School of Economics Working Paper
2022-11

University of East Anglia

SCHOOL OF ECONOMICS

Positive Disruption? Meritocratic Principal Selection and Student Achievement

Oana Borcan*
James Merewood**

*School of Economics, University of East Anglia
**RAND Europe

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

Positive Disruption? Meritocratic Principal Selection and Student Achievement

Oana Borcan
University of East Anglia[1]

James Merewood
RAND Europe[2]

Principals are the gatekeepers of education and can influence student achievement through management practises. However in many countries discretionary staff appointments, corruption and inefficiency undermine the quality of management and education. Meritocratic selection in public service has been advocated as a tool to elevate management quality. We analyse the short-term impact of the 2016 introduction of merit-based selection for Romanian state school principals on students' school-leaving test scores. Employing a staggered difference-in-difference strategy, we study the impact of competitively selected principals (compared to those appointed), and the impact of new principals (compared to principals who retain their position). The average treatment effect is small and insignificant immediately after the policy, with some evidence that new principals begin to improve outcomes two years on, particularly in schools with average historical performance. Since principals have limited management autonomy, this improvement is likely due to strategic selection of students into sitting the exam, but additional survey data also suggests the policy selects principals that are more motivated for the job. The evidence points to benefits and limitations of merit-based recruitment policies in education.

JEL codes: I21, I28, O15, M54

Keywords: Merit-based selection, public sector recruitment, school principals, test scores.

[1] Email: o.borcan@uea.ac.uk
[2] Email: james_merewood@randeurope.org. Work presented in this document was written whilst James Merewood was affiliated with the University of East Anglia.
The authors thank Andreea Mitrut, Mikael Lindahl and Sheheryar Banuri, Amrish Patel, Chris Starmer and Christa Brunnschweiler for their useful comments. We are grateful to Andrei Munteanu, who shared with us data on school admissions.

1. Introduction

Principals are the gatekeepers of education, overseeing the allocation of material and human resources to ensure quality learning, equity and student performance. In 2017, pre-university education expenditure was, on average, 3.5% of the GDP in OECD countries (OECD, 2020). School principals manage a large share of the national education budgets; it is therefore critical that principals are able to enhance school productivity and student attainment. A small number of recent studies show that education management practises are linked to student outcomes (e.g. Cilliers *et al,* 2022*,* Jacob *et al.*, 2018; Bohlmark *et al.*, 2016 Di Liberto *et al.*, 2015; Coelli and Green, 2011); for example, Bohlmark *et al.* (2016) demonstrate that students in schools with higher quality principals have improved outcomes.

As a result, the selection process of school principals is vitally important in order to ensure that capable and motivated individuals are entrusted with leadership roles. Despite this, many education systems outside high-income countries rely on discretionary appointments, either based on evaluative interviews (which can result in subjective and inconsistent assessments), or based on criteria such as nepotism and political affiliation, adversely affecting management performance.[3][4] Standardised and objective principal selection processes are an underexplored avenue for improving student learning and test scores. This is partly because meritocratic selection policies for principals utilising standardised testing are relatively rare. Moreover, national or state policies on principal (or more generally public sector) recruitment do not typically leave room for experimental variation in meritocratic or competitive selection processes.

In our study, we examine the impact of a gradually introduced meritocratic selection of principals in Romanian secondary schools on student outcomes using a quasi-experimental design. In the economics literature, there is emerging evidence linking objective rule-based selection of teaching staff to improved job performance (Estrada, 2019). In the private sector, employee recruitment is known to deliver better quality hires when utilising objective meritocratic selection, than when using discretionary appointments (Hoffman *et al.*, 2018). However, very few studies investigate how merit-based selections of school principals affect management performance or student outcomes, and none offers compelling causal estimates.

Romania provides a relevant context because its post-communist education landscape has been marred by corruption and nepotistic appointments. For instance, in Romanian universities, 23% of staff say that the person obtaining a job at the institution is often known before interviews are held (PEIS, 2007). Against this backdrop, a series of anti-corruption policies have been implemented since 2011, beginning with the introduction of CCTV monitoring of exams and tougher punishments for teachers and students (Borcan *et al.*, 2017). Another recent reform was a nationwide meritocratic principals selection policy which ran between 2016 and 2017. The policy was a hurried attempt by a provisional technocrat government to purge schools of political influence and came as shock to the education sector. In September 2016 *all* principal posts across the country were vacated and a national competition including standardised cognitive and management competency tests was open from

[3] Whilst not the direct focus of this paper, for more information on patronage or nepotistic appointments see Colonnelli *et al*. (2018), Scoppa (2009), Huseyin and Mustafa (2008).

[4] For example, Gurmu (2020) documents that principalship in Ethiopia is often denied to those who have the relevant level of education, and suggests that political affiliation is the main criteria used for selecting principals into positions. Similarly, Walker and Kwan (2012) highlight the role of cultural and religious affiliation in the selection of principals in Hong Kong; despite this, prospective principals in Hong Kong are required to provide proof of a pre-principal certification and must undertake an interview (Walker and Kwan, 2012). In many districts of the US, evidence suggests that most principals perceive merit-based selection to be the main practise (Palmer and Mullooly, 2015); however, several participants still perceive inequality in the selection process (Palmer and Mullooly, 2015).

October 2016. National elections in December saw a new political party in government, which halted the exam process before all schools were assigned principals, and deferred the competition over remaining schools to July 2017. A vast majority of principals who were eventually confirmed in a post in 2016 and 2017 (i.e., 87% of all principals), had to undergo the standardised exam, and were competitively selected based on their results.

Using administrative data and an independently conducted survey, we study the effects of this policy on the student outcomes in the national school-leaving exam, the Baccalaureate. Since the policy took place in two waves (one in 2016, one in 2017), the identification strategy uses a staggered difference-in-difference model to assess the change in student outcomes in schools with and without an exam selected principal pre-/post-policy (this is our first treatment – T1). Since recent literature (see Baker, Larcker and Wang, 2022) has highlighted the biases of the two-way fixed effects (TWFE) estimator, we use the estimator developed by Callaway and Sant'Anna (2021) which estimates group-time-specific treatment effects with the correct control groups.

This approach has the advantage that we can use nearly the entire population of Romanian schools, since we know all schools which have had an exam selected principal. The interpretation of this treatment is limited by the fact that the policy has de facto changed only around 30% of principals. Therefore, using a sample for which we have additional information on principals, we also study the impact of a change in principal on student test scores. Specifically, for a sample of schools with exam selected principals, we estimate the difference in student outcomes between schools with principals who retained their position (henceforth legacy principals) and those who replaced an existing principal as a result of the policy (new principals; this is our second treatment, T2).

First we estimate the effect of T1, to understand whether the policy overall had any impact on student outcomes. We find overall zero average treatment effects of competitively selected principals in a large sample of Romanian schools.[5] This is likely to be due to the fact that a large share (~70%) of schools with exam selected principals retain their pre-policy principal.

We then estimate the effect of T2 by comparing student outcomes for new principal schools, compared with those with legacy principals post-policy. We find that new principals have a positive and significant impact on exam pass rates in the longer-term (two to three years after the policy); as expected, these effects are very small and insignificant in the short-term (one year after the policy). Importantly, we find that the average treatment effect doubles in magnitude when comparing treatment effects from one year after the policy to those three years after the policy. The average treatment effects on the treated (ATTs) of new principals in 2019, the last year in our data, are 0.07 of a SD for the final Baccalaureate score, and approximately 8 percentage points higher passing rates compared to students in schools with legacy principals in 2012-2016. These findings are in line with literature which suggests that it takes time for principals to have an impact on student outcomes (Coelli and Green, 2011).

We next turn to the potential mechanisms through which this change may be brought about. As in many low- and middle-income countries, Romanian principals have little autonomy, with most decisions on curriculum or staff hiring and dismissal being made nationally or at the county level.[6] Even slight improvements in outcomes driven by new principals in low- to mid-performing schools can contribute to reducing education inequality. To explore the mechanisms through which new principals improve attainment in poor performing schools, we provide supplementary survey data on public

[5] We discuss the definition of our working sample in section 4.1.

[6] In high-income countries principals have considerably more autonomy. For example, Bohlmark *et al.* (2016) show that in Sweden, principals can recruit and dismiss teachers and have control over the hiring process.

service motivation and management practices inspired by Perry (1996) and the World Management Surveys (WMS, Bloom and Van Reenen, 2007), which we collected in 2017 after the second round of the competitive selection. The survey reveals little evidence that new competitive principals adopt different management practices in the short run. There is some suggestive evidence that new principals are more prosocial, indicating that they feel more motivated to fulfil their public mission. Despite this, using additional administrative data on Baccalaureate exam enrolments, we find strong evidence of strategic behaviour of principals: enrolment rates are significantly lower in schools with a new principal for both the 2018 and 2019 Baccalaureate exam sessions. This suggests that new principals are restricting permission for low-performing students to sit the main Baccalaureate exam. Whilst students have the opportunity to revise and sit the exam later in the year, restricting students in this way mechanically improves the formal Baccalaureate performance measures, with ambiguous effects on student learning. This finding raises questions about the power of meritocratic selection to enhance learning and student achievement in the absence of autonomy or effective reforms in managerial practices.

We also examine the policy impact across the distribution of school performance. We divide schools into three different percentile groups based on the schools' pre-policy average Baccalaureate scores: poor performance schools (<25th percentile), midrange performance schools (25th-75th percentile), and top performance schools (>75th percentile). For new principals selected in 2017, we find a significant positive impact on student outcomes in mid-performing (25th-75th percentile) schools. Specifically mid-performing schools with a 2017 new principal see exam pass rates which are 11.3% greater than the control group mean (a 5.1 percentage point increase).

In order to verify the validity of our results, we undertake several robustness checks. In addition to parallel trend tests, we run alternative estimations changing the control group (from never to not-yet treated). We also estimate OLS TWFE estimates and, based on these, run placebo policy tests which indicate that our results are indeed a consequence of the meritocratic selection policy.

Our paper contributes to at least two strands of literature. Firstly, we present complementary evidence to studies linking education management practices and performance to student outcomes. Several studies show a positive correlation between principal performance and student outcomes (e.g. Meyer *et al.*, 2020; Agasisti *et al.*, 2018; Masci *et al.*, 2018). Only a few studies have analysed this relationship through a causal lens, using quasi-experimental designs or tightly controlled econometric models (Coelli and Green, 2011; Dhuey and Smith, 2014; Bloom *et al.*, 2015; Di Liberto *et al.*, 2015; Bohlmark *et al.*, 2016). Generally, principals can influence outcomes by improving management practises (Di Liberto *et al.*, 2015), and higher management quality is shown to be positively associated with outcomes (Bloom *et al.*, 2015, Cilliers *et al.*, 2022). Coelli and Green (2011), demonstrate that principals influence outcomes for students who are already dedicated to improvement; specifically, principals who switch between schools typically have a greater impact in higher performing schools than in poorer performing schools, since students in higher performing schools are more dedicated to improvement. Importantly, the evidence suggests that the principal effects takes a few years to show, which the authors attribute to the time needed for a principal to change the ethos and management practises of the school. Dhuey and Smith (2014) show that principals can improve outcomes in math and reading exams; similar to Coelli and Green (2011), they show that the effect is strongest in higher performing schools.

One of the most important channels principals in many education systems have to influence student outcomes is the hiring and firing of teachers (Jacob *et al.*, 2010). In addition, principals must be able to utilise the skills of their teachers (Agasisti *et al.*, 2018). Meyer *et al.* (2020) highlight the impact which principals can have on the collaboration of teachers when the collective efficacy of

teachers is in line with the beliefs of the principal. However, the influence of principals on outcomes depends greatly on the level of autonomy which they have over teacher recruitment and other school decisions. In this vein, our study contributes complementary findings that there are limited short-term gains from meritocratic principal recruitment in a setting where principals have little decision power over key education inputs. The positive impact we find on student test scores with a lag in schools with new principals is most likely related to strategic management decisions to restrict permission for students to sit the exam. However, we cannot rule out that introducing a more prosocial environment with the school may also have contributed to improved outcomes.

Secondly, we contribute one of the first studies to investigate the effects of meritocratic staff selection in public education. Since principals are shown to influence student outcomes, selecting high quality principals is paramount. Meritocratic recruitment processes inherently favour the selection of managers with the highest ability. Ruiz-Tagle (2019) empirically link the meritocratic selection of principals in Chile with student outcomes 6 years after the introduction of the policy. Whilst they document no changes in management practises within these schools, student outcomes are shown to have improved since the introduction of the policy. However, while in Ruiz-Tagle's (2019) setting the policy was gradually rolled out across the country over a several waves, the TWFE estimator used raises concerns about possible bias in the estimates. Hsiao *et al.* (2012) examine a similar policy in Taiwanese secondary schools, but do not link the policy to student outcomes. Using interview techniques they find differences in management styles; competitively selected principals are more willing to transform management practises and encouraging democratic participation in management than traditionally appointed principals. We provide a significant contribution by using a robust econometric identification strategy to identify a causal link between meritocratic principal selection and student outcomes.

Significantly more literature exists in the meritocratic selection of teachers. Jacob *et al.* (2018) use data from Washington DC schools to show that performance of prospective teachers in screening measures, such as written assessments and sample lessons, is highly predictive of job performance. Similarly, Bruno and Strunk (2019) demonstrate that performance in specific screening measures is predictive of contributions to student achievement in the same areas of study. Despite this, Jacob *et al.* (2018) suggest that meritocratic selection of teachers is significantly under-utilised in the hiring decisions of public schools. Estrada (2019) consider differences between discretion based and rule based appointments of public school teachers in Mexico. The findings of Estrada (2019) are particularly strong, since they undertake analysis in a very large education system (of around 1 million teachers), and the estimation strategy controls for possible selection. Estrada (2019) demonstrate that teachers hired based on the discretion of the appointing manager perform considerably worse on the job than those appointed using a rule-based system. As previously discussed, whilst the literature highlights the importance of meritocratic selection of teachers on outcomes, there is a gap in the literature surrounding the meritocratic selection of their managers (principals); our study is, to our knowledge, the first to empirically study meritocratic selection at this level.

Outside of the education sector, meritocratic selection is shown to improve performance (Dahis *et al.*, 2020), increase legitimacy (Cortazar *et al.*, 2016), increase the quality of new hires (Hoffman *et al.*, 2018), and reduce discrimination (Tan, 2008). Charron *et al.* (2017) argue that meritocratic selection shifts the motivation of public sector workers away from political criteria and towards professional criteria (see also Yeboah-Assiamah *et al*., 2014). Dahis *et al.* (2020) show that even among state judges in Brazil, meritocratic selection is helpful in selecting the most competent candidates. Similarly, Hoffman *et al.* (2018) demonstrate meritocratic selection in the private sector improves the overall quality of new hires.

There are considerable difficulties when testing the causal impact of meritocratic selection of principals on student outcomes. The self-selection of principals into specific types of schools, such as inexperienced principals in poor performing schools (Loeb *et al.,* 2018; Branch *et al.*, 2012) threatens the identification of treatment effects. To counter these selection issues, Bohlmark *et al.* (2016) utilise a principal switching system in Sweden, which allows the inclusion of principal and school fixed effects as they change between schools to identify the effect of principals on outcomes. Whilst we do not observe the school placement of principals pre-policy, by using the new Callaway and Sant'Anna (2021) estimator, we obtain difference-in-difference estimates that are clean of the bias of two-way fixed effect estimates. In so doing, our main contribution is to bring causal evidence of the impact of school principals on student outcomes in the short- and medium term, using a unique policy and a rigorous identification strategy. We reinforce this evidence with alternative OLS difference-in-difference estimations including school fixed effects and county-specific trends, similar to Estrada (2019).

The paper is structured as follows. Background information, and detail on the policy are outlined in section 2. We present our data sources, working sample and survey in section 3. Section 4 provides detail on the empirical strategy, whilst section 5 presents the results and mechanisms. Section 6 concludes the paper.

2. Background

2.1 Meritocratic Selection of Principals

Prior to 2016, principal appointments and dismissals were decentralised to the county inspectorates and were often made in a haphazard manner, lacking clear competency criteria and transparency. The typical recruitment process included a CV and operational plan assessment. Contracts were typically three years, but contracts would often be extended on the inspector's order, and on an annual basis. Anecdotal evidence suggests that often appointments and extensions would be granted based on political connections rather than candidate ability. The introduction of meritocratic competition in 2016 sought to rectify this.

The year 2016 was a time of numerous attempted reforms to the Romanian public administration. Following mass protests as a result of the deaths of 64 young people in a club fire in November 2015, which was linked to corruption, the former prime-minister Ponta was forced to resign. As a result, a new interim government, entirely comprised of politically independent technocrats, held power from November 2015 till December 2016. In an attempt to remove the existing political influence within school management, this government introduced the nationwide meritocratic selection process for all public school principals.

In August 2016, the Romanian Ministry of Education announced that the new competition would screen both cognitive and managerial skills; the process applied to both principal and deputy-principal positions in all pre-university institutions, and effectively meant the dismissal of those currently holding the position. Any principal currently in position was therefore forced to run in the competition to continue on as principal of that school; other principals were also able to apply for the post, giving no guarantee that the existing principal would remain. As part of the policy, four year fixed terms were introduced for meritocratically selected principals.

The exam was comprised of three main components: a written test; a CV screening; an interview. The written test was held at a national level, and focused on assessing both cognitive skills (a logical reasoning test, receiving a 66% weight in the overall test score) and knowledge of the newest

management literature (33% weight) of applicants.[7] The CV screening and interview were judged by a county panel with the oversight of nationally appointed inspectors; these components focused on the motivation and management competencies of applicants, reflected in their one year operational plan. This process was transparent, and samples of all work were held and monitored by officials. Each component of the exam was awarded a maximum score of 50 points. In order to pass the exam, candidates had to score at least 35 points in each component and therefore a minimum score of 105 points overall – the highest written test score would break any tie. Principals were able submit multiple applications tailored for their preferred schools, however they could sit the test only once. Eligibility criteria to enter the competition included: having the relevant degree; have a "very good" professional, managerial, and moral record for the past 4 years; have had no disciplinary issues for the past 3 years.[8]

The meritocratic selection process was announced on 31st August 2016, and principals began to sit the exam as soon as October 2016 (applications were open between September 13th and October 2nd). The very short time period between announcement and enforcement left little time for prospective principals to gain an in-depth understanding of the selection process or to game the system. In addition, in this limited period, exam entrants were given considerable novel literature to learn, upon which parts of the written test were based. As a result, many schools saw no candidates apply for the position in 2016, whilst others saw candidates apply but fail the meritocratic selection process. These schools ultimately retained their principals for the rest of the year (and they form our control group for T1); the rest had a principal meritocratically selected in 2016 (around 75% of all schools).[9]

A resulting second wave of meritocratic selection occurred in July 2017, meaning that only schools which had not successfully appointed a principal through selection in 2016 were reopened for candidates to apply. In some instances, principals who failed the exam in 2016 were able to apply to the same school and be selected by exam in 2017. The exam process in 2017 was very similar to that in 2016; however, the written test covered knowledge of education legislation and managerial competencies based on different literature and excluded the logical reasoning test in 2017. Principals selected in 2017, took office in September 2017. Following the selection process in 2017, schools which either had no applicants or candidates selected by exam were appointed a principal by the county school inspector. This staggered intervention allows us to study the effects of competitive principals across time and different schools, and compare effects of the two meritocratic selection waves.

The timing of the exam selection of principals has significant impact in the way we must analyse the policy, a graphical representation of this timeline is shown in Figure 1. Since the selection process took place late in 2016, principals were placed in schools during the 2016/2017 academic year and were unable to influence student test scores for exams which took place in the summer of 2016 (2015/2016 academic year). Similarly, the meritocratic selection in 2017 took place after the 2017

[7] Since the written test is a purely objective form of assessment, the introduction of such a test (in the absence of leakage from other sources) purely increases the probability of a principal being meritocratically selected.

[8] One might be concerned that meritocratic selection may still lead to biased appointments. Stravakou (2019) argues that the interview portions of meritocratic selection policies are still likely to lead to subjective assessments. Despite this, we argue that the policy which we study in Romania is a discrete change from the previous system of appointments at the discretion of regional education regulators (often based on obscure criteria or clientelistic relationships) and towards a more merit-based form of selection.

[9] For the remaining 25% of schools, the technocratic government ruled that neither the incumbent nor the unsuccessful candidate could hold the position of interim principal; this meant that another person must be appointed. This stipulation was overruled by the new Social-Democratic government (appointed in December 2016), and was changed to allow incumbent principals to be reinstated until the meritocratic selection process could run again in 2017.

Baccalaureate exam. This is important for our analysis as effects of the policy must be considered on the following years student outcomes (2016 selected principals first impact on 2017 outcomes; 2017 selected principals first impact on 2018 outcomes). The timeline (Fig. 1) visually demonstrates the event horizon of the policy implementation, including when our survey was collected (details in the next section).

2.2 The Role of Principals in Student's Performance

Romanian school principals have very little autonomy in resource allocation and curriculum design. Romania's level of school autonomy is below the OECD average (OECD, 2014). The responsibilities of principals in Romanian schools are regulated through an Education Ministry's act,[10] and are limited to organising the entire educational process and delivering national education objectives. They are accountable for the school's performance, end-of-year evaluation and quality assurance processes. Principals also decide the annual budget and procurement processes, subject to approval from the school executive committee.

In terms of hiring decisions, principals have limited autonomy. They can propose new posts or submit vacancies to the county inspectorate, but the latter decide the final placement of tenured teachers in schools; however the principal's consultation with their school's executive committee regarding the new hires is often symbolic. Principals have some degree of freedom in recruiting substitute teachers (see ROFUIP 2020, Art. 21, paragraph 3). Also under the principal's remit are the training, integration and motivation of staff. Other responsibilities linked to the teaching and learning activities include: drafting internal regulations; assigning form teachers, school and extracurricular project coordinators; forming working groups and coordinating teaching and learning organisation such as timetabling; facilitating the professional development of staff. In terms of staff motivation and sanctions, principals have little autonomy. Salary scales and bonuses in state schools are regulated nationally, while contracts are protected by worker's rights. In extraordinary circumstances, the principals may propose to terminate staff contracts, but this must be approved by the inspectorate. However, principals contribute their statement in teachers' end of year evaluations, which may influence the inspectorate's allocation of merit bonuses. In terms of the routine monitoring of teacher's activities, principals are mandated to observe the teacher attendance records on a daily basis and make recommendations.

In sum, principals of Romanian state schools can, in the long run, influence education quality by maintaining a good school resource base and improving the school performance and reputation; in the short run, principals may improve staff discipline and attendance keeping, use non-pecuniary rewards to motivate staff and lead by example.

2.3 The Baccalaureate Exam

At the start of high school, Romanian students are sorted into subject specific tracks based on previous performance measures (this process is computerised), these are: Theoretical – including humanities and sciences; Technological – including technical training, and natural resource/environment focuses; Vocational – including arts, military and sports. Regardless of their track, Romanian high school students are required, at the end of their 12th grade, to undertake a nationwide, standardized test: The Baccalaureate. This exam process is especially important as strong weighting is applied on the grade obtained in the Baccalaureate during university and labour market considerations. The Baccalaureate exam takes place in June-July (first sit) and August (reassessment

[10] The Ministry's act for the organisation and functioning of the pre-university education establishments - "Regulamentul-cadru de organizare si functionare a unitatilor de invatamant preuniversitar" (ROFUIP 2020).

for students who fail or miss the first sit) each year, and includes oral and written tests; whilst some exams differ between tracks, all students (regardless of track) take identical exam papers for Romanian language and literature. This allows for a direct comparison between students across tracks. The Baccalaureate scores were problematic before 2011, because the grades were inflated through cheating and corruption; since 2012 tougher punishments have been introduced and the exams have been monitored with CCTV devices. This means that Baccalaureate test scores from the first sit have become a reliable measure of student performance (for an account of the impact of the anti-corruption campaign and further details about the Baccalaureate exam, see Borcan *et al.*, 2017). Therefore, we analyse the impact of the meritocratic selection of principals on student outcomes in the Baccalaureate exam from 2012-2019.

3. Data

3.1 Data Sources

We examine the impact of meritocratic selection on student test scores and explore the management mechanisms through which principals could enact change. To do so we use several sources to compile our outcomes, controls and to generate several exploratory variables. These sources are:

1) Principal selection data from each Romanian county and school from the Ministry of Education's administrative records of the meritocratic selection of principals in 2016 and 2017. [11] This data includes information about which candidates successfully passed the exam, and which were selected for the position.[12] From this data, we know which exam wave the principal was selected in, or whether there was no successfully appointed principal. Throughout the paper, we define a principal who was selected as part of the meritocratic selection policy as a "competitively selected principal" and any principals who were appointed without undertaking the meritocratic selection process as a "non-competitively selected principal". Whilst in principle we could use the entire universe of schools (1,637 schools), we restrict attention to a sample of schools for which we have more information on principal characteristics (see the working sample below).

2) Our working (reduced) sample consists of schools with data collected from county inspectorates; this data includes information on the year in which principals started their post. We use this data to divide the sample into principals who took their post as a result of the meritocratic selection policy (new principals), and those who remained managing the same school pre- and post-policy (legacy principals). Data was collected from 16 (out of 42) counties, who answered our Freedom of Information Act requests, and covers all schools within these counties. Our working sample is representative of the full sample (above) in terms of student outcomes both pre- and post- policy (Appendix Table A1). Within the working sample, 71% of schools (71.5% of students) had a Legacy Principal selected whilst 29% of schools (28.5% of students) had a new principal selected (Appendix Table A2).

[11] Whilst the meritocratic selection took place in all 41 Romanian counties, administrative data is only publicly available for 39 of 41 counties, and therefore two counties are excluded from analysis as we are unable to determine any principal appointments through meritocratic selection.

[12] This data also includes a full breakdown of exam test scores for each applicant (in some cases only final mark is available); whilst we do not use this data we do account for principal and school selection issues (see section 4).

3) Student outcome data, from by the Ministry of Education, which includes all students who enrolled in the Baccalaureate exam from 2012 to 2019.[13] In total, this data contains the following for over 1 million students in the full sample and 431,940 students in our working sample: the student's school and track; the exam subjects; the breakdown of exam scores; the final outcome of the overall Baccalaureate exam (including whether they passed); the rank of the student on a county and country wide level.

4) Survey data we collected independently from a randomly selected sample of 303 high school principals (around 20% of all high schools in Romania), carried out by the Romanian "Institute for Social and Political Studies" (hereafter ISSPOL).[14] We identify competitively and non-competitively selected principals and new or legacy principals (also cross-checked with county inspectorate data used to form our reduced working sample). Importantly, the survey includes useful information about the principals' characteristics (demographics and motivation for public service) as well as their managerial practises, which we use to examine mechanisms linking the policy to student outcomes.

3.2 Summary Statistics

Our reduced sample includes 510 schools (and 431,940 students) across 16 counties of Romania, over the 2012-2019 period. In the Appendix Table A1 we show that there are strong similarities between the average student outcomes in the full sample and the reduced sample. In the full sample, the average overall Baccalaureate score between 2012-2016 was 6.118, whilst in the reduced sample it was 6.176. The similarities also hold for post-policy student outcomes (2017-2019), where the full sample average was 6.832 whilst the reduced sample average was 6.846, as well as for other student outcome metrics examined. Of the schools contained within the reduced sample, 90.78% (463) had an exam selected principal whilst 9.22% had a non-competitively selected principal before/after the policy (Appendix Table A2). Of the 90.78% schools with an exam selected principal, 71% had a Legacy Principal whilst 29% had a New Principal. Significantly more new principals were appointed in the 2017 policy wave (48.4% of schools had a new principal in this wave) than in the 2016 wave (24% of schools had a new principal in this wave).

Within the reduced sample, schools with competitively selected versus non-competitively selected principals are comparable pre-policy in terms of the overall Baccalaureate score (student's average overall Baccalaureate score was 6.160 and 6.178 respectively; see Table 1), but schools with competitively selected principals have slightly better Romanian scores and pass rates. Conditional on having a competitively selected principal, schools with a Legacy Principal had significantly lower pre- and post-policy student outcomes than schools with a new principal (pre-policy student average was 6.137 for legacy compared with 6.280 for new principals, and post-policy outcomes were 6.787 compared with 7.045; see Table 1).

Whilst it is smaller than the reduced sample, the survey sample contains information from a greater number of counties across Romania (303 schools from 39 counties).[15] In Appendix Table A2

[13] Only exam results from 2012 onwards are used as in 2012 all Baccalaureate exams in Romania became monitored by camera in an attempt to curb cheating, see Borcan *et al.* (2017).

[14] The Romanian name is "Institutul de Studii Sociale si Politice". The Institute's website link (fully translatable) is available at: <www.isspol.ro>.

[15] The survey sample may not constitute a fully representative sample because there is a selection of schools who decided to respond to the survey questionnaire. For this reason, as well as the sample size, our main results are based on the reduced sample. However, the survey sample gives us extra information on the principals' characteristics and practise, which are useful for examining the mechanisms behind the results.

we show that 277 (90.5%) of surveyed schools have exam selected principals, whilst 29 (9.5%) did not. Of the exam selected principals, 25.6% began as a new principal in the school and the remainder were legacy principals. As in our reduced sample, considerably more new principals were selected during the 2017 policy wave in surveyed schools than in the 2016 wave (43% vs 13%; Appendix Table A2). Outcomes both pre- and post-policy are slightly lower in our survey sample than in the full and reduced samples, but are still very similar (Appendix Table A1).

These descriptive statistics make it clear that there was systematic selection into competition and change of management. Schools with a competitively selected principal in the 2016 wave were better on average, and new principals tend to be matched with higher performing schools. This has important implications for the identification of the competition impact on schools and student performance, as discussed in the following sections.

3.3 The Survey

Our survey was held by telephone interview, designed by us and conducted by ISSPOL – a private agency specialised in social research. The survey was carried out in November 2017, after both waves of the competition were over and when all principals were in office. The main advantages of telephone interviews are that they are less prone to social desirability bias and we could reach principals from a sample spanning almost the entire country. Interviews were between 15 and 20 minutes, and were held one-to-one with principals in Romanian high schools. The questions were adapted from the Perry (1996) survey on motivation for civil service and from the World Management Surveys (WMS; Bloom and Van Reenen, 2007). Our survey has four independent sections: General Questions; Motivation for the Role and Management Activity; Evolution of Management and School Performance; Meritocratic Selection Process (see the full survey in the appendix).

In the general questions section we collected information such as the name, age, gender and general statistics about the running of the school. We use the first year in post to identify new and legacy principals. The motivation for the role and management activity module of the survey enquires about the beliefs of the principal regarding their suitability for the role, including ethical concerns, service to the public, their leadership style and management of working relationships. Questions in this portion of the survey are seven-dimension Likert-scales as in Perry (1996), with equal positive and negative worded items and a neutral term. In the evolution of management and school performance part of the survey, principals were asked to rate the performance of management practises compared to pre-policy years on a 3-point scale from “better” to “worse”. These questions provide us with suggestive evidence of evolution in management practises which may have influenced changes in student outcomes. The final section of the survey is comprised of questions about meritocratic selection. Here the focus is on the principal’s motivation to take part in the meritocratic selection policy (if they did take part), and their perceptions of the policy.

4. Empirical Strategy

4.1 The Model

We employ a Difference-in-Difference (DiD) strategy that exploits the gradual introduction of the policy in order to estimate the impact of the compulsory meritocratic selection of principals on student outcomes. We estimate the effects of two treatments. First, we compare schools with a competitively versus non-competitively selected principal; we are able to implement a staggered DiD due to the gradual filling of principal posts: a large share of schools received a competitively selected

principal in 2016, but the rest deferred the competitive selection by exam until 2017 or had no valid candidates in 2016. This is our first treatment (T1).[16] Second, for schools with competitively filled posts, we estimate the impact of having a new principal (as opposed to a Legacy Principal) on student outcomes. Legacy (i.e. incumbent) principals remain in post upon passing the competitive selection process. Hence, having a new principal as a result of the competitive selection is our second treatment (T2).

The difference-in-difference two-way fixed effects (TWFE) estimator, which is commonly used to estimate of the average treatment effect on the treated (ATT) is based on an OLS regression:

$$Y_{isct} = \alpha + \beta T_{st} + \varphi_t + \theta_s + \theta_c \cdot t + \gamma X_{icst} + \varepsilon_{icst} \quad (1)$$

Here, Y_{isct} is the exam outcome for student i, in school s, situated in county c in year t. T is either T1 or T2, two indicators which are 1 when the units are treated (some begin to be treated in 2016, others in 2017), and 0 before they are treated. θ_s and θ_c are school and county fixed effects (which can also be interacted with a linear trend) and X_{icst} are several student-level pre-treatment covariates: chosen track (theoretical and technologic, where vocational is the omitted category) and, where we have the information, full-time student.

The unbiasedness of the TWFE estimator β (commonly used in such settings to capture the average treatment effect on outcomes) has been challenged by Goodman-Bacon (2021), who showed that it is in fact the "weighted average of all possible two-group/two-period DiD estimators in the data". Importantly, the TWFE estimator is comprised of several 2x2 DiD comparisons, some of which involve comparisons between groups treated at different points in time ("timing only" 2x2s). In our context, we have the gradual introduction of principals selected based on merit (and therefore also a gradual introduction of new principals as opposed to continuing ones), with two treated groups: the schools treated in the 2016 principal selection and those treated in the 2017 selection, while a small share of schools remained untreated through the entire period. This means that the TWFE estimator is an average of the following 2x2 comparisons: between the 2016 treated group vs never-treated, between the 2017-group and never-treated, but also between 2016 and 2017 treated groups (one where the 2017 group *before it is treated* serves as control for the 2016 group, and another where the 2016 group *after it is treated* serves as control group for the 2017 group).

The last "timing-only" component is particularly problematic, because when treatment effects are different over time between early and late treated groups, this component may capture these dynamic effects. Since OLS applies variance weighting on each 2x2 component to estimate the sample TWFE DiD, the weights on all comparisons will be positive, and this means the problematic 2x2 estimate may bias the average treatment effect (and may even change the sign of the average ATT). The specific bias is sample-dependent, with components for which there is a higher treatment variation receiving a higher weight (thus changing the length of the panel and implicitly the variation in treatment for certain treated groups can alter the estimates). Baker, Larcker and Wang (2022) highlight the fact that the problems associated with dynamic treatment effects and the 2x2 component from comparing the late-treated to the early treated group (as control) ca be mitigated when never-treated groups account for a substantial portion of the sample. In our case, the never-treated group for T1 (schools with principals confirmed in post without a merit-based competition) represent around 13% of the sample, and the never-treated group for T2 (schools with legacy principals) represent 71.06% of the sample (See Appendix Table A2). Thus, we expect a reduced bias

[16] This was either because there were no candidates who applied to the post, or no candidates who passed the exam.

in estimating the TWFE impact of T2 on outcomes. Initially, we produce estimates which use only those who are never treated as a comparison group, however we provide additional results which use those 'not yet treated' in the appendix for robustness.

In section 3.2 we showed that the early treated groups are systematically different from the later treated groups, and schools with new principals are different to those with legacy principals prior the selection. We may also expect that the treatment effects are heterogeneous across groups, and also that the effects change over time, as suggested in the literature on principals' impact on student performance, which may take time to be felt. For these reasons, the naïve OLS TWFE in model (1) may be biased even if parallel trends are satisfied. Therefore, we need a reliable estimator which can exclude the problematic 2x2 comparisons from the estimation of the overall ATT.

Several recent papers propose alternative estimators that correct the static and dynamic TWFE estimator, using different methodologies, depending on the context and treatment adoption (see Baker, Larcker and Wang (2022) for an overview of the main methods developed by Callaway and Sant'Anna, 2021, Sun and Abraham, 2021, de Chaisemartin, D'Haultføeuille, 2020). We apply the estimator developed by Callaway and Sant'Anna, 2021, which essentially estimates group-time-specific treatment effects through simple 2x2 comparisons with clean control groups (either never-treated, or not-yet-treated units). Each 2x2 comparison is a valid ATT for that specific group and time period, so long as there is no anticipation and there are unconditional parallel trends. The overall ATT for the sample is obtained by aggregating the group-timing 2x2s. The advantages with the Callaway and Sant'Anna estimator are: i) it is particularly suitable for our context where units once treated remain treated for the entire period; ii) it works with repeated cross-sections, as is the case in our data; and iii) it aggregates 2x2s both for each group and for each post-treatment period, such that we can infer the heterogeneous and dynamic treatment effects. This is important because it allows us to capture the changes in the impact of merit-selected and new principals on student outcomes.

To summarise, for our main analysis, we use the estimator from Callaway and Sant'Anna (2021), based on the doubly-robust estimator from Sant'Anna and Zhao (2020), and cluster bootstrap standard errors, clustered at county level. The main outcomes of interest Y_{isct} are: the overall student score of the Baccalaureate exam (observed between 2012-2019), whether the student passed the exam and the scores in a standardised written Romanian exam. We estimate this model on our working sample of 510 schools in the main specifications.

One disadvantage with this approach is that the standard errors are less efficient than the OLS standard errors (Wooldridge, 2021). For comparison, we also report the OLS regression TWFE estimates from model (1) in the appendix. We run the Goodman-Bacon (2021) decomposition (see appendix table A4), which shows that our TWFE estimator is in large part based on the outcome differences between the treated schools and the never-treated schools (90% weight). [17] This indicates that the bias in the classic DiD estimator is likely to be small, despite the problematic 2x2 comparisons between late and early treated groups.

We are also interested in whether the estimates differ along the distribution of pre-treatment school performance. For this heterogeneity analysis, we present the Calloway and Sant'Anna (2021) estimates, and for simplicity we also estimate variants of model (1) for different percentile groups

[17] Due to the requirement to have a strongly balanced panel, we retain only the schools that have Baccalaureate data for all eight years of the study time period, and we conduct the analysis on school-level aggregate scores and characteristics, weighted by the number of students in 2019.

(bottom, middle, and top) in the school average exam score distribution in 2012-2015 (see section 5.2).

Our setting is susceptible to selection bias, owing to the possibility that principals of different abilities selected the wave of the competition or the schools for which to compete. The concern is that we may have seen better performance in the schools with exam-selected principals or new principals, even in the absence of the policy. Table 1 indicates that pre-policy student performance is indeed higher in the exam-selected treatment group, and in the new principals (compared to legacy) group. The difference-in-difference estimates automatically account for the pre-policy gap in school performance. In the next section we examine the outcome trends before the policy, to rule out differential trajectories of the treatment groups before the competition was introduced. In terms of selection on unobserved characteristics, we also report OLS TWFE estimates controlling for school fixed effects and country trends.

5. Results

5.1 The Effect of a Principal Being Competitively Selected on Student Outcomes

First, we examine graphically the evolution of student outcomes over the study period in schools with a principal that was meritocratically selected in either wave of the policy, compared to schools where the principal was appointed (Figure 2). We see a very slight trend difference in overall Baccalaureate exam scores post-2016 based on whether the principal sat the exam in 2016 or 2017, however schools with a competitively selected principal in 2016 have higher student outcomes across the whole study period. This signals that there was selection of motivated principals into higher performing schools (see section 4.2); despite this, Figure 2 also displays fairly parallel trends between the two groups pre-policy.

We also examine graphically the group-period ATTs estimates obtained using Callaway and Sant'Anna (2021) estimators, without conditioning for pre-treatment covariates, for each outcome and for each group (2016 and 2017 competition, respectively). In Figure 3, the top row figures display the ATTs for the 2016 competition group, while the bottom row figures display ATTs for the 2017 group. All pre-treatment ATTs are insignificantly different from zero. Most of the pre-treatment ATTs are close to zero or negative. The post-treatment ATTs for the 2016 competition group are close to zero in the first period after treatment, and become positive and larger in the next periods. The post-treatment ATTs for the 2017 competition group are close to zero or slightly negative. None of these post-treatment 2x2 DiD estimates are significantly different from zero.

We confirm these results in Table 2 in which we present the Calloway and Sant'Anna (2021) overall sample average ATT estimate (in column 1), the average ATT estimates for the 2016 treatment group and the 2017 treatment group (columns 2 and 3) and the average ATT estimates for each post-treatment period 2017-2019 (columns 4-6). In Panel A we display the estimates for the overall Baccalaureate score, Panel B displays estimates for the probability of passing the exam, and Panel C displays the estimates for the written Romanian exam. The overall sample ATTs are insignificant, and all group and period ATTs are insignificant. The period ATTs become larger the more time has passed since the competitive selection of a principal for the school, but never exceed 0.03 of a SD for the overall Baccalaureate score and pass, or 0.07 for the Romanian exam score. Thus the effects are precisely estimated, albeit fairly small. The formal tests of the unconditional parallel trend assumption confirm that we have parallel trends in terms of two outcomes: the probability of passing the exam and the Romanian exam score. We can reject the null hypothesis of parallel trends for the overall

Baccalaureate score; however, this is mainly driven by the outcomes in the first period – the year 2012; all subsequent periods have ATTs insignificantly different from zero, as Figure 3 shows.

In Appendix Table A3, we show the Goodman-Bacon (2021) decomposition for treatment timing groups in our estimation specification. The Goodman-Bacon (2021) decomposition provides a full breakdown of the weighting of the four comparisons made between groups when using a TWFE DiD estimator. We denote the weights on the comparison of: our early treated group (2016) to our untreated group (no exam) as "W never vs 2016"; our late treated group (2017) to our untreated group as "W never vs 2017"; our early treated group to our late treated group pre-policy as "W treatment timings"; and our late treated group to our early treated group post-policy as "W within". As shown in Appendix Table A3 (column 1), only 2.7-3.4% of the main ATT components come from the problematic comparison. We expect the bias in the OLS TWFE estimate to be small, owing to the fact that its main ATT components that make up most of the effect (around 90%) are coming from comparing the treated groups with the never treated group, or with the not yet treated group. Therefore, in Appendix Table A4 we present the OLS results from model (1) on T1, whether a principal was competitively selected, for our full and working samples.[18] In columns (1) and (3) we present results with year fixed effects (FE) with 2016 as a reference, school FE and our student-level controls (track dummies and full-time student dummy); in column (2) and (4) we present results with the same FE and include county specific trends. Consistent with the results in Table 2, we do not find any significant difference in any of the three outcomes in our full sample, and no significant differences in passing probability and Romanian exam score in the working sample. We only find a small effect on the overall Baccalaureate score (around 0.01 of one SD of the no exam pre-policy mean) that is significant at 10 percent significance level in the working sample.[19]

Overall, we estimate a null average treatment effect of the policy of assigning a school a competitively selected principal on student outcomes up to three years after the policy, given the small magnitude of the coefficients and the narrow confidence intervals.

5.2 The Effect of a New Competitively-Selected Principal on Student Outcomes

The second treatment we examine is having a New Principal, compared to having a Legacy Principal conditional on the principal being confirmed in post through merit-based selection. It is more likely that new principals change school policies or processes than legacy principals.[20] However, the proportion of principals who remain in their post following the exam is larger than the proportion of new principals, which means that the overall effect of the exam may conceal the difference in impacts between existing and new hires. Figures 4.a and 4.b display the overall Baccalaureate score trends in schools of new and legacy principals who were competitively selected in 2016 and 2017 respectively; the proportion of new principals in the 2016 exam session was 24%, compared with 48.4% in 2017 (Appendix Table A2). Both figures show little difference in pre-policy trends of student outcomes for treatment and control schools. Figure 5 shows graphs of the Calloway and Sant'Anna (2021) estimated ATTs for each group and period, suggesting that most pre-policy ATTs are insignificant (parallel trends),

[18] For T1 (schools with competitively selected vs non-competitively appointed principals), we are able to compare DiD estimates in the full and reduced samples, since we know the treatment status for all schools. For T2 (schools with new principals vs legacy principals), we only have the treatment status data for the reduced sample of counties (see section 4). We use the T1 effects comparison as supporting evidence for the representativity of the reduced sample.

[19] We display estimates of T1 on our survey sample of schools in appendix table A5. We find no significant differences between competitively selected and appointed principals in any student outcome metric.

[20] For more information on the changes to school policies which Romanian principals can make, see section (2.1).

except for the third period pre-treatment ATT for the overall Baccalaureate score, which is insignificant, but positive and a bit larger than the rest of the pre-treatment ATTs.

The post-treatment ATTs are insignificant in the first period post-treatment, but become positive and larger in the second and third periods, and some are significant (for example, for pass probability, for the 2016 group).

In Table 3, following the same format as in Table 2, we present the Calloway and Sant'Anna (2021) overall sample average DiD ATT estimate for T2 (new compared to Legacy Principal schools, in column 1) and again the average ATT estimates for the 2016 treatment group and the 2017 treatment group (columns 2 and 3) and the average ATT estimates for each post-treatment period 2017-2019 (columns 4-6). Note that the sample is now restricted to include only schools which had a principal selected through the competition. The overall sample ATTs for the Baccalaureate score and Romanian exam are insignificant, but the overall ATT for the pass probability is 2 percentage points and statistically significant at 10% significance level. The results differ slightly for the 2016 and 2017 groups, particularly in the Romanian exam score, where the 2016 group ATT is three times larger, albeit still insignificant. The overall sample ATT of having a new principal suggests an increase in overall Baccalaureate score by 0.035 of a SD (a 1.3% increase on the average score in legacy schools' in pre-policy years) on the Legacy Principal schools in the pre-policy period. The effect on the pass probability is larger, amounting to 0.04 of a SD (3.6% increase on the legacy schools mean pass rate before the policy).

One interesting result is that we see dynamic effects. The period ATTs of having a new principal become larger the more time has passed since the competitive selection of a principal for the school, and are statistically significant in the last period at 10% significance level for the Baccalaureate exam score and for the Romanian score, and significant at 5% for the pass probability. The last period ATTs are double the magnitude of the overall sample ATTs (an increase by 0.07 of a standard deviation increase for the overall Baccalaureate score, pass rates that are around 8 percent larger than in the control group, and an increase by 0.08 of a SD for Romanian exam scores). This indicates that new principals take some time to produce effects in terms of student outcomes, which is in line with the literature (e.g., Coelli and Green, 2011).

Parallel pre-treatment trends tests confirm the assumption holds for the probability of passing the exam and the Romanian exam score, but not for the overall Baccalaureate score. Nevertheless, we see consistently significant and higher treatment effects for all outcomes in the last period. In Appendix Table A6, we present estimates of our OLS DiD specification using T2, whether a principal is a new or Legacy Principal. Overall, we find similar or slightly smaller magnitudes to the overall ATT estimates in Table 3, none of which are statistically significant.

Having found null contemporaneous and a small, delayed effects, we explore the potential channels by which competitive principals may have made a difference on student outcomes. We examine the differences between new and legacy principals in the next section.

5.3 Mechanisms

We examine the differences between management styles and motivations of new and legacy principals in our dataset of surveyed schools; for detailed information on our telephone survey see section 3.3. Of the 277 surveyed schools which had an exam selected principal, 206 schools had a Legacy Principal whilst 71 had a New Principal. One caveat of our survey is the absence of one particular reference category: former principals who were replaced by a new principal but who were not competitively selected for another post. However, we are able to compare new principals to

competitively selected legacy principals, who likely share some similar traits by virtue of being instated through the policy and managing similar schools. We analyse the difference in survey responses between competitive new and legacy principals in our surveyed schools (Table 4).

First, we consider differences in responses to questions related to the motivation for the management role (adapted from Perry's 1996 motivation for civil service survey). These range from beliefs about the importance of ethical behaviour, self-righteousness, and own interpersonal skill, measured on a 7-point Likert scale (e.g. "The ethical behaviour of school directors is as important to me as their competencies"; see the complete questionnaire in the Appendix). On average, new principals are more supportive of peers (more willing to "fight for the rights of others"; $p<0.1$), are typically less rude to others ($p<0.05$), and see financial success as less important ($p<0.05$). These traits might suggest that new principals are more service-minded and more likely to foster an environment of support and cooperation between teachers and management; in particular, their lesser concern with financial success suggests they are more prosocial and more motivated to fulfil their duties.

Despite these findings, we also find suggestive evidence that legacy principals are more likely to believe that teachers in their school trust each other ($p<0.1$). In addition, we find some differences between new and legacy principals which may be imprecisely estimated due to a small sample size: legacy principals are more likely to believe that teachers trust management, share common values with staff and trust teachers enough to delegate authority more often than it is the case in schools with new principals. Note that one of our main findings highlights the time it takes for new principals to have an impact on outcomes. Therefore, it is unsurprising that new principals perceive trust between teachers and management to be weaker than legacy principals do, as productive relationships within the organisation may take time to build.

We also analyse the differences in responses to questions about the (perceived or observed) change in management practises and school performance indicators since taking on the role. These questions ask whether the principal believes that specific metrics have declined, stayed the same or improved in 2017-2018 compared to previous years (on a scale 1-3, where 1 is decreased and 3 is improved). These metrics are: adapting learning to student needs (targeted teaching), tracking teacher performance, monitoring objectives, teacher absence, hiring skilled staff, use of sanctions for staff, motivating staff, mobilizing staff, the level of authority over staff, absence amongst students, student dropout rate, graduation rate, baccalaureate enrolment rate. Table 4 shows that new principals report significantly less improvement in adapting learning to student needs ($p<0.1$), but that there are no other significant differences between new and legacy principals, neither in terms of teacher or student absenteeism, nor on monitoring and motivating staff. Some of the similarities between new and legacy principals' schools in terms of management practises (especially on motivating staff) is unsurprising, given that school principals have little autonomy in decision making; most key decisions for education quality are made at the county inspectorate or ministry level, including: remuneration, hiring and firing of teachers and principals; school budgets.[21] Overall there is little evidence that new principals implemented superior management practices to those of legacy principals (which is not surprising in the very short run), but they have potentially facilitated a more collaborative environment.

We can go one step further and examine whether the objective Baccalaureate enrolment rates have changed differentially in schools with new compared to legacy principals. Principals can decide strategically whether students at the margin can graduate and sit the exam. Schools are ranked

[21] The main limitation of the survey is its susceptibility to social desirability bias, which is a concern if it is displayed differently by new and legacy principals (however, in that case, we would expect new principals to score higher on more measures than the select few that we see).

by the passing rates in the first sit, and improvements in the passing rates are monitored on an annual basis. Preventing students predicted to fail the exam from taking the first sit (leaving them the option of a taking sitting the August reassessment) improves the passing rates and test scores, and therefore the objective measures of school performance. To test for this, we need data on the Baccalaureate enrolment rate, i.e. the share of students who sit the Baccalaureate of the students who are enrolled in the 12th grade in each school. For our measure of the number of students enrolled in the 12th grade, we use high school admissions data from Munteanu (2021).[22] The data contains the number of admitted students per each track, but excludes students who were in vocational tracks (e.g. theology, pedagogy etc.). As a result, for some schools which do not offer a vocation track we have the total number of students admitted to the high school four years prior the Baccalaureate. However, for schools which also have vocational tracks, we only have the number for the subset of the students admitted to their technical and theoretical tracks. We create the Baccalaureate enrolment rate as the ratio of the number of theoretical or technical track students who sat the Baccalaureate (corrected by the number of returning students) to the number of students admitted to that particular school and track four year prior. Thus, we deal with the missing data by excluding vocational track students from the sample. The disadvantage of restricting the observations in this way is that we lose all vocational schools (typically lower performance schools). We re-estimate the results for the sample of technical and theoretical schools.

The results in Table 5 show that the new principal DiD estimates on the Baccalaureate enrolment rate from the restricted sample are negative and significant, both in terms of the overall ATT, implying an overall reduction in Baccalaureate enrolment rates by 8.4 percentage points (p-value 0.02), and in terms of the 2018 and 2019 outcomes (the drop is 13.6 percentage points the 2018). The estimates suggest that one of the levers new principals use is the permission for students to enrol in the Baccalaureate exam. This has the effect that school Baccalaureate performance measures mechanically improve. The effects on students' learning and exam performance are ambiguous: in theory students have more time to revise and a better chance at passing the exam in the second sit, but in practice there is a risk that they drop out and forego the chance of a high school diploma.

Thus, we provide evidence that the new competitive principals may influence Baccalaureate exam scores through a channel which artificially improves school performance. This is because principals have little autonomy in changing fundamental inputs in education, such as recruiting and incentivising teachers, curriculum and pedagogical innovation. This raises concerns that centralisation may curtail the potential of sound recruitment policies to make an impact on learning and reduce education inequalities, despite the fact that they raise the quality of school principals. Nevertheless, we cannot rule out that the new principals also brought about real improvements in students' performance, given the evidence that new principals are more motivated to fulfil their public mission.

5.4 Heterogeneity Analysis in School's Pre-Policy Performance

In order to analyse the differential impact of new and legacy principals depending on past performance of schools, we split our working sample into three groups on the full sample distribution of school average final exam score over the 2012-2016 period: below the 25th percentile (<25th); between the 25th and 75th percentile (25th-75th); above the 75th percentile (>75th) in school average exam scores. We divide the distribution in this way in order to capture the bottom and top performing schools in distinct groups from the midrange.

[22] Obtained from Diana Coman's repository of Romanian public education data (www.ossasepia.com).

In columns (1-3) of Table 6 we present the Calloway and Sant'Anna (2021) average DiD ATT estimates for T2 (new compared to Legacy Principal schools) for schools with pre-policy performance under the 25th percentile. In columns (4-6) we present results for the middle part of the distribution, and in columns (7-9) we present ATT estimates for schools with pre-policy performance above the 75th percentile. We report results assuming unconditional parallel trends, We also report the ATT estimates for the 2016 treatment group and the 2017 treatment group (columns 2-3/5-6/8-9 respectively). As with the results in Table 3, the sample is now restricted to include only schools which had a principal selected through the competition.

We find negative but insignificant overall ATT's of new principals on student outcomes for schools in the group below the 25th percentile. The new principal effects are positive for the 2017 policy wave for this group, modest in magnitude for the overall exam score and large for the Romanian exam score (the estimate for the Romanian exam score is 0.31, or 7.6% increase compared to the mean score in the legacy principals' schools before 2017, and the p-value is 0.157). Within the 25th-75th percentile group, the overall ATT estimates are positive, modest in magnitude and insignificant. In this group, we find significant positive ATTs for the 2017 policy wave, for the Baccalaureate exam score (estimate 0.168, equivalent to a 3.2% increase on the control group mean score before 2017; p-value of 0.067) and for pass rates (estimate 0.051, an 11.3% increase compared to the pass rate of 0.451 in legacy principals' schools before 2017; p-value of 0.029);[23] By contrast, we do not find a significant impact of new principals in schools belonging to the group above the 75th percentile. Indeed, many of these estimates are close to 0.

The caveat in reading these results is that when splitting the sample in this way, we cannot guarantee parallel pre-policy trends in each group. For the low performing schools, which sit below the 25th percentile, as well as for two outcomes for schools in the 25th -75th percentile, we could reject the null hypothesis of parallel pre-trends (both unconditional and conditional). For this reason, the estimates should be read with caution. We can reliably conclude that there were no effects of new principals on top performing schools. However, there is some evidence that middle-performing schools in the 2017 policy wave benefitted significantly from having new principals, at least in terms of pass rates (for which the parallel trend assumption holds), and possibly in terms of the other outcomes. The results for the low-performing schools are inconclusive, and cannot be reconciled with the OLS TWFE estimates in Appendix Table A7, which display positive and significant estimates for this low performing group.

Given that we have found an overall positive treatment effect emerges two years after the policy, these results make it plausible to suspect that the policy shifted the left side of the performance distribution up, and that the policy has helped to reduce the performance gap between students in low-to-middle and top schools. We also examine the impact of the overall policy (comparing competitively and non-competitively selected principals) on students in different regions of the school performance distribution, and we display the results in Appendix Table A8. We find no impact of the overall policy on any student outcome in any percentile group, which suggests that the principal exam policy only worked insofar as it enabled a change in principals in some of the lower performing schools.

[23] For a complete set of reference group student average outcomes, pre- and post-policy, please see Appendix Table A3.

5.5 Robustness Checks

In order to verify that our main results are econometrically robust we undertake a series of robustness checks including using the not yet treated as an alternative control group, and running the traditional TWFE OLS models with placebo policies.

We estimate the Calloway and Sant'Anna ATT, for the overall sample, and by timing groups and dynamic effects using the not yet treated units as a control group. This means that units never treated and units who were only treated in 2017 become the control group for units treated in 2016. For the units treated in 2017, the never treated units remain the control group. The results displayed in appendix Tables A9 and A10 (for T1 – exam and T2- new principals) are very similar to the main results in Tables 2 and 3.

In sections 5.1. and 5.2 we showed that the TWFE OLS models show results broadly consistent with the Calloway and Sant'Anna ATT. The advantage of OLS is that we can control for time-varying and invariant characteristics of schools and students. One additional check to understand whether the OLS estimates reflect a trend that predated the policy is a placebo test where we artificially instate the policy one year earlier (2015 replacing 2016, and 2016 replacing 2017 selection wave). We would expect to see insignificant coefficients of the placebo policy. Appendix Table A11 shows no significant result from our placebo tests in any student outcome metric in the overall sample. There is one significant estimate of the placebo policy on overall Baccalaureate scores for schools which received a New Principal, relative to those who retained legacy principals in 2016, but the effect is very small and significant only at 10%. This is further supporting evidence that our main results are not confounded by diverging trends started just before the policy, which partly alleviates the concern that new principals selected schools based on their prior performance trajectory.

Finally, for the OLS TWFE estimator, we turn our attention to the standard errors. Our main estimation includes 19 county clusters, which is lower than a minimum standard number of clusters required to estimate unbiased standard errors. A small number of clusters with a high degree of inter-cluster correlation, as is likely to be the case within Romanian counties, can lead to an underestimation of standard errors and falsely rejecting a true zero effect (Cameron *et al.* 2008). We therefore run alternative regressions (examining T2 – the impact of new principals) with the wild bootstrap correction for the standard errors. We report these in Appendix Tables A12. The wild bootstrap correction of the standard errors does not compute new coefficients, however Table A12 displays new t- and p-values. We find no significant results when estimating bootstrap standard errors; this may be expected as our earlier results suggest that new principals only have an impact in the long-term, but our bootstrap standard errors are calculated based on a TWFE estimate which does not include a breakdown of ATTs (as in Callaway and Sant'Anna, 2021).

6. Conclusions

This paper aims to understand whether meritocratic selection of state school principals can improve student outcomes. We test whether (1) having a competitively selected principal, and (2) having a new principal selected based on merit, improves a range of student outcome metrics: overall Baccalaureate scores; pass rates; and written Romanian scores.

Specifically, we study the effect of the Romanian meritocratic selection policy, which was introduced in two waves (2016 and 2017 independently), on outcomes from 2012-2019. We utilise administrative data to analyse effects of principals who became new principals in the school,

compared to those who remain principal of the same school pre- and post-policy. Further, we examine survey responses of principals and additional outcomes to understand the possible mechanisms through which change is enacted.

At first glance, we do not find a significant impact of the overall policy; having an exam selected principal does not inherently improve student outcomes in the short run. However, we do find some evidence that those competitive principals who were new in post had a positive impact on test scores. In line with the literature, our results suggest that it takes time for new principals to have an impact on outcomes since the ATT continues to increase in the years post-policy, and this effect is driven by schools with low or average historical performance.

Survey evidence suggests that new principals are more prosocial individuals and are more motivated to fulfil their public mission through their work, but otherwise similar to legacy principals in terms of managerial practices. Since the survey was held immediately after the competitive selection of principals, it was unlikely that changes would have been introduced or taken effect. More problematically, owing to the lack of autonomy in decision making, there remains a possibility that principals would game the system to mechanically enhance performance indicators. One of the few margins which principals can influence in our context is the student's enrolment into the Baccalaureate exam. Our results suggest showing a lower enrolment rate in new principals' schools is consistent with this mechanism. By imposing that low performing students defer the first sit, principals help schools achieve higher passing rates (and possibly also better subsequent outcomes for the students).

There is scope for future research to evaluate policies which combine meritocratic selection with more discretionary powers for principals to shape in-class instruction, as well as management training focused on enhancing teacher engagement with effective pedagogical practices, such as targeted teaching (which has proven effective in primary education, see e.g. Banerjee et al, 2016, Beg, Fitzpatrick and Lucas, 2019) or remedial education (see e.g. Banerjee et al, 2007). Further research avenues include studying the longer term impact of merit-based selection of principals on student outcomes, which was not possible in this study due to the systematic changes to the education process following the COVID-19 pandemic.

On balance, the meritocratic selection process allowed more motivated candidates to replace complacent ones and signal that high management quality is an important government objective. Despite this, there were relatively fewer new principals in low-performing schools compared to the rest. Therefore, merit-based selection combined with incentives for new principals to consider posts in low-performing schools, and policies to instil effective management practices are promising avenues to improving student outcomes and bridging the performance gap between high and poor performing schools. More broadly, our paper provides robust evidence from a sudden and widespread policy change that merit-based selection of public sector managers can bring renewed motivation to public service delivery, but is not free of the risk that managers explore bureaucratic loopholes when constrained in their decisions.

Bibliography

Agasisti, T; Bowers, A; Soncin, M. (2018). *School Principals' Leadership Types and Student Achievement in the Italian Context: Empirical Results from a Three-Step Latent Class Analysis.* Educational Management Administration and Leadership, Vol. 47, No. 6, Pp 860-886.

Banerjee, A., Banerji, R. Berry, J. , Duflo, E., Kannan, E., Mukherji, E., Shotland, M. and Walton, M. (2016, October). *Mainstreaming an effective intervention: Evidence from randomized evaluations of teaching at the right level in India*. Working Paper 22746, National Bureau of Economic Research.

Banerjee, A., Cole, S., Duflo, E. and Linden, L. (2007). *Remedying education: Evidence from two randomized experiments in India.* The Quarterly Journal of Economics.

Baker, A; Larcker, D; Wang, C. (2022). *How Much Should we Trust Staggered Difference-in-Differences Estimates?*. Journal of Financial Economics, Vol. 114, No. 2, Pp 370-395.

Beg, S., Fitzpatrick, A., and Lucas, A. M. (2019). *The Importance of Management Support for TeacherLed Targeted Instruction in Ghana,* Innovations for Poverty Action Brief.

Bertrand, M; Schoar, A. (2003). *Managing with style: The effects of managers on firm policies*. The Quarterly Journal of Economics, Vol. 118, No. 4, Pp 1169-1208.

Blase, J. (1988). *The politics of favouritism: A qualitative analysis of the teachers' perspective*. Educational Administration Quarterly, Vol. 24, No. 2, Pp 152-177.

Bloom, N; Lemos, R; Sadun, R; Van Reenen, J. (2015). *Does management matter in schools?*. The Economic Journal, Vol. 125, Pp 647-674.

Bloom, N; Van Reenen, J. (2007). *Measuring and explaining management practises across firms and countries*. The Quarterly Journal of Economics, Vol. 122, No. 4, Pp 1351-1408.

Bohlmark, A; Gronqvist, E; Vlachos, J. (2016). *The headmaster ritual: The importance of management for school outcomes*. The Scandinavian Journal of Economics, Vol. 118, No. 4, Pp 912-940.

Bohlmark, A; Lindahl, M. (2007). *The impact of school choice on pupil achievement, segregation and costs: Swedish evidence*. IZA Discussion Papers, No. 2786.

Borcan, O; Lindahl, M; Mitrut, A. (2017). *Fighting Corruption in Education: What Works and Who Benefits?* American Economic Journal: Economic Policy, Vol. 9, No. 1, Pp 180-209.

Branch, G; Hanushek, E; Rivkin, S. (2012). *Estimating the effect of leaders on public sector productivity: The case of school principals*. NBER Working Paper No. 17803.

Bruno, P; Strunk, K. (2019). *Making the Cut: The Effectiveness of Teacher Screening and Hiring in the Los Angeles Unified School District*. Educational Evaluation and Policy Analysis, Vol. 41, No. 4, Pp 426-460.

Callaway, B; Sant'Anna, P. (2021). *Difference-in-Differences with Multiple Time Periods*. Journal of Econometrics, Vol. 225, No. 2, Pp 200-230.

Cameron, C; Gelbach, J; Miller, D. (2008). *Bootstrap-Based Improvements for Inference with Clustered Standard Errors*. Review of Economics and Statistics, Vol. 90, No. 3, Pp 414-427.

Charron, N; Dahlstrom, C; Fazekas, M; Lapuente, V. (2017). *Careers, connections and corruption risks: Investigating the impact of bureaucratic meritocracy on public procurement processes*. The Journal of Publics, Vol. 79, No. 1, Pp 89-104.

Cilliers, J., Dunford, E. and Habyarimana, J. (2022). *What Do Local Government Education Managers Do to Boost Learning Outcomes?,* The World Bank Economic Review, Vol. 36, Issue 3, August 2022, Pp 629–645.

Coelli, M; Green, D. (2011). *Leadership effects: School principals and student outcomes*. Economics of Education Review, Vol. 31, No. 1, Pp 92-109.

Colonnelli, E; Teso, E; Prem, M. (2018). *Patronage in the allocation of public sector jobs*. Job Market Paper available at: <https://www.dropbox.com/s/r3qaegqfreorcfb/EdoardoTeso_JMP.pdf?dl=0>

Cortazar, J; Fuenzalida, J; Lafuente, M. (2016). *Merit Based Selection of Public Managers: Better Public Sector Performance*. Inter-American Development Bank.

Dahis, R; Schiavon, L; Scot, T. (2020). *Selecting Top Bureaucrats: Admission Exams and Performance in Brazil*. Working Paper, SSRN: 3584725.

De Chaisemartin, C; D' Haultføeuille, X. (2020). *Two-Way Fixed Effects Estimators with Heterogenous Treatment Effects*. American Economic Review, Vol. 110, No. 9, Pp 2964-2996.

Dhuey, E; Smith, J. (2014). *How important are school principals in the production of student achievement?* Canadian Journal of Economics, Vol. 47, No. 2, Pp 635-662.

Di Liberto, A; Schivardi, F; Sulis, G. (2015). *Managerial practises and student performance*. Economic Policy, Vol. 30, No. 84, Pp 683-728.

Evens, P; Rauch, J. (1999). *Bureaucracy and Growth: A Cross-National Analysis of the Effect of Weberian State Structure on Economic Growth*. American Sociological Review, Vol. 64, Pp 748-765.

Estrada, R. (2019). *Rules Versus Discretion in Public Service: Teacher Hiring in Mexico*. Journal of Labor Economics, Vol. 37, No. 2, Pp 545-579.

Faucette, N; Graham, G. (1986). *The impact of principals on teachers during in-service education: A qualitative analysis*. Journal of Teaching in Physical Education, Vol. 5, No. 2, Pp 79-90.

Goodman-Bacon, A. (2021). *Difference-in-Differences with Variation in Treatment Timing*. Journal of Econometrics, Vol. 225, No. 2, Pp 254-277.

Gurmu, T. (2020). *Primary School Principals in Ethiopia: Selection and Preparation*. Educational Management Administration and Leadership, Vol. 48, No. 4, Pp 651-681.

Hoffman, M; Kahn, L; Li, D. (2018). *Discretion in Hiring*. The Quarterly Journal of Economics, Vol. 133, No. 2, Pp 765-800.

Hsiao, H; Lee, M; Tu, Y. (2012). *The Effects of Reform in Principal Selection on Leadership Behaviour of General and Vocational High School Principals in Taiwan*. Educational Administration Quarterly, Vol. 49, No. 3, Pp 421-450.

Huseyin, A; Mustafa, T. (2008). *Nepotism, favouritism and cronyism: A study of their effects on job stress and job satisfaction in the banking industry of north Cyprus*. Social Behaviour and Personality: An International Journal, Vol. 36, No. 9.

Jacob, B. (2010). *Do Principals Fire the Worst Teachers?*. Educational Evaluation and Policy Analysis, Vol. 33, No. 4, Pp 403-434.

Jacob, B; Rockoff, J; Taylor, E; Lindy, B; Rosen, R. (2018). *Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools*. Journal of Public Economics, Vol. 116, Pp 81-97.

Loeb, S; Kalogrides, D; Horng, E. (2010). *Principal Preferences and the Uneven Distribution of Principals Across Schools*. Educational Evaluation and Policy Analysis, Vol. 32, No. 2, Pp 205-229.

Masci, C; De Witte, K; Agasisti, T. (2018). *The Influence of School Size, Principal Characteristics, and School Management Practises on Educational Performance: An Efficiency Analysis of Italian Students Attending Middle Schools*. Socio-Economic Planning Sciences, Vol. 61, Pp 52-69.

Meyer, A; Richter, D; Hartung-Beck, V. (2020). *The Relationship Between Principal Leadership and Teacher Collaboration: Investigating the Mediating Effect of Teachers' Collective Efficacy*. Educational Management Administration and Leadership.

OECD. (2014). *PISA 2012 Results: What Makes a School Successful?* Resources, Policies, and Practises, Vol. 4.

OECD. (2020). *Education at a Glance 2020*. Available at: https://doi.org/10.1787/69096873-en

Palmer, B; Mullooly, J. (2015). *Principal Selection and School District Hiring Cultures: Fair or Foul?* Journal of Education and Social Policy, Vol. 2, No. 2, Pp 26-37.

PEIS. (2007). *The Survey on Higher Education Policies*. Open Society Foundation; Gallup Romania.

Perry, J. (1996). *Measuring public service motivation: An assessment of construct reliability and validity*. Journal of Public Administration Research and Theory, Vol. 6, No. 1, Pp 5-22.

Poocharoen, O; Brillantes, A. (2013). *Meritocracy in Asia Pacific: Status, issues and challenges*. Review of Public Personnel Administration, Vol. 33, No. 2, Pp 140-163.

Rauch, J; Evans, P. (2000). *Bureaucratic structure and bureaucratic performance in less developed countries*. Journal of Public Economics, Vol. 75, No. 1, PP 49-71.

Ruiz-Tagle, C. (2019). *Selection of School Principals Based on Competitive Processes: Evidence from One Policy in Chile*. Calidad en la Educacion, No. 51, Pp 85-130.

Sant'Anna, P; Zhao, J. (2020). *Doubly Robust Difference-in-Differences Estimators*. Journal of Econometrics, Vol. 219, No. 1, Pp 101-122.

Scoppa, V. (2009). *Intergenerational transfers of public sector jobs: A shred of evidence on nepotism*. Public Choice, Vol. 141, No. 1-2, Pp 167-188.

Stravakou, P. (2019). *Selecting School Principals in Greece in the Last Fifteen Years – A Theoretical Approach*. Journal of Advances in Education and Philosophy, Vol. 3, No. 8, Pp 277-282.

Strouse, D. (2004). *A qualitative case study of the impact of principal leadership and school performance awards on eight Maryland schools*. University of Maryland Digital Repository: Teaching, Learning, Policy & Leadership Theses and Dissertations

Sun, L; Abraham, S. (2021). *Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Effects*. Journal of Econometrics, Vol. 225, No. 2, Pp 175-199.

Sundell, A. (2014). *Are formal civil service examinations the most meritocratic way to recruit civil servants? Not in all countries*. Public Administration, Vol. 92, No. 2.

Tan, K. (2008). *Meritocracy and Elitism in a global city: Ideological shifts in Singapore.* International Political Science Review, Vol. 29, No. 1, Pp 7-27.

Walker, A; Kwan, P. (2012). *Principal Selection Panels: Strategies, Preferences and Perceptions*. Journal of Educational Administration, Vol. 50, No. 2, Pp 188-205.

Wooldridge, J. (2021). *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators*. Available at: *SSRN 3906345*.

Yeboah-Assiamah, E; Asamoah, K; Osei-Kojo, A. (2014). *Corruption Here, Corruption There, Corruption Everywhere: A Framework for Understanding and Addressing Public Sector Corruption in Developing African Democracies*. Journal of Public Administration and Governance, Vol. 4, No. 3.

FIGURES AND TABLES

Figure 1. Timeline of Baccalaureate exam and competitive selection of principals

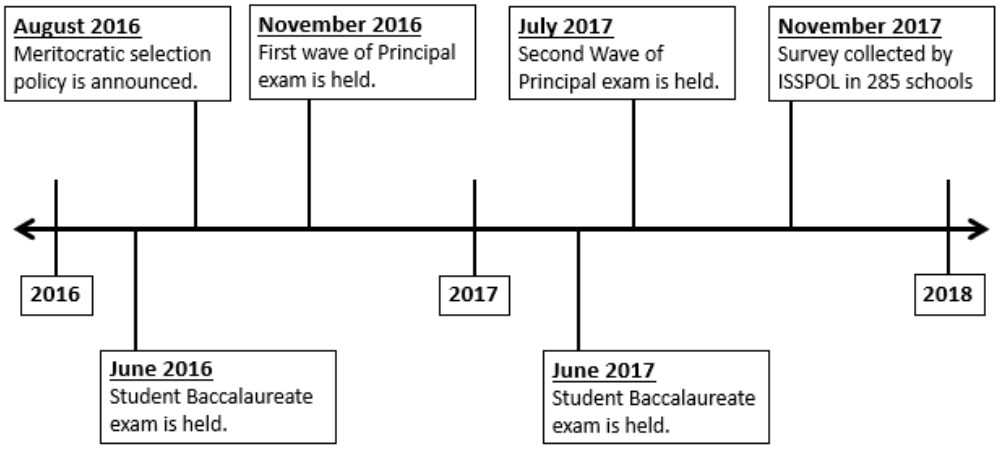
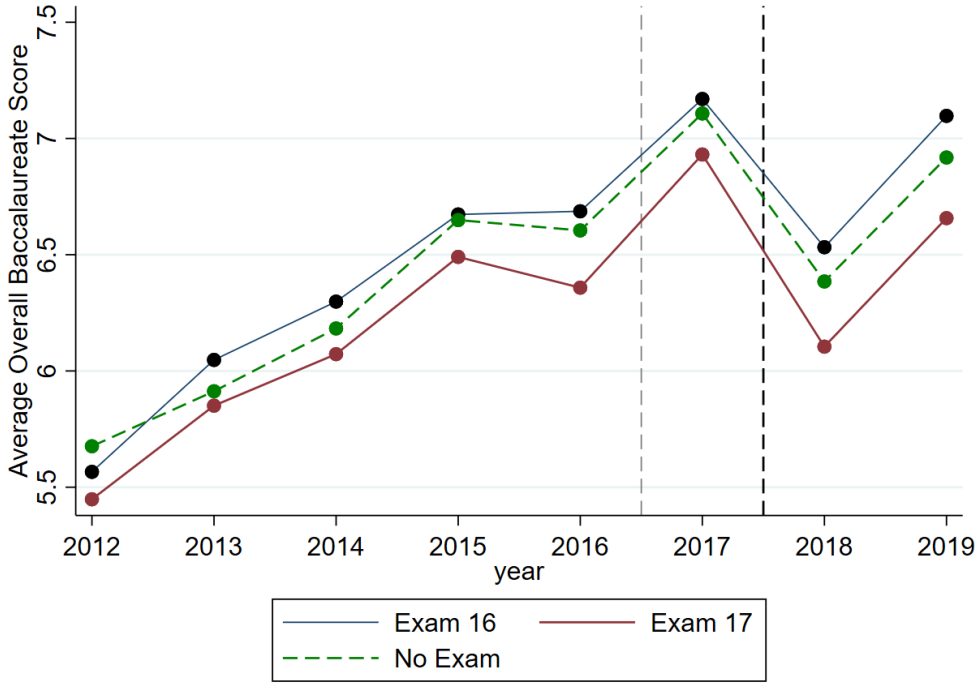
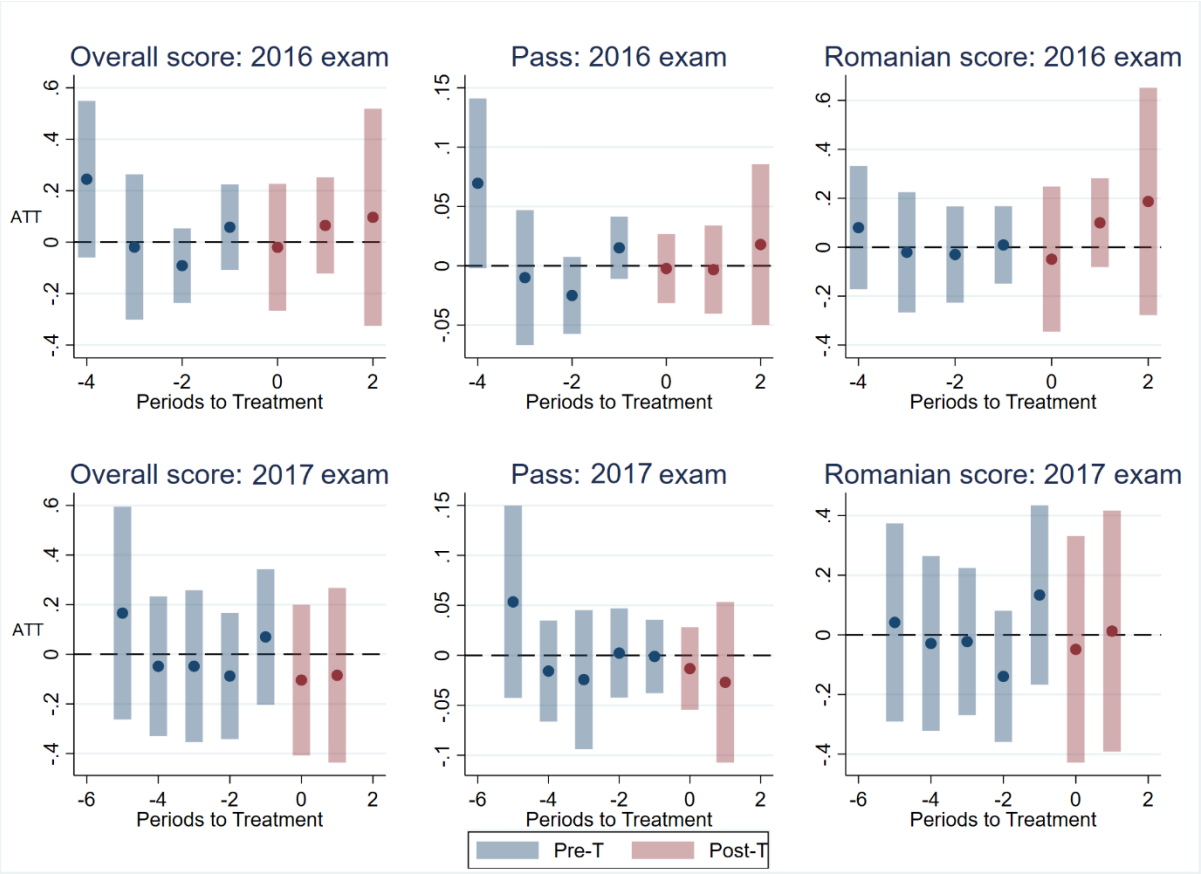


Figure 2 – Exam 2016 vs Exam 2017 (Reduced Sample)



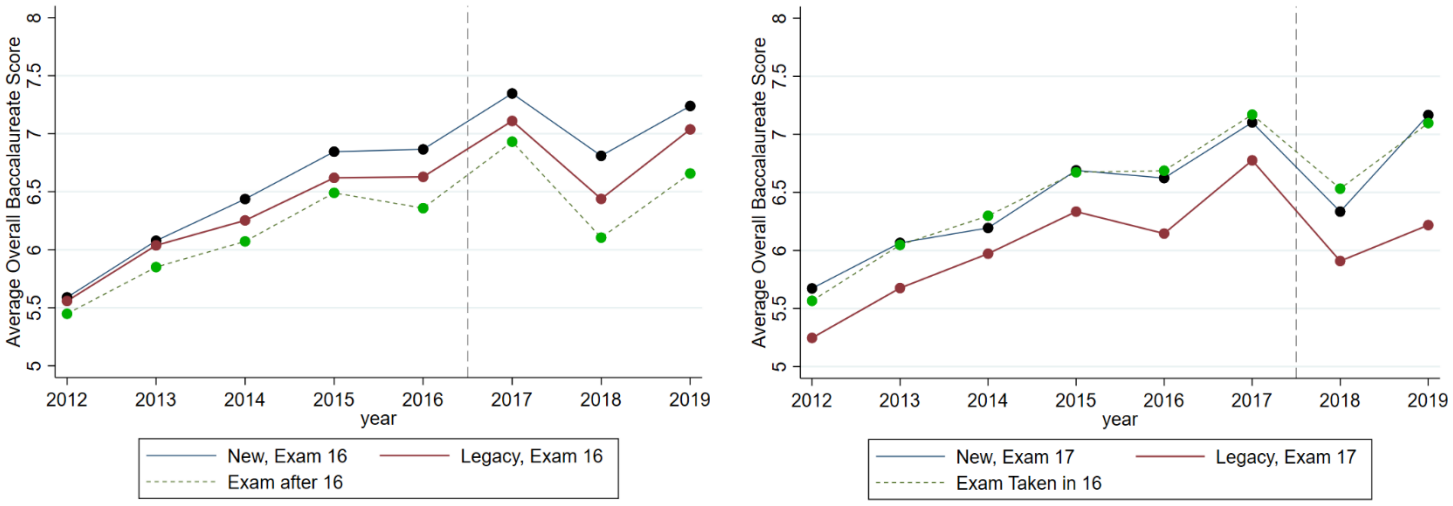
Notes: The figure shows average overall Baccalaureate scores for our reduced (working) sample between 2012-2019. We show the average scores for schools with principals selected in the 2016 exam (in blue), and schools with principals selected in the 2017 exam (in red). The light dashed vertical line demarcates the timing of the 2016 exam, whilst the dark dashed vertical line demarcates the timing of the 2017 exam.

Figure 3: DID plots (T1=principal exam) – dynamic heterogeneous effects by groups



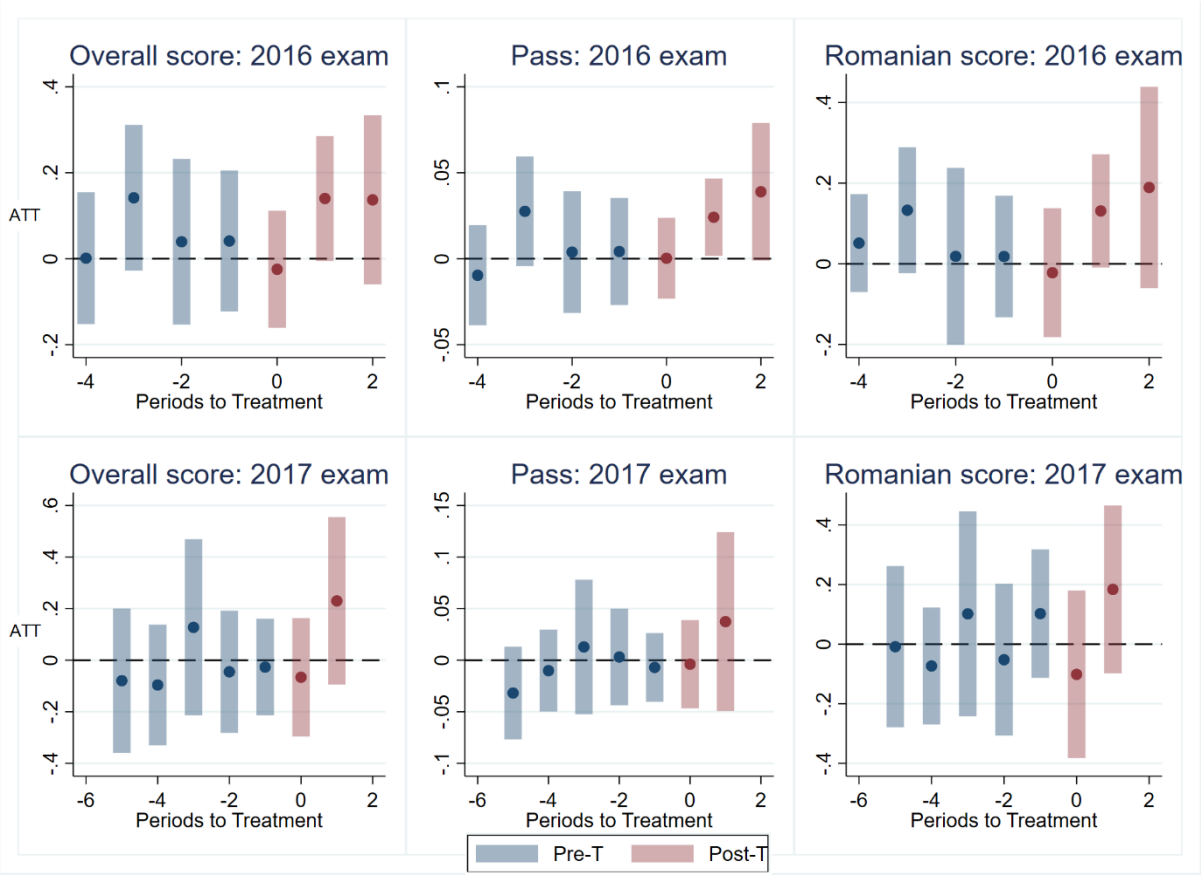
Notes: The figure shows group-period ATTs from Callaway and Sant'Anna (2021) estimates for the effect of having an exam selected principal on student outcomes. We show the overall Baccalaureate scores, pass rates, and Romanian test scores for schools with principals selected in 2016 (top figures), and those selected in 2017 (bottom figures) separately. Note that these results do not condition for pre-treatment covariates. Pre-treatment ATTs are shown in blue, whilst post-treatment ATTs are shown in red.

Figure 4.a – Exam 2016 New vs Exam 2016 Legacy vs Exam 2017 (all); Figure 4.b – Exam 2017 New vs Exam 2017 Legacy vs Exam 2016 (all)



Notes: The figure shows average overall Baccalaureate scores for our reduced (working) sample between 2012-2019. Figure 4.a shows average scores for schools with principals selected in the 2016 exam; these are displayed for new principals (in blue) and legacy principals (in red). In addition, we show average outcomes for principals selected in the 2017 exam as a reference (in green). Figure 4.b shows average scores for schools with principals selected in the 2017 exam; these are displayed for new principals (in blue) and legacy principals (in red). In addition, we show average outcomes for principals selected in the 2016 exam as a reference (in green). The dashed vertical lines demarcate the timing of the relevant exam.

Figure 5: DID plots (T2=new principal) – dynamic heterogeneous effects by groups



Notes: The figure shows group-period ATTs from Callaway and Sant'Anna (2021) estimates for the effect of having a new principal (compared to a legacy principal) on student outcomes. We show the overall Baccalaureate scores, pass rates, and Romanian test scores for schools with principals selected in 2016 (top figures), and those selected in 2017 (bottom figures) separately. Note that these results do not condition for pre-treatment covariates. Pre-treatment ATTs are shown in blue, whilst post-treatment ATTs are shown in red.

Table 1 – Student Outcome Variables by (Panel A) No Exam compared with Exam, and (Panel B) Legacy principals compared with New principals; Reduced Sample

Panel A	**(1) No Exam**	**(2) Exam**		
	Mean (sd)	Mean (sd)	Difference (1)-(2)	P Value
	(1)	(2)		
2012-2016 Overall Baccalaureate Score	6.160 (2.290)	6.178 (2.257)	-0.018	0.259
2017-2019 Overall Baccalaureate Score	6.786 (2.130)	6.851 (2.160)	-0.065	0.004***
2012-2016 Romanian Score	5.913 (2.285)	6.072 (2.376)	-0.159	0.000***
2017-2019 Romanian score	6.519 (2.141)	6.729 (2.282)	-0.21	0.000***
2012-2016 Pass rate	0.539 (0.497)	0.558 (0.498)	-0.022	0.000***
2017-2019 Pass rate	0.639 (0.473)	0.662 (0.480)	-0.023	0.000***
Panel B	**(1) Legacy Principals**	**(2) New Principal**		
	Mean (sd)	Mean (sd)	Difference (1)-(2)	P Value
	(1)	(2)		
2012-2016 Overall Baccalaureate Score	6.137 (2.318)	6.280 (2.278)	-0.143	0.000***
2017-2019 Overall Baccalaureate Score	6.767 (2.073)	7.045 (2.148)	-0.278	0.000***
2012-2016 Romanian Score	6.033 (2.319)	6.174 (2.270)	-0.141	0.000***
2017-2019 Romanian score	6.638 (2.096)	6.941 (2.153)	-0.303	0.000***
2012-2016 Pass rate	0.550 (0.494)	0.577 (0.497)	-0.027	0.000***
2017-2019 Pass rate	0.646 (0.460)	0.697 (0.478)	-0.051	0.000***

Notes: The table displays means and differences for all three student outcome metrics considered, along with a p value for a completed t-test, for schools in our reduced (working) sample. Panel A displays results for competitively selected (exam) and non-competitively selected (no exam), both pre- (2012-2016) and post- (2017-2019) policy. Panel B displays results for new principals and legacy principals, again pre- and post- policy. Standard errors at displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2 – Exam vs No Exam; ATT and dynamic effects

	Sample	Group 2016	Group 2017	ATT 2017	ATT 2018	ATT 2019
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall Baccalaureate Exam Score						
ATT (T = Principal Exam)	0.029	0.046	-0.095	-0.020	0.035	0.066
	(0.116)	(0.123)	(0.144)	(0.126)	(0.092)	(0.199)
P-Value	[0.803]	[0.710]	[0.510]	[0.874]	[0.703]	[0.740]
Observations	428,061					
Chi2 H_0: parallel pre-trends	6.542					
P-Value H_0: parallel pre-trends	[0.000***]					
Panel B: Pass rate						
ATT (T = Principal Exam)	0.001	0.004	-0.020	-0.002	-0.005	0.010
	(0.018)	(0.018)	(0.028)	(0.015)	(0.018)	(0.034)
P-Value	[0.965]	[0.836]	[0.479]	[0.876]	[0.780]	[0.764]
Observations	431,940					
Chi2 H_0: parallel pre-trends	11.962					
P-Value H_0: parallel pre-trends	[0.215]					
Panel C: Romanian written exam score						
ATT (T = Principal Exam)	0.064	0.076	-0.020	-0.049	0.073	0.157
	(0.128)	(0.135)	(0.178)	(0.151)	(0.092)	(0.223)
P-Value	[0.615]	[0.573]	[0.908]	[0.746]	[0.423]	[0.480]
Observations	429,674					
Chi2 H_0: parallel pre-trends	65.450					
P-Value H_0: parallel pre-trends	[0.685]					

Notes: The table displays ATT estimates from difference-in-difference specifications of the effect of T1 (the meritocratic selection policy) on student outcomes, using the double-robust inverse probability weighting estimator from Calloway and Sant'Anna (2021). Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present estimates for the entire working sample in column (1), ATTs by treatment groups in columns (2) and (3), and ATTs by period in columns (4)-(6). County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table 3 – New vs Legacy Principal; Reduced Sample; ATT and dynamic effects

	Sample	Group 2016	Group 2017	ATT 2017	ATT 2018	ATT 2019
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall Baccalaureate Score						
ATT (T = New Principal)	0.081	0.083	0.070	-0.025	0.081	0.162
	(0.061)	(0.068)	(0.095)	(0.069)	(0.070)	(0.093)
P-Value	[0.186]	[0.220]	[0.467]	[0.724]	[0.247]	[0.081*]
Observations	394,409					
Chi2 H_0: parallel pre-trends	52.868					
P-Value H_0: parallel pre-trends	(0.000***)					
Panel B: Pass						
ATT (T = New Principal)	0.020	0.021	0.015	0.000	0.016	0.039
	(0.011)	(0.012)	(0.027)	(0.012)	(0.011)	(0.019)
P-Value	[0.085*]	[0.088*]	[0.566]	[0.982]	[0.151]	[0.047**]
Observations	398,013					
Chi2 H_0: parallel pre-trends	14.365					
P-Value H_0: parallel pre-trends	(0.110)					
Panel C: Romanian written score						
ATT (T = New Principal)	0.084	0.097	0.030	-0.022	0.064	0.188
	(0.067)	(0.079)	(0.100)	(0.082)	(0.069)	(0.100)
P-Value	[0.210]	[0.215]	[0.765]	[0.787]	[0.355]	[0.062*]
Observations	395,925					
Chi2 H_0: parallel pre-trends	9.1576					
P-Value H_0: parallel pre-trends	(0.421)					

Notes: The table displays ATT estimates from difference-in-difference specifications of the effect of T2 (the new principal policy) on student outcomes, using the double-robust inverse probability weighting estimator from Calloway and Sant'Anna (2021). Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present estimates for the entire working sample in column (1), ATTs by treatment groups in columns (2) and (3), and ATTs by period in columns (4)-(6). All specifications display the chi2 and p-values from the tests for the unconditional parallel trends assumption. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table 4 – All Survey Responses for Surveyed Schools; P-Value for T-Test between New and Legacy principals.

	Legacy Principal		New Principal		
Surveyed Schools with Exam	Mean	SD	Mean	SD	P Value
Values and Beliefs					
Ethics are Important	6.828	0.531	6.900	0.302	0.164
Fights for Rights of Others	6.197	1.211	6.414	0.807	0.093*
Civic Issues are Moral Obligations	6.527	0.940	6.657	0.720	0.232
Care for Welfare of Strangers	3.813	2.074	3.943	2.139	0.659
Financial Success is Important	2.616	1.752	2.171	1.569	0.050**
Effective Problem Solver	6.030	0.949	6.143	0.889	0.368
Sometimes Rude to Others	2.739	2.028	2.129	1.769	0.018**
Likes to Listen to and Help Others	6.443	0.796	6.400	0.969	0.737
Easy to Work with Others	6.631	0.800	6.443	1.072	0.183
Working Relationships					
Teacher Trust	5.581	1.013	5.257	1.259	0.054*
Director Trust	5.980	0.802	5.871	0.992	0.409
Share Common Values	6.054	0.961	5.971	1.063	0.566
Trust Enough to Delegate	6.192	0.877	6.186	0.952	0.961
Men Make Better Managers	1.872	1.443	1.529	0.944	0.025**
Management Techniques Compared to Pre-Policy					
Adapt Learning to Student Needs	2.726	0.469	2.594	0.577	0.089*
Tracking Teacher Performance	2.594	0.502	2.623	0.571	0.707
Tracking Objectives	2.764	0.459	2.754	0.526	0.889
Teacher Absence	2.335	0.654	2.265	0.704	0.471
Hiring Skilled Staff	2.470	0.557	2.368	0.621	0.229
Use of Staff Sanctions	2.097	0.551	2.118	0.441	0.756
Motivate Staff	2.502	0.575	2.638	0.593	0.102
Mobilization and Initiative Spirit	2.631	0.523	2.735	0.507	0.146
Manage Authority	2.577	0.505	2.559	0.583	0.818
Student Absence Rate	2.269	0.719	2.294	0.648	0.786
Student Dropout Rate	2.343	0.655	2.235	0.649	0.240
Graduation Rate of 12th Grade	2.624	0.553	2.559	0.608	0.438
Registration Rate for Baccalaureate	2.361	0.671	2.456	0.679	0.322
Observations	203		74		

Notes: The table displays mean survey responses for a full set of survey questions, along with a p value for the difference in means test, for our survey sample. Panel A displays results for the set of survey questions related to principal beliefs about morality and their own qualities in relation to work and co-workers. Panel B displays results for the set of questions related to working relationships. Panel C displays results for the set of questions related to how management techniques have changed compared to pre-policy. In column 1 we show responses for new principals (with standard deviations in column 2), and in column 3 we show responses for legacy principals (with standard deviations in column 4). *** p<0.01, ** p<0.05, * p<0.01

Table 5 – New vs Legacy Principal; Reduced Sample; ATT and dynamic effects on Baccalaureate Enrolment rates and initial cohort size

	Baccalaureate Enrolment rate					
	Sample	Group 2016	Group 2017	ATT 2017	ATT 2018	ATT 2019
Panel A: Baccalaureate Enrolment	(1)	(2)	(3)	(4)	(5)	(6)
ATT (T=New Principal)	-0.084	-0.069	-0.135	-0.010	-0.136	-0.082
	(0.036)	(0.042)	(0.053)	(0.023)	(0.071)	(0.029)
P-value	[0.021**]	[0.101]	[0.011**]	[0.662]	[0.055*]	[0.004***]
Observations	2,848					
Chi2 H_0: parallel pre-trends	5.022					
P-Value H_0: parallel pre-trends	[0.832]					
Panel B: Cohort size						
Admission ATT (T=New Principal)	3.876	4.083	3.165	0.979	5.655	4.142
	(3.605)	(3.984)	(9.855)	(3.872)	(4.563)	(5.480)
P-value	[0.282]	[0.306]	[0.748]	[0.800]	[0.215]	[0.450]
Observations						
Chi2 H_0: parallel pre-trends	18.42					
P-Value H_0: parallel pre-trends	[0.030**]					
Controls	No	No	No	No	No	No

Notes: The table displays Calloway and Sant'Anna (2021) estimates of the ATT of new principals on Baccalaureate Enrolment rates (panel A) and the initial cohort size (number of students admitted four years prior to the Baccalaureate exam, in panel B), including a test for unconditional parallel trends. We present estimates for the entire working sample in column (1), ATTs by treatment groups in columns (2) and (3), and ATTs by period in columns (4)-(6). All specifications display the chi2 and p-values from the tests for the unconditional parallel trends assumption. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table 6 – New vs Legacy principals by school pre-policy performance; ATT

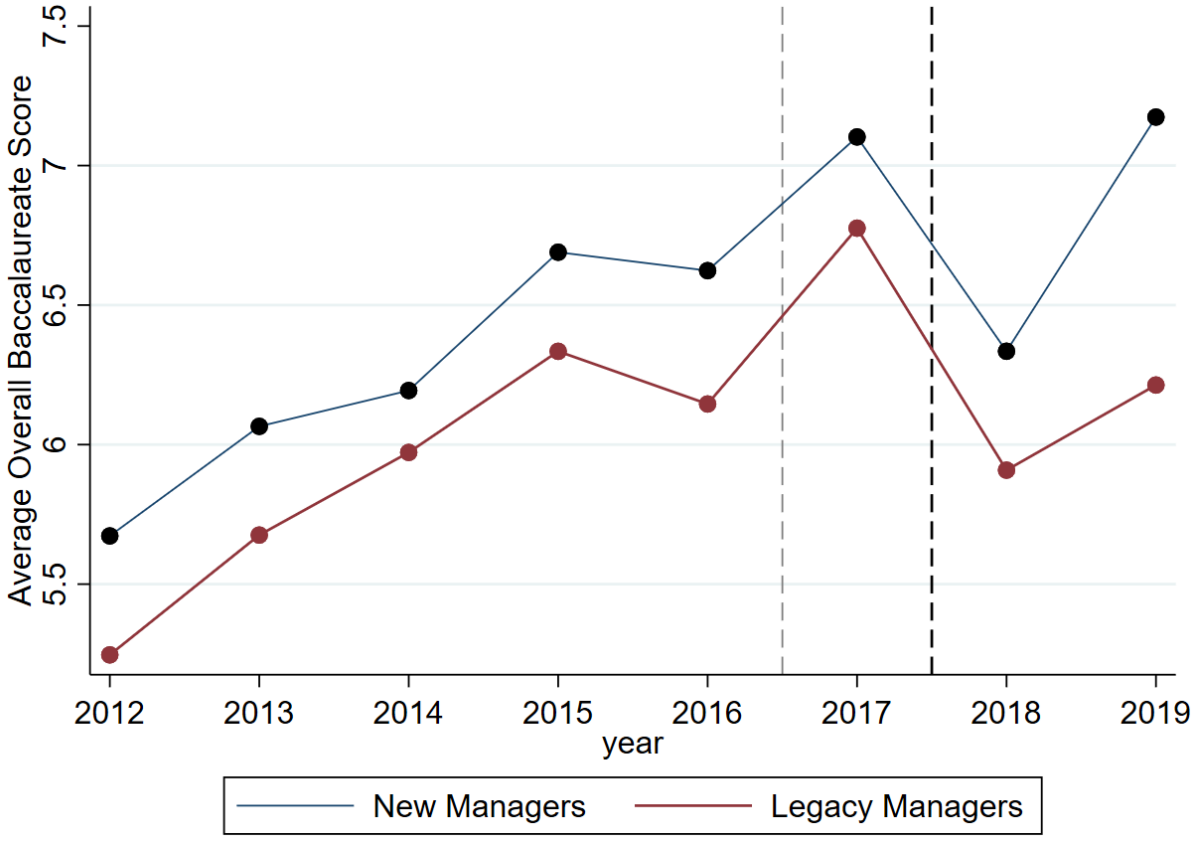
	>25th percentile			25th-75th percentile			>75th percentile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Sample	2016	2017	Sample	2016	2017	Sample	2016	2017
Panel A: Overall Baccalaureate Score									
ATT	-0.076	-0.114	0.052	0.075	0.050	0.168	0.026	0.029	0.013
	(0.073)	(0.089)	(0.164)	(0.082)	(0.088)	(0.092)	(0.057)	(0.062)	(0.105)
P-Value	[0.301]	[0.197]	[0.750]	[0.363]	[0.569]	[0.067*]	[0.650]	[0.643]	[0.901]
Observations	46,953			190,848			156,339		
Chi2 H_0: parallel pre-trends	244.762			25.718			10.731		
P-Value H_0: parallel pre-trends	[0.000***]			[0.002***]			[0.295]		
Panel B: Pass Rate									
ATT	-0.024	-0.024	-0.025	0.022	0.014	0.051	0.004	0.006	-0.002
	(0.014)	(0.018)	(0.037)	(0.019)	(0.019)	(0.023)	(0.007)	(0.009)	(0.019)
P-Value	[0.093*]	[0.177]	[0.491]	[0.238]	[0.471]	[0.029**]	[0.529]	[0.491]	[0.908]
Observations	47,969			193,099			156,661		
Chi2 H_0: parallel pre-trends	47.683			8.162			9.605		
P-Value H_0: parallel pre-trends	[0.000***]			[0.518]			[0.383]		
Panel C: Written Romanian Score									
ATT	-0.025	-0.128	0.310	0.038	0.045	0.014	0.042	0.046	0.024
	(0.114)	(0.111)	(0.219)	(0.105)	(0.104)	(0.185)	(0.070)	(0.072)	(0.152)
P-Value	[0.829]	[0.250]	[0.157]	[0.715]	[0.667]	[0.940]	[0.548]	[0.523]	[0.873]
Observations	47,265			191,878			156,504		
Chi2 H_0: parallel pre-trends	94.556			33.421			8.036		
P-Value H_0: parallel pre-trends	[0.000***]			[0.000***]			[0.531]		

Notes: The table displays ATT estimates from difference-in-difference specifications of the effect of T2 (the new principal policy) on student outcomes by pre-policy performance groups, using the double-robust inverse probability weighting estimator from Calloway and Sant'Anna (2021). Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. The sub-samples in terms of the 2012-2016 average exam score are displayed as follows: the low-performing schools (below the 25th percentile) in columns (1)-(3), middle-

performing schools (25^{th}-75^{th} percentile) in columns (4)-(6) present, and top-performing (above the 75^{th} percentile) in columns (7)-(9). We present estimates for the entire working sub-sample in columns (1), (4) and (7), ATTs by treatment groups in columns (2) – (3), (5) – (6), and (7)-(8). All specifications display the chi2 and p-values from the tests for the unconditional parallel trends assumption. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Supplementary Material (Online Appendix)

Appendix Figure AF1 – New vs Legacy Principal; <25th Percentile; Reduced Sample



Notes: The figure shows average overall Baccalaureate scores in schools <25th percentile for our reduced (working) sample between 2012-2019. We show the average scores for schools with new principals (in blue), and schools with legacy principals (in red). The light dashed vertical line demarcates the timing of the 2016 exam, whilst the dark dashed vertical line demarcates the timing of the 2017 exam.

Table A1 – Student Outcomes Pre- and Post-Policy for the Full Sample, Reduced Sample and Survey Sample

	Full Sample	**Reduced Sample**	**Survey Sample**
	Mean (sd)	Mean (sd)	Mean (sd)
	(1)	(3)	(4)
2012-2016 Overall Baccalaureate Score	6.118	6.176	5.990
	(2.348)	(2.288)	(2.283)
2017-2019 Overall Baccalaureate Score	6.832	6.846	6.621
	(2.153)	(2.132)	(2.165)
2012-2016 Romanian Score	6.082	6.060	5.963
	(2.331)	(2.293)	(2.278)
2017-2019 Romanian score	6.739	6.713	6.540
	(2.137)	(2.153)	(2.135)
2012-2016 Pass rate	0.548	0.556	0.519
	(0.498)	(0.497)	(0.500)
2017-2019 Pass rate	0.657	0.660	0.618
	(0.474)	(0.474)	(0.486)
Observations	1,260,671	431,940	248,899

Notes: The table displays mean student outcomes, taken for all student outcome data, for overall Baccalaureate score, written Romanian score and pass rates in our: full sample; reduced (working) sample; and survey sample. Mean outcomes are split into pre- (2012-2016) and post- (2017-2019) student outcomes with standard deviations displayed in parentheses.

Table A2 – Proportion of Schools with New and Legacy principals (Conditional on Competition)/Exam Selected Principals vs No Exam Selected Principals

	Legacy		New		Exam		No Exam	
	Number	Proportion	Number	Proportion	Number	Proportion	Number	Proportion
Full Sample								
Schools	-	-	-	-	1,425	87.02%	213	12.98%
Students	-	-	-	-	1,098,079	87.08%	162,572	12.92%
Reduced Sample								
Schools	329	71.06%	134	28.94%	463	90.78%	47	9.22%
Students	284,673	71.52%	113,340	28.48%	398,013	92.15%	33,927	7.85%
Schools (2016)	281	75.95%	89	24.05%	370	79.91%	-	-
Students (2016)	245,913	75.23%	80,979	24.77%	326,892	82.13%	-	-
Schools (2017)	48	51.61%	45	48.39%	93	20.09%	-	-
Students (2017)	38,760	54.50%	32,361	45.50%	71,121	17.87%	-	-
Survey Sample								
Schools	206	74.37%	71	25.63%	277	90.52%	29	9.48%
Students	161,855	75.65%	52,111	24.35%	213,966	85.96%	34,933	14.04%
Schools (2016)	182	86.67%	28	13.33%	210	75.81%	-	-
Students (2016)	148,832	85.66%	24,909	14.34%	173,741	81.20%	-	-
Schools (2017)	24	35.82%	43	64.18%	67	24.19%	-	-
Students (2017)	13,023	32.38%	27,202	67.62%	40,225	18.80%	-	-

Notes: The table displays raw number and percentages of legacy and new principals in the overall reduced (working) sample and survey sample; these are further broken down into groups who took part in the 2016 wave and 2017 wave of the competition. In addition, we provide the raw number and percentage of competitively selected (exam) and non-competitively selected (no exam) principals for the full sample, reduced (working) sample, and survey sample; these are again presented separately for the two waves of the policy (2016/2017). Note that in separate waves of the policy, we do not present non-competitively (no exam) selected principals, since due to the definition of principals selected as part of the policy wave, all principals were competitively selected.

Table A3. Bacon-Goodman Diff-in-Diff estimate and decomposition of effects. Strongly balanced panel of schools).

	(1) Exam Policy	(2) New Principal Policy
Panel A: Overall Baccalaureate Score		
DD Exam	0.026	
	(0.048)	
DD New Principal		0.029
		(0.035)
Observations	2,936	2,664
Panel B: Pass		
DD Exam	0.005	
	(0.012)	
DD New Principal		0.006
		(0.007)
Observations	2,936	2,664
Panel C: Romanian Exam Score		
DD Exam	-0.012	
	(0.073)	
DD New Principal		0.011
		(0.034)
Observations	2,936	2,664
W timing groups	0.431	0.040
W never vs 2016	0.472	0.720
W never vs 2017	0.070	0.205
W within	0.027	0.034
School FE	Yes	Yes
Year FE	Yes	Yes
County-specific trends	Yes	Yes
Controls	Yes	Yes

Note: The table displays the Goodman-Bacon (2021) decomposition of the effects into effects from pairwise conditional mean differences (between the early and late implementing groups, and between the implementers and never-implementers). We use a restricted sample of schools which form a strongly balanced panel (i.e., retaining only schools which have data for the entire time period 2012-2019), which is a requirement for implementing the Goodman-Bacon decomposition. Column (1) presents estimates of T1 (exam) and column 2 estimates of T2 (New Principal). All estimations include school and year fixed effects, country specific linear trends and controls (the share of theoretic and technical track students), and regressions are weighted by the number of students in 2019.

Table A4. Full Sample and Reduced Sample

	Full Sample		Reduced Sample	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)
Exam * Policy	0.044	-0.019	0.070	0.077*
	(0.040)	(0.039)	(0.041)	(0.040)
Constant	5.027***	5.081***	4.989***	4.953***
	(0.102)	(0.118)	(0.085)	(0.083)
Observations	1,247,350	1,247,350	428,061	428,061
R-squared	0.505	0.507	0.481	0.484
Pass Rate				
Exam * Policy	0.005	-0.006	0.013	0.014
	(0.008)	(0.008)	(0.008)	(0.009)
Constant	0.395***	0.399***	0.362***	0.354***
	(0.019)	(0.021)	(0.023)	(0.023)
Observations	1,260,651	1,260,651	431,940	431,940
R-squared	0.389	0.391	0.369	0.371
Written Romanian Score				
Exam * Policy	0.067	0.000	0.033	0.027
	(0.045)	(0.034)	(0.069)	(0.058)
Constant	4.552***	4.576***	4.487***	4.364***
	(0.082)	(0.101)	(0.093)	(0.087)
Observations	1,252,704	1,252,704	429,674	429,674
R-squared	0.421	0.423	0.415	0.417
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T1 (the meritocratic selection policy) on student outcomes. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present estimates for the full sample (all Romanian students) in columns 1 and 2, as well as our reduced (working) sample in columns 3 and 4. All specifications include year and school fixed effects, controls (theoretical and technical track dummies, and full-time student dummy) and county specific trends are included in columns 2 and 4. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.

Table A5 – Exam vs No Exam; Survey Sample

	Survey Sample	
Overall Baccalaureate Score	(1)	(2)
Exam * Policy	0.040	0.022
	(0.048)	(0.060)
Constant	5.101***	5.116***
	(0.179)	(0.211)
Observations	246,178	246,178
R-squared	0.471	0.475
Pass Rate		
Exam * Policy	0.001	-0.006
	(0.009)	(0.012)
Constant	0.387***	0.376***
	(0.035)	(0.040)
Observations	248,897	248,897
R-squared	0.363	0.366
Written Romanian Score		
Exam * Policy	0.028	-0.015
	(0.050)	(0.060)
Constant	4.672***	4.588***
	(0.161)	(0.222)
Observations	247,375	247,375
R-squared	0.395	0.397
Year FE	Yes	Yes
School FE	Yes	Yes
Controls	Yes	Yes
County Specific Trends	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T1 (the meritocratic selection policy) on student outcomes. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for overall survey sample in columns 1 and 2. Both specifications include year and school fixed effects and columns 2, includes county-specific trends. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table A6 – New vs Legacy Principal; Reduced Sample; Exam Selected Principals for 2016 and 2017

	Reduced Sample		Exam 2016		Exam 2017	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)	(5)	(6)
New Principal *	0.058	0.053	0.074	0.062	0.070	0.142
Policy	(0.043)	(0.036)	(0.051)	(0.043)	(0.109)	(0.104)
Constant	4.914***	4.916***	4.903***	4.874***	4.939***	6.509***
	(0.114)	(0.114)	(0.133)	(0.132)	(0.294)	(0.335)
Observations	394,409	394,409	324,102	324,102	70,307	70,307
R-squared	0.483	0.486	0.484	0.487	0.471	0.477
Pass Rate						
New Principal *	0.010	0.010	0.011	0.010	0.015	0.027
Policy	(0.007)	(0.006)	(0.009)	(0.009)	(0.018)	(0.018)
Constant	0.354***	0.345***	0.345***	0.331***	0.388***	0.546***
	(0.027)	(0.026)	(0.032)	(0.031)	(0.059)	(0.065)
Observations	398,013	398,013	326,892	326,892	71,121	71,121
R-squared	0.372	0.375	0.376	0.379	0.350	0.355
Written Romanian Score						
New Principal *	0.047	0.042	0.054	0.053	0.137	0.151
Policy	(0.036)	(0.029)	(0.044)	(0.033)	(0.102)	(0.093)
Constant	4.436***	4.344***	4.454***	4.344***	4.361***	5.531***
	(0.115)	(0.113)	(0.131)	(0.129)	(0.305)	(0.359)
Observations	395,925	395,925	325,330	325,330	70,595	70,595
R-squared	0.413	0.415	0.409	0.411	0.422	0.426
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T2 (the impact of new principals) on student outcomes. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for the entire reduced (working) sample (columns 1 and 2), as well as separately for the two waves of the policy (columns 3-4 for the 2016 selection and columns 5-6 for the 2017 selection). All specifications include year and school fixed effects and columns 2, 4 and 6 include county-specific trends. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table A7 – New vs Legacy Principals; Reduced Sample; Percentiles Calculated from School Average Final Score 2012-2016

	< 25th Percentile		25th – 75th Percentile		> 75th Percentile	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)	(5)	(6)
New Principal * Policy	0.188	0.263*	0.094	0.044	-0.000	-0.024
	(0.121)	(0.134)	(0.068)	(0.067)	(0.054)	(0.047)
Constant	3.783***	4.380***	4.131***	4.160***	4.936***	4.820***
	(0.337)	(0.374)	(0.171)	(0.176)	(0.287)	(0.294)
Observations	46,953	46,953	190,848	190,848	156,339	156,339
R-squared	0.143	0.151	0.210	0.217	0.187	0.192
Pass Rate						
New Principal * Policy	0.024*	0.039**	0.018	0.010	0.003	0.001
	(0.012)	(0.014)	(0.012)	(0.012)	(0.010)	(0.009)
Constant	0.202***	0.204***	0.181***	0.156***	0.247***	0.250***
	(0.052)	(0.055)	(0.040)	(0.040)	(0.054)	(0.055)
Observations	47,969	47,969	193,099	193,099	156,661	156,661
R-squared	0.060	0.063	0.151	0.156	0.087	0.090
Written Romanian Score						
New Principal * Policy	0.253**	0.238*	0.057	0.049	-0.006	-0.061*
	(0.114)	(0.130)	(0.077)	(0.077)	(0.062)	(0.031)
Constant	3.317***	3.582***	3.811***	3.678***	4.026***	3.894***
	(0.237)	(0.259)	(0.143)	(0.145)	(0.428)	(0.431)
Observations	47,265	47,265	191,878	191,878	156,504	156,504
R-squared	0.160	0.164	0.197	0.200	0.201	0.207
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T2 (the impact of new principals) on student outcomes in our reduced (working) sample. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for schools in the <25th percentile (columns 1 and 2), 25th-75th percentile (columns 3 and 4) and >75th percentile (columns 5 and 6); percentile groups are calculated based on pre-policy (before 2016) overall Baccalaureate scores. All specifications include year and school fixed effects and columns 2, 4 and 6 include county-specific trends. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table A8 – Exam vs No Exam; Reduced Sample; Percentiles Calculated from School Average Final Score 2012-2016

	< 25th Percentile		25th – 75th Percentile		> 75th Percentile	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)	(5)	(6)
Exam * Policy	0.166	0.141	-0.001	-0.022	-0.016	0.016
	(0.181)	(0.156)	(0.077)	(0.073)	(0.077)	(0.079)
Constant	3.776***	4.407***	4.132***	4.170***	4.936***	4.815***
	(0.339)	(0.367)	(0.172)	(0.172)	(0.285)	(0.290)
Observations	46,953	46,953	190,848	190,848	156,339	156,339
R-squared	0.143	0.151	0.210	0.217	0.187	0.192
Pass Rate						
Exam * Policy	0.017	0.013	-0.001	-0.001	0.004	0.007
	(0.023)	(0.022)	(0.022)	(0.022)	(0.016)	(0.016)
Constant	0.201***	0.208***	0.181***	0.158***	0.247***	0.250***
	(0.053)	(0.055)	(0.040)	(0.039)	(0.053)	(0.054)
Observations	47,969	47,969	193,099	193,099	156,661	156,661
R-squared	0.060	0.063	0.151	0.156	0.087	0.090
Written Romanian Score						
Exam * Policy	0.055	0.039	-0.116	-0.121	-0.057	-0.020
	(0.182)	(0.152)	(0.136)	(0.126)	(0.114)	(0.105)
Constant	3.311***	3.611***	3.811***	3.692***	4.025***	3.884***
	(0.231)	(0.249)	(0.143)	(0.144)	(0.426)	(0.428)
Observations	47,265	47,265	191,878	191,878	156,504	156,504
R-squared	0.159	0.164	0.197	0.200	0.201	0.207
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T1 (the meritocratic selection policy) on student outcomes in our reduced (working) sample. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for schools in the <25th percentile (columns 1 and 2), 25th-75th percentile (columns 3 and 4) and >75th percentile (columns 5 and 6); percentile groups are calculated based on pre-policy (before 2016) overall Baccalaureate scores. All specifications include year and school fixed effects and columns 2, 4 and 6 include county-specific trends. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table A9 – Exam vs No Exam; ATT and dynamic effects; Robustness control group not yet treated

	Sample	Group 2016	Group 2017	ATT 2017	ATT 2018	ATT 2019
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Overall Baccalaureate Exam Score					
ATT (T=principal exam)	0.014	0.029	-0.095	-0.068	0.035	0.066
	(0.098)	(0.101)	(0.144)	(0.056)	(0.092)	(0.199)
P-Value	[0.884]	[0.771]	[0.510]	[0.225]	[0.703]	[0.740]
Observations	428,061					
Chi2 H_0: parallel pre-trends	71.354					
P-Value H_0: parallel pre-trends	[0.000***]					
	Panel B: Pass rate					
ATT (T=principal exam)	0.001	0.004	-0.020	-0.002	-0.005	0.010
	(0.016)	(0.015)	(0.028)	(0.009)	(0.018)	(0.034)
P-Value	[0.949]	[0.795]	[0.479]	[0.860]	[0.780]	[0.764]
Observations	431,940					
Chi2 H_0: parallel pre-trends	13.127					
P-Value H_0: parallel pre-trends	[0.157]					
	Panel C: Romanian written exam score					
ATT (T=principal exam)	0.037	0.045	-0.020	-0.140	0.073	0.157
	(0.115)	(0.117)	(0.178)	(0.094)	(0.092)	(0.223)
P-Value	[0.748]	[0.700]	[0.908]	[0.134]	[0.423]	[0.480]
Observations	429,674					
Chi2 H_0: parallel pre-trends	6.638					
P-Value H_0: parallel pre-trends	[0.675]					

Notes: The table displays ATT estimates from difference-in-difference specifications of the effect of T1 (the meritocratic selection policy) on student outcomes, using the double-robust inverse probability weighting estimator from Calloway and Sant'Anna (2021) with the not yet treated schools as a control group. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present estimates for the entire working sample in column (1), ATTs by treatment groups in columns (2) and (3), and ATTs by period in columns (4)-(6). All specifications display the chi2 and p-values from the tests for the conditional parallel trends assumption. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Table A10 – New vs Legacy Principal; ATT and dynamic effects; Robustness control group not yet treated

	Sample	Group 2016	Group 2017	ATT 2017	ATT 2018	ATT 2019
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall Baccalaureate Score						
ATT (T=principal exam)	0.081	0.084	0.070	-0.022	0.081	0.162
	(0.061)	(0.068)	(0.095)	(0.068)	(0.070)	(0.093)
P-Value	[0.183]	[0.216]	[0.467]	[0.745]	[0.247]	[0.081*]
Observations						
Chi2 H_0: parallel pre-trends						
P-Value H_0: parallel pre-trends						
Panel B: Pass						
ATT (T=principal exam)	0.020	0.021	0.015	0.001	0.016	0.039
	(0.011)	(0.012)	(0.027)	(0.011)	(0.011)	(0.019)
P-Value	[0.083*]	[0.082*]	[0.566]	[0.932]	[0.151]	[0.047**]
Observations						
Chi2 H_0: parallel pre-trends						
P-Value H_0: parallel pre-trends						
Panel C: Romanian written score						
ATT (T=principal exam)	0.081	0.094	0.030	-0.033	0.064	0.188
	(0.066)	(0.077)	(0.100)	(0.078)	(0.069)	(0.100)
P-Value	[0.221]	[0.223]	[0.765]	[0.674]	[0.355]	[0.062*]
Observations						
Chi2 H_0: parallel pre-trends	9.180					
P-Value H_0: parallel pre-trends	[0.421]					

Notes: The table displays ATT estimates from difference-in-difference specifications of the effect of T2 (the new principal policy) on student outcomes, using the double-robust inverse probability weighting estimator from Calloway and Sant'Anna (2021) and the not yet treated schools as a control group. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present estimates for the entire working sample in column (1), ATTs by treatment groups in columns (2) and (3), and ATTs by period in columns (4)-(6). All specifications display the chi2 and p-values from the tests for the unconditional parallel trends assumption. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A11 – New vs Legacy Principal; Placebo Year Robustness Test (Policy = 2015/2016); Reduced Sample; Exam Selected Principals for 2016 and 2017

	Reduced Sample		Exam 2016		Exam 2017	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)	(5)	(6)
New Principal * Policy	0.016	0.013	0.006	0.003	0.103	0.097
	(0.034)	(0.034)	(0.037)	(0.037)	(0.117)	(0.112)
New Principal * Policy t-1	0.053	0.051	0.083*	0.072*	-0.039	0.058
	(0.039)	(0.035)	(0.046)	(0.039)	(0.100)	(0.105)
Constant	4.903***	4.903***	4.882***	4.852***	4.939***	6.500***
	(0.117)	(0.118)	(0.136)	(0.136)	(0.294)	(0.338)
Observations	394,409	394,409	324,102	324,102	70,307	70,307
R-squared	0.483	0.486	0.484	0.487	0.471	0.477
Pass Rate						
New Principal * Policy	0.008	0.008	0.004	0.004	0.023	0.019
	(0.006)	(0.006)	(0.007)	(0.007)	(0.024)	(0.023)
New Principal * Policy t-1	0.003	0.004	0.009	0.007	-0.009	0.010
	(0.009)	(0.009)	(0.008)	(0.008)	(0.025)	(0.026)
Constant	0.353***	0.344***	0.343***	0.328***	0.388***	0.544***
	(0.027)	(0.027)	(0.032)	(0.032)	(0.059)	(0.066)
Observations	398,013	398,013	326,892	326,892	71,121	71,121
R-squared	0.372	0.375	0.376	0.379	0.350	0.355
Written Romanian Score						
New Principal * Policy	-0.000	0.000	0.030	0.032	0.108	0.096
	(0.040)	(0.037)	(0.045)	(0.039)	(0.124)	(0.121)
New Principal * Policy t-1	0.060	0.054	0.029	0.026	0.034	0.069
	(0.036)	(0.034)	(0.032)	(0.032)	(0.101)	(0.106)
Constant	4.424***	4.330***	4.447***	4.336***	4.361***	5.520***
	(0.118)	(0.116)	(0.135)	(0.133)	(0.305)	(0.361)
Observations	395,925	395,925	325,330	325,330	70,595	70,595
R-squared	0.413	0.415	0.409	0.411	0.422	0.426
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T2 (the impact of new principals) on student outcomes. Placebo treatment variables are included, enacting the policy one year before it took place. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for reduced (working) sample (columns 1 and 2), 2016 exam selected principals (columns 3 and 4) and 2017 exam selected principals (columns 5 and 6). All specifications include year and school fixed effects and columns 2, 4 and 6 include county-specific trends. County-clustered standard errors in parentheses for 19 clusters. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A12 – New vs Legacy Principals; Wild Bootstrap; Reduced Sample; Percentiles Calculated from School Average Final Score 2012-2015

	Reduced Sample		Exam 2016		Exam 2017	
Overall Baccalaureate Score	(1)	(2)	(3)	(4)	(5)	(6)
New Principal *	0.058	0.053	0.074	0.062	0.070	0.142
Policy	[1.356]	[1.484]	[1.462]	[1.443]	[0.639]	[1.366]
	(0.201)	(0.157)	(0.171)	(0.168)	(0.600)	(0.217)
Constant	4.914***	4.916***	4.903***	4.874***	4.939***	6.509***
	(0.114)	(0.114)	(0.133)	(0.132)	(0.294)	(0.335)
Observations	394,409	394,409	324,102	324,102	70,307	70,307
R-squared	0.483	0.486	0.484	0.487	0.471	0.477
Pass Rate						
New Principal *	0.010	0.010	0.011	0.010	0.015	0.027
Policy	[1.418]	[1.707]	[1.238]	[1.115]	[0.836]	[1.494]
	(0.178)	(0.110)	(0.236)	(0.309)	(0.456)	(0.173)
Constant	0.354***	0.345***	0.345***	0.331***	0.388***	0.546***
	(0.027)	(0.026)	(0.032)	(0.031)	(0.059)	(0.065)
Observations	398,013	398,013	326,892	326,892	71,121	71,121
R-squared	0.372	0.375	0.376	0.379	0.350	0.355
Written Romanian Score						
New Principal *	0.047	0.042	0.054	0.053	0.137	0.151
Policy	[1.309]	[1.464]	[1.214]	[1.596]	[1.336]	[1.625]
	(0.220)	(0.166)	(0.224)	(0.140)	(0.216)	(0.114)
Constant	4.436***	4.344***	4.454***	4.344***	4.361***	5.531***
	(0.115)	(0.113)	(0.131)	(0.129)	(0.305)	(0.359)
Observations	395,925	395,925	325,330	325,330	70,595	70,595
R-squared	0.413	0.415	0.409	0.411	0.422	0.426
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Specific Trends	No	Yes	No	Yes	No	Yes

Notes: The table displays OLS estimates from difference-in-difference specifications of the effect of T2 (the impact of new principals) on student outcomes in our reduced (working) sample. Panel A displays results for the overall Baccalaureate Score, Panel B for the passing rate, and Panel C for the standard Romanian Written Exam scores. We present the estimates for schools in the <25th percentile (columns 1 and 2), 25th-75th percentile (columns 3 and 4) and >75th percentile (columns 5 and 6); percentile groups are calculated based on pre-policy (before 2016) overall Baccalaureate scores. All specifications include year and school fixed effects and columns 2, 4 and 6 include county-specific trends. Wild bootstrap t-value is shown in square parentheses and wild bootstrap p-value is shown in round parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Full Survey Instrument conducted in November 2017 in collaboration with ISSPOL

Questionnaire on the evolution and challenges of management in the pre-university school environment

Section 0 - Introduction

Hello! My name is …… and I am calling from ISSPOL. We are currently conducting a survey on the evolution and challenges of management in a pre-university school environment. Do you have a few moments to answer this questionnaire?

Section 1 – General Questions

I. 1. How long have you been the director in your school? _____

I. 2. How many years of experience do you have in school management?

_____

I. 3. How many years of experience do you have in pre-university education?

_____

I. 4. How many students are currently enrolled in your high school*? [approximately]*

_____

I. 5. How many teachers are employed in your institution, and what is the proportion of men/women?

I. 5. *Number of teachers:* _____

I. 5. A. *Percent male:* _____

I. 5. B. *Percent female:* _____

Section 2 – Management activity in 2016-17 and comparison with previous years

On a scale of 1 to 7, where 1 means total disagreement and 7 means total agreement, please choose the option that best suits your answer.

II. 1. The function of school principal	Totally disagree	Disagree	Partially disagree	Neither agree nor disagree	Partially agree	Agree	Totally Agree
a. The ethical behaviour of school directors is as important to me as their competencies.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
b. I'm not afraid to fight for the rights of others, even if that means being ridiculed.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
c. I think that all people have a moral obligation to get involved in civic issues no matter how busy they are.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
d. I rarely think of the	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7

welfare of people that I do not know personally.							
e. It is more important for me to succeed financially than to do good deeds.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
f. I think that I am an effective problem solver.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
g. I think that I am a person who can sometimes be rude to others.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
h. I think that I am a person who listens to and helps others.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
i. It is easy for me to work with the Deputy Director and other colleagues.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
II. 2. Working Relationships	**Totally disagree**	**Disagree**	**Partially disagree**	**Neither agree nor disagree**	**Partially agree**	**Agree**	**Totally Agree**
a. Teachers in my high school trust each other.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
b. Teachers in my high school trust the decisions of the directors.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
c. If I think about how the teachers in my high school solve challenges, I believe that we have a common set of values.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
d. I trust colleagues enough to delegate assignments.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
e. I think that a man is more capable than a woman at running a high school.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7

II. 3. How do you feel that the following areas of your high school performance have evolved in 2016-17 when compared to pervious years.

	Better	Worse	The Same	N/A
a. Adapting learning activities to student needs.	❑ 1	❑ 2	❑ 3	❑ 4
b. Tracking and evaluating the performance of teachers and school activities.	❑ 1	❑ 2	❑ 3	❑ 4
c. Establishing, communicating and tracking school objectives related to student outcomes.	❑ 1	❑ 2	❑ 3	❑ 4
d. Absence among teachers.	❑ 1	❑ 2	❑ 3	❑ 4
e. Hiring and retaining high-skilled staff members.	❑ 1	❑ 2	❑ 3	❑ 4
f. Use of staff sanctions (e.g. reprisals, weak evaluations, administrative sanctions, layoffs).	❑ 1	❑ 2	❑ 3	❑ 4
g. Motivating your staff (e.g.) through praise, good ratings, bonuses, promotions).	❑ 1	❑ 2	❑ 3	❑ 4
h. Mobilization and initivative spirit	❑ 1	❑ 2	❑ 3	❑ 4
i. Managing authority (ease of implementation, respect, staff discipline over decisions made by executives).	❑ 1	❑ 2	❑ 3	❑ 4
j. Absence among students.	❑ 1	❑ 2	❑ 3	❑ 4
k. School dropout rate.	❑ 1	❑ 2	❑ 3	❑ 4
l. Gradutation rate of the 12th grade.	❑ 1	❑ 2	❑ 3	❑ 4
m. Registration rate for the baccalaureate.	❑ 1	❑ 2	❑ 3	❑ 4

II. 3_j_1. What is the absentee rate in your institution? (in the school year 2016-17)
_____%

II. 3_k_1. What is the dropout rate in your institution? (in the school year 2016-17)
_____%

II. 3_m_1. What is the baccalaureate enrollment rate in your institution? (in the school year 2016-17)

_____%

II. 4. Does your high school have any initiatives in place to improve the school performance of students from disadvantages backgrounds, or high school drop-outs? (e.g. remedial hours, after school hot meal)

YES in 2016/17. Examples ... Pass to II.4.1	❑ 1
Yes before 2016	❑ 2
No	❑ 3

II. 5. (Original: If your high school has applied and has been approved funds in the ROSE project, funded by the World Bank, what amount have you received?)
(in link) Did your high school apply to the ROSE program, financed by the World Bank, and have funds approved?

Did not apply	❑ 1
Was not approved	❑ 2
Yes. Go to II.5.1	❑ 3

II. 5. 1. How much funding did you receive through the ROSE project? _____

ROSE is a grant project for eligible high schools (those with low baccalaureate promotion rates and many underprivileged students) aimed at increasing graduate exam outcomes and discouraging school dropouts. The funds are around 70,000-150,000 Euros for each beneficiary school. More details can be found at: https://www.edu.ro/271-de-licee-beneficiaz%C4%83-de-granturi-%C3%AEn-cadrul-proiectului-privind-%C3%AEnv%C4%83%C8%9B%C4%83m%C3%A2ntul-secundarromania

Section 3:

Running the manager contest – Tick all the options that apply to you.

III. 1. Have you run in the competition to become director?

A. Yes, in 2016	❑ 1
B. Yes, in 2017	❑ 2
C. No, I was seconded to the interest of education	❑ 3

III. 2. If you answered 1.A or 1.B: How did you decide to join the competition?

A. Own initiative	❑ 1
B. At the request of my colleagues	❑ 2
C. To prevent someone from another school gaining the post in my high school	❑ 3
D. On recommendation of the inspectors/other people in the education system	❑ 4
E. Other reasons (please provide short description)_____	❑ 5

III. 3. If you answered 1.B or 1.C: Why did you decide not to register for the competition in 2016?

A. Too little time to prepare	❑ 1
B. The literature was completely new	❑ 2
C. I was advised not to sign up	❑ 3
I felt that I would be disfavoured/discriminated against during the evaluation	❑ 4
E. Other reasons (please provide short description) _____	❑ 5

III. 4. Consider the session running October-November 2016. Winning this competition for the post of director in Romania was based upon:

Based on merit (theoretical knowledge or capabilities)	❑ 1
Partially based on merit, partially on political knowledge/connections	❑ 2
Based on political knowledge/connections	❑ 3

III. 5. Consider the session running July-August 2017. Winning this competition for the post of director in Romania was based upon:

Based on merit (theoretical knowledge or capabilities)	❑ 1
Partially based on merit, partially on political knowledge/connections	❑ 2
Based on political knowledge/connections	❑ 3

III. 6. Do you know people who retired or did not participate at all (in the 2016 or 2017 competitions) because they felt that they would be disfavoured/discriminated against?

Yes	❑ 1
No	❑ 2

III. 7. Select the one that best suits your experience:

	Totally disagree	Disagree	Partially disagree	Neither agree nor disagree	Partially agree	Agree	Totally Agree
a. The competition tested the skills expected from a school manager.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
b. The competition attracted candidates with high motivation and abilities and increased the quality of management in schools.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7
c. A candidate with previous experience as a manager, or worked in that school, has a competitive advantage	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7

over other candidates.							
d. It would have been necessary/helpful to take part in a training program in order to prepare for this exam.	❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6	❑ 7

III. 8. Section 3, Question 1, if YES: After winning the post by contest/appointment, what has changed in your high school? (You can refer to any aspect, from efficiency in administration, to ease of money management, to attracting funds and funding, to the attitude of teachers or parents, to student behaviour and school performance.)

_____

III. 9. What would you like to change at the administrative level in your high school over the next four years?

_____

Demografice:

D1. Name: _____

D2. Age:_____

D3. Gender:_____

D4. School/High School: _____

D5. Town: _____

D6. County:_____

D7. Email address: _____

D8. Telephone Number: _____

Operator Code: _____