually monitoring some students to gauge their progress, the time allocated to answering extra clarifications/questions about the material taught, the time spent resolving problematic situations created by certain more disruptive students, and so on. Please rank the above 10 students in terms of how much time and effort you devote to each.

Place the student for whom you spend the MOST time and effort first, and the student for whom you spend the LEAST time and effort last. The list below must include the names of the students mentioned above.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

3.) Some students are more disruptive than others. For example, some students disrupt the learning process by making too much noise, harassing their neighbors, etc. In the list below, place the students listed above. Place the one that is the MOST disruptive at the top, and the student who is the LEAST disruptive at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

4.) Some students are more self-confident than others. For example, some students are more likely to express their opinions and to take risks. They might ask or answer questions even if there is a risk of ridicule, or they may trust themselves more when taking an exam. In the list below, place the students listed above. Place the one who is the MOST self-confident at the top, and the student who is the LEAST self-confident at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

The next three questions will ask you to rank the list of 10 students in terms of three dimensions: **effort**, **ability**, and **performance**. Most researchers of education believe that ability and effort are the two separate ingredients shaping academic performance. For example, some students have a very high raw ability: they have a very easy time understanding difficult concepts; we think of them as being highly intelligent. At the same time, student effort is also considered an important driver of academic success. For example, some students spend a lot of time learning and reviewing the materials either at home, doing homework, or working hard and paying attention in school. In this sense, there are four types of students: (1) some are very intelligent and hardworking, (2) some are not as bright but nonetheless work really hard, (3) some are intelligent but lazy, and finally (4) some are neither intelligent nor hardworking. Please answer the three questions below.

5. Some students work harder on their studies than others. For example, some students—even if they are not necessarily among those with the highest academic performance (grades) or the highest innate intelligence—work very hard on homework or pay a lot of attention in class. In the list below, place the students listed above. Place the one who exerts the MOST effort at the top, and the student who exerts the LEAST effort at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

6. Consider these students in terms of their native ability or native intelligence. In the list below, place the students listed above. Place the one that has the highest ability at the top, and the student who has the lowest ability at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

7. The “Bacc” exam is a measure of students’ academic performance. Consider how you expect your students will do in terms of their total Bacc score. In the list below, place the students listed above. Place the one that you think will have the HIGHEST Bacc score at the top, and the student who will have the LOWEST scores at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

8. What is the average grade that the students on the list below received/will receive in the school year 2016-2017 in your class?

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

Now I would like to ask you some general questions about your experiences teaching two classes this year: 9_X and 9_XX in High School Y.

9. Consider classes 9_X and 9_XX, and think about the effort you exert. For example, you prepare lessons, you use your own energy in lecturing, you monitor students' performance, etc. Which class would you say you have devoted more total energy to (CIRCLE ONE)?

9_X compared to 9_XX

10. Consider classes 9_X and 9_XX, and think of the level of difficulty of instruction. Which class would you say that you teach at a higher level of difficulty (CIRCLE ONE)?

9_X compared to 9_XX

11. Consider classes 9_X and 9_XX, and think of the number of disruptive students in these classes. Which class would you say has more disruptive students (CIRCLE ONE)?

9_X compared to 9_XX

12. Consider classes 9_X and 9_XX, and think of the effort that students exert in these classes. Which class would you say has more students who exert a lot of effort (CIRCLE ONE)?

9_X compared to 9_XX

13. Consider classes 9_X and 9_XX, and think of the native ability of the students in these classes. Which class would you say has more students with high native ability (CIRCLE ONE)?

9_X compared to 9_XX

14. Consider classes 9_X and 9_XX, and think of the academic performance of the students in these classes. Which class would you say has more academically high-performing students (CIRCLE ONE)?

9_X compared to 9_XX

15. Consider classes 9_X and 9_XX, and think of your expectations in terms of achieving certain learning objectives in these classes. For which class would you say you had higher expectations to reach high learning objectives? (CIRCLE ONE)?

9_X compared to 9_XX

16. Consider two classes you have taught this year: 9_X and 9_XX, and think of how much you enjoyed instruction in these classes. Which class would you say that you have enjoyed teaching more? (CIRCLE ONE)

9_X compared to 9_XX

Finally, I would like to ask you a question about where you target your instruction for a typical class. For example, some teachers spend most of their time targeting the academically most able students. They like to teach difficult material that challenges the most able students in the class. Other teachers teach on most days easier material that even the least able students in the class can understand and follow. Finally, some other teachers try to target the students in the class who are in the middle in terms of their academic ability. These students in the middle are neither the most nor the least academically capable students in the class.

17. I would now like to ask you a question about your experiences teaching 9_X at High School Y this year. To which set of students do you aim your instruction? (CIRCLE ONE)

- 1 Bottom 1/3 of students
- 2 Middle 1/3 of students
- 3 Top 1/3 of students
- 9 I treat them all the same

18. I would now like to ask you a question about your experiences teaching 9_XX at High School Y this year. To which set of students do you aim your instruction? (CIRCLE ONE)

- 1 Bottom 1/3 of students
- 2 Middle 1/3 of students
- 3 Top 1/3 of students
- 9 I treat them all the same

Appendix Section 3: Targeted Students Survey

CHILD SURVEY PART I

Questionnaire pupils

THE GALLUP ORGANIZATION
ROMANIA
London str. 6, sector 1
Bucharest
Tel: 021-318 1497

Commune Code
City:
Name
op. interview
High school high
school Code:
Code

Interviewer:
Interview date: (day)
Supervisor Name
Duration: minutes

Hello, I'm Gallup interview operator Romania. Gallup analysis at the level of high school students from Romania. The survey is conducted in partnership with the Ministry of Education and research. I would like us to reply to some questions related to the beginning of high school you're driving. Your answers shall be confidential.

Part A: Socio-demographic Questions about the student and their family

Name and surname of the student :
Class:

A0: Place of birth:

98 . Urban :

99 . Rural:

(according to the list D):

Code county (judet)

A1: Sex:

1 . Boy 2. Girl

A2: What is your nationality?

1 . Romanian 2. Hungarian 3. Roma 4. German 5. Another

A3: Date of birth:

Year: Month: Day:

A.4.1: How many brothers and/or sisters do you have?

A.4.2: How many of your brothers and/or sisters are older than you?

Now I would like you to respond to some questions about your parents:

Table 1. The student's parents:

	Month and year of birth		Nationality 1 . Romanian 2 . Hungarian 3 . Rome 4 . German 5 . Another	Marital status 1 . Married 2 . We live together 3 . Divorced 4 . Widowed 5 . Not married	Occupation (in the month) (coded according to the codes in box B)	Field of activity (in the month) (coded according to the codes in box C)	The last school graduated (coded according to the codes in box A)
	Month (NUMBER)	Year					
THE STUDENT'S FATHER			1 2 3 4 5	1 2 3 4 5			
THE STUDENT'S MOTHER			1 2 3 4 5	1 2 3 4 5			

Box A. Last school GRADUATED (column 6, section 1, Table 1-student's parents)

1. No official school year
2. The primary school (grades 1-4)
 3. Secondary school (grades 5-8)
 4. Vocational school or apprenticeship
 5. First stage in high school (grades 9-10)
 6. High school (grades 9-12)
 7. School of craftsmen or post-secondary school education of foremen
 8. State College
 9. Private College
 10. Master's degree or higher (e.g., PhD) from a State school
 11. Master's degree or higher (e.g., PhD) at a private college

Box B. Occupation (column 4, section 1, Table 1-student's parents)

1. Employee
2. *Employee in another country (working abroad)*
3. Patron
4. Freelancer, not in agriculture
5. On their own, in agriculture
6. Membership in the non-agricultural cooperative
7. Member of the agricultural cooperative
8. Unemployed
9. Retiree
10. Pupil, student
11. Housewife
12. In military service
13. Other (elderly person, disabled, addicted, etc.)

Box C. activity field (column 5, section 1, Table 1-student's parents)

1. Agriculture, forestry, fishing, hunting
2. Extractive industry
3. Processing industry
4. Electricity, heat, gas, water
5. Construction
6. Wholesale trade and retail trade, hotel and restaurant industry
7. Transport, storage, communications
8. Finance, banks, insurance
9. Real estate, rentals, and administrative services for companies
10. Public administration and defense, social security
11. Education
12. Healthcare and social assistance
13. Other social and personal services
14. Domestic help services contract
15. Activities of international organizations and agencies
16. Maintenance and repair

Box D. County

1. Alba 10. Brăila 19. Galati 28. Mehedinți 37. Teleorman
2. Arad 11. Buzau 20. Galt 29. Mures 38. Timis
3. Baton Rouge, 12. Caraș-Severin 21. 30, gorj County. Neam, 39. Tulcea
4. Bacău 13. Calarasi 22. Harghita 31. OLT 40. Vaslui
5. Bihor 14. Cluj 23. Hunedoara 32. Prahova 41. Vâlcea
- 6.15, Bistrița-Năsăud. Konstanz 24. Ialomița 33. Satu Mare 42. Vrancea
7. Brisbane 16. Covasna 25. Iasi, 34. Sălaj **43. Overseas**
- 8. Bucharest** 17. Dâmbovița 26. Ilfov, 35. Sibiu
9. Brasov 18. Dolj 27. Maramures 36. Suceava

A5: In general, how satisfied are you with your current school?

- 1 . Very satisfied
- 2 . Satisfied
- 3 . Dissatisfied
- 4 . Very dissatisfied

A6: How long does it take from your home to your school on an ordinary day? (in minutes)

minutes:

A7: Do you intend to attend college?

- 1 . Yes
- 2 . No

A8: Your parents enforce rules regarding the manner in which you study (for example, how much time you devote to studies versus how much time you devote to other activities outside of

school (sports). Do they allow you to do certain things if they are satisfied with your performance in school?

- 1 . Yes
- 2 . No

A9: How often do they follow these rules? (*ONLY ONE ANSWER POSSIBLE*)

- 1 . Never
- 2 . Very rarely
- 3 . From time to time
- 4 . Quite often
- 5 . All the time

[SPONTANEOUSLY, DO NOT READ] 6 . You don't need anybody to pressure/constrain- I do them voluntarily anyway

A10: How often are your parents asking about your situation in school, checking or helping with homework, etc.?

1. never
2. Very rare
3. From time to time
4. Quite often
5. All the time

A11: During this school year, your parents:

1. ... Have volunteered for various activities in class, your school, or library? YES/NO. ...
2. ... Discussed with the project supervisor? YES/NO. ...
3. Discussed with another teacher of yours? YES/NO. ...

A12: During this school year, have you received private tutoring?

1. **No**
2. **Yes**

If Yes, on which subjects:

- a) Math
- b) Romanian
- c) English
- d) French
- e) Physics
- f) Chemistry
- g) History
- h) Geography
- i) Other subject _ _ _ _

A13: What is your overall (expected) average grade in the current school year, 2016-2017?

--

A14:

Some questions about your experiences with other classmates from the school.

Throughout this school year, how often did your classmates...?

<i>OPERATOR!</i> <i>READ EACH VARIANT</i> <i>ONE POSSIBLE ANSWER ON</i> <i>EACH LINE</i>	Never	Very rarely	Rare	Often	Very often	Always
1. Helped/worked together to study/do homework?	0	1	2	3	4	5
2. Spent time with you outside school?	0	1	2	3	4	5
3. Verbally abused you or said nasty things?	0	1	2	3	4	5
4. Hit you?	0	1	2	3	4	5
5. Took your belongings (e.g. money, sandwich) without asking for your permission?	0	1	2	3	4	5
6. Excluded you from activities with friends (for example, sports activities, outings) intentionally?	0	1	2	3	4	5

A15:

Please tell me to what extent you agree or disagree with each of the following statements:

<i>OPERATOR!</i> <i>READ EACH VARIANT</i> <i>ONE POSSIBLE ANSWER ON</i> <i>EACH LINE</i>	Strongly agree	Agree	Disagree	Strongly disagree	I do not know
1. I feel that I'm a valuable/important person, at least as much as others	4	3	2	1	0
2. I feel that I have a few qualities/strengths	4	3	2	1	0
3. In general, I tend to believe that I'm no good	4	3	2	1	0
4. I can do things just as well as others.	4	3	2	1	0
5. I feel that I have good reasons to feel proud of myself	4	3	2	1	0
6. I have a positive attitude toward myself	4	3	2	1	0
7. In general, I am satisfied/happy with myself	4	3	2	1	0
8. I wish I could have more self-respect	4	3	2	1	0
9. I sometimes feel that I'm useless/worthless	4	3	2	1	0
10. Sometimes I think that I'm good for nothing	4	3	2	1	0

CHILD SURVEY PART II

Name of School:

County:

Grade:

Name of student:

Student ID:

1.) You have been in the same grade as the following 10 students (including yourself) this year
[CONFIRM THAT YOU KNOW THEM]:

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

2.) Some students are more disruptive than others. For example, some students disrupt the learning process by making too much noise, harassing their neighbors, etc. In the list below, rank yourself and your fellow students in terms of disruption. Place a 1 next to the student who is the MOST disruptive, and a 10 next to the student who is the LEAST disruptive

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

3.) Some students are more self-confident than others. For example, some students are more likely to express their opinions and to take risks. They might ask or answer questions even if there is a risk of ridicule, or they may trust themselves more when taking an exam. In the list below (which includes yourself), place the students listed above. Place the one who is the MOST self-confident at the top, and the student who is the LEAST self-confident at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

The next three questions will ask you to rank the list of 10 students in terms of three dimensions: **effort**, **ability**, and **performance**. Most researchers of education believe that ability and effort are the two separate ingredients shaping academic performance. For example, some students have a very high raw ability: they have a very easy time understanding difficult concepts; we think of them as being highly intelligent. At the same time, student effort is also considered an important driver of academic success. For example, some students spend a lot of time learning and reviewing the materials either at home, doing homework or working hard and paying attention in school. In this sense, there are four types of students: (1) some are very intelligent and hardworking, (2) some are not as bright but nonetheless work really hard, (3) some are intelligent but lazy, and finally (4) some are neither intelligent nor hardworking. Please answer the three questions below.

4.) Some students work harder on their studies than others. For example, some students—even if they are not necessarily among those with the highest academic performance (grades) or the highest innate intelligence—work very hard on homework or pay close attention in class. In the list below (which includes yourself), place the students listed above. Place the one who exerts the MOST effort at the top, and the student who exerts the LEAST effort at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

5.) Consider these students in terms of their native ability or native intelligence. In the list below (which includes yourself), place the students listed above. Place the one that has the highest ability at the top, and the student who has the lowest ability at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J

6.) The “Bacc” exam is a measure of students’ academic performance. Consider how you expect you and your fellow students will do in terms of the Bacc score. In the list below (which includes yourself), place the students listed above. Place the one that you think will have the HIGHEST Bacc score at the top, and the student who will have the LOWEST scores at the bottom.

1. STUDENT_A
2. STUDENT_B
3. STUDENT_C
4. STUDENT_D
5. STUDENT_E
6. STUDENT_F
7. STUDENT_G
8. STUDENT_H
9. STUDENT_I
10. STUDENT_J