Who Leaves and Who Stays? The Impact of the Teacher Wage Gap on Teacher Quality

Desislava Tartova*

Job Market Paper

October 30, 2025

Click here for the most recent version.

Abstract

This paper studies how uniform increases in teacher wages affect the average quality of the teaching workforce. I exploit a 2014 French reform that substantially increased wages in highly disadvantaged schools but only modestly in slightly less disadvantaged ones to identify the heterogeneous exit elasticities with respect to wages across teachers of different productivity levels. I show that high-productivity teachers—defined using plausibly unbiased measures of teacher value-added—are three times more responsive to wage increases in their exit decisions than low-productivity teachers. I develop a simple labor-supply framework showing that high-productivity teachers are more responsive to wages because their preferences between teaching and the outside option are less dispersed and, on average, closer to the indifference threshold. I find empirical support for a mechanism that can rationalize the latter: teachers with better average outside options have more incentives to collect information about them, thus reducing the dispersion in their relative preferences. Finally, policy counterfactuals show that while uniform wage increases raise aggregate teacher quality, given the differential elasticities, targeting increases toward high-productivity teachers delivers larger gains at lower fiscal cost.

*Paris School of Economics (PSE). Email address: desislava.tartova@psemail.eu. I am very grateful to Luc Behaghel, Thomas Piketty and Jonah Rockoff for their guidance and support. I also thank Gabrielle Fack, Manon Garrouste, Nina Guyon, Antoine Hubert de Fraisse, Sylvie Lambert, Karen Macours, David Margolis and Adrien Matray for their comments and suggestions. This work has been funded by a French government subsidy managed by the French National Research Agency (ANR) under the framework of the "Investments for the Future" program (reference ANR-17-EURE-001). The provision of MENJS-DEPP data was made possible by the Innovation, Data and Experimentation in Education (IDEE) project, funded by the ANR under the framework of the "Investments for the Future" program (reference ANR-21-ESRE-0034).

Teacher wages around the world fall short of those of similarly educated workers, with secondary-school teachers in OECD countries today earning on average 90 cents for every dollar their peers make (OECD 2023)—as low as 66 cents in the United States. This gap has been steadily increasing over time—tripling in size since the late 1970s in the United States (Allegretto 2025). This stylized fact is mirrored by the widening of the feasible teacher wage gap—the wage gap between teaching and occupations that teachers can move to after leaving the profession—which has doubled over the last decade in the case of France. Such variation in relative compensation is likely to be an important determinant in the decision to enter or exit the teaching sector.

In the context of teachers' exit decisions, studies suggest that paying teachers higher wages, thus reducing this wage gap, increases retention in the profession.[1] However, a question remains: *does a uniform increase in teacher wages increase the quality of the teaching force by affecting its composition?* This matters because teacher quality is an important driver of human capital accumulation that affects students' long-term earnings (Chetty, Friedman, & Rockoff 2014b). The answer is a priori not obvious. The effect of higher wages on the quality of the teaching sector may depend on how teachers' preferences for wage and non-wage amenities vary with teaching productivity. Some teachers would remain regardless of pay ("never movers"), while others would leave even if wages rose ("always movers"). If more productive teachers have stronger outside options, they may include many "always movers", while less productive teachers, facing weaker alternatives, may include many "never movers". Only the composition of "potential movers"—those whose decision is sensitive to changes in wages—would affect the quality of the teaching force in response to a change in wages. Whether these potential movers are predominantly high- or low-productivity teachers is theoretically ambiguous and empirically understudied.

Providing an answer to this question is empirically challenging for at least two reasons. First, identifying a causal effect requires a shock to the wage gap that is exogenous to other determinants of teachers' labor supply decisions; in practice, variation in either teacher or outside wages is often endogenous to local economic conditions and individual productivity. Second, identifying productivity-specific effects requires unbiased measures of teacher productivity. The standard proxy is teacher value-added (TVA)—the causal effect of a teacher on students' scores on standardized exams—but typical TVA estimators (Rockoff 2004; Kane & Staiger 2008; Chetty, Friedman, & Rockoff 2014a) rely on annual standardized tests to address non-random sorting of students to teachers, whereas in most countries such tests are given only at key stages several years apart.

This paper provides the first causal evidence on whether high- and low-productivity teachers differ in their propensity to exit the profession when teacher wages increase

[1] See Dolton and van der Klaauw (1995); Benhenda and Sims (2022); Feng and Sass (2018); Clotfelter, Glennie, Ladd, and Vigdor (2008); Falch (2011, 2017); Cowan and Goldhaber (2018); Hendricks (2014).

uniformly, and rationalizes the results in a static discrete-choice framework of labor supply with two sectors—teaching and an outside option. Using a difference-in-differences approach, I exploit a plausibly exogenous shock to teacher wages from a 2014 French reform that significantly raised wages across teachers in highly disadvantaged ("REP+") middle schools, while only marginally increasing wages in other disadvantaged ("REP") middle schools. I find that the exit decisions of high-productivity teachers are three times as elastic to wages as those of low-productivity teachers, where I proxy for teacher productivity by a novel, plausibly unbiased measure of teacher value-added (Tartova 2023), which does not rely on annual standardized testing. These results show that uniformly increasing teacher wages raises aggregate teaching quality. I confirm this in the static discrete-choice framework. Taking the estimated elasticities as given, I show in policy counterfactuals that while uniform increases in wages do raise aggregate teacher quality, targeted increases towards high-productivity teachers deliver larger gains at half the fiscal cost. The framework implies that the higher elasticity of high-productivity teachers is a product of a lower dispersion of the distribution of preferences for the outside option relative to teaching, and a mean of the distribution closer to the indifference threshold. I propose a mechanism that would explain why high-productivity teachers are more elastic to changes in wages: they have better outside options. Teachers with better outside options have more incentives to learn about them, which increases the precision with which they make a decision to leave and, in turn, moves a larger mass of teachers in response to changes in the wage gap. I find empirical evidence in support of this, showing that teachers with higher pre-reform outside options are more responsive to the change in wages introduced by the reform.

To causally identify teachers' labor supply response to wage changes, I exploit a 2014 reform in France and the institutional features of the French teacher labor market. The reform introduced a sizable differential treatment between REP+ and REP teachers, resulting in a 7% average increase in wages for the average REP+ teacher over the first six years of the reform, relative to less than a 2% increase for REP school teachers. The treatment is also characterized by its large scope, with REP+ and REP teachers representing, respectively, 5% and 20% of all public middle school teachers. To plausibly isolate labor supply decisions, I focus on decisions to exit the teaching profession by tenured teachers. This is because tenured teachers (more than 90% of all teachers) are civil servants who cannot be dismissed for performance, meaning that their observed exit is a choice. By contrast, because entry and internal mobility within the teaching sector are constrained by France's centralized teacher allocation system, analyzing these margins would conflate labor supply responses with labor demand and institutional constraints.

I estimate the causal effect of the wage increase induced by the reform on teacher exit using an event-study difference-in-differences design, comparing exit decisions of REP+

teachers to those of REP teachers, based on panel data covering the universe of tenured teachers in these schools. My baseline empirical approach relies on instrumenting for actual exposure to the reform by exploiting the pre-reform allocation of teachers to schools, and thus estimating the Intention-to-Treat (ITT) effect. Focusing on incumbent teachers at REP+ and REP schools at the time of the reform rules out biases that may arise from post-reform mobility from either within the educational sector or from outside. My preferred specification controls for teacher characteristics, school fixed effects, and commuting zone–by–year fixed effects.

The reduced-form estimates indicate that, on average over the six years following the announcement, the wage bonus reduced the exit rate of teachers initially assigned to REP+ schools by 25 percent relative to the counterfactual exit rate predicted in the absence of the differential wage increase, compared to teachers initially assigned to REP schools. Instrumenting wage changes with the school-type-specific bonus implied by the reform in a Two-Stage Least Squares (2SLS) analysis implies an exit elasticity with respect to the wage for the average teacher for whom the treatment was intended of -5.5 over the first six years. In other words, a 1 percent increase in wages reduced the probability of the average teacher to exit the teaching profession by 5.5 percent, indicating a strong negative responsiveness of exits to changes in wages. Because compliance is imperfect—teachers designated as treated may move out of treated schools—I use the rate of compliance to identify the local average treatment effect (LATE) for teachers who were initially assigned to the REP+ schools and who stayed in these schools after the reform was implemented. Given an average compliance rate of 77.5 percent over the six years following the introduction of the bonus, the implied exit elasticity to the wage bonus for the initially assigned teachers who actually received the bonus is -7.

The validity of the difference-in-differences design relies on the parallel trends assumption, which requires that, absent the reform, exit rates of REP+ and REP teachers would have followed similar trajectories after 2014. While the pre-trend in the baseline ITT cannot be observed because, by definition, the treatment is defined based on the subsample of teachers that had not exited by the time of the reform (in other words, selection of treatment is based on survival until the last pre-reform year), I perform two diagnostics of the identifying assumption using alternative definitions of treatment. First, I fix treatment at the teacher level using each teacher's last observed pre-reform school (irrespective of the year of that last observation). This removes the mechanical pre-period zeros induced by conditioning on presence in 2013–2014 and allows estimation of pre-trends for fixed cohorts ("last-REP+" vs "last-REP"). Second, I use the contemporaneous school assignment, which traces the pre-reform evolution of exits among teachers actually in REP+ versus REP schools in each year. Across both approaches, the pre-period exhibits a small positive differential linear trend in exits at REP+ relative to REP that reverses sharply at the reform date. This structural break in the relative probability of exit at the time of the reform is consistent with a causal impact of the policy, rather than a continuation of pre-existing trends. Thus, my baseline results could be considered a lower bound of the true effects, if the relative evolution of exits would have continued increasing linearly in the post-period, had it not been for the differential wage increase.

To study the heterogeneous effect by teacher productivity, I proxy teacher productivity with a novel measure of teacher value-added (TVA). Standardized tests in French middle schools were administered for Math and French only at the end of 9th grade during the reform period, making it impossible to use standard unbiased TVA estimators, which would require lagged student scores in standardized exams to control for the potential student-teacher sorting. To address this issue, I estimate TVA using the network method developed in Tartova (2023). This approach relies on linking students' cross-sectional standardized test scores in Math and French to their teachers, and achieves unbiasedness by exploiting "networks" of teachers who share students across subjects. Intuitively, if two Math teachers both teach classes with the same French teacher, the common French teacher's effect cancels out when comparing the difference in their students' Math and French test scores, thus isolating the relative contribution of the two Math teachers. By transitivity, these pairwise comparisons can be linked together within schools, and—using teacher mobility—across schools. The unbiasedness of the TVA estimates relies on the assumption that there is no sorting on subject-specific ability—that is, differences in students' relative abilities in Math and French—a type of sorting for which there is little to no evidence in middle schools, based on tests on observables (Tartova 2023).[2]

I find that high-TVA teachers are significantly more responsive to the reform-induced change in teacher wages, where high-TVA teachers are defined as those with TVA above the median in each subject. Specifically, the estimated exit elasticity to the wage for high-TVA teachers is three times as large as that of low-TVA teachers. This result is robust to a range of alternative specifications, controls and a rich set of fixed effects. The heterogeneous effect between high- and low-TVA teachers persists after accounting for potential differential treatment effects across other observable dimensions, including experience, age, gender, and qualification level. The latter suggests that the observed heterogeneity by TVA reflects a broader productivity gradient rather than being driven by one particular characteristic.

I develop a static discrete-choice model of labor-supply decisions with two sectors (teaching and outside option) and two teaching productivity types (high- and low-TVA), in order to rationalize the heterogeneous responsiveness to wages of teachers with different levels of productivity found empirically, and in order to quantify the

[2]The method provides a reliable proxy for the standard TVA estimator (Kane & Staiger 2008), as shown in both Monte Carlo simulations and observational data from New York City—where annual testing allows for the two methods to be compared (Tartova 2023).

implications of changes in wages for aggregate quality in the teaching sector. Utility from working in either sector depends linearly on three ingredients: sectoral wages, non-pecuniary preferences for each sector, and idiosyncratic sector-specific taste shocks. I assume that (i) taste shocks follow Gumbel distributions that are specific to each TVA type, (ii) teaching wages do not vary by TVA, while outside wages are allowed to vary with TVA, and (iii) both types share the same non-pecuniary preferences.[3]

In this framework, the equilibrium probability of exiting teaching for each teacher type depends positively on the wage gap between outside options and teaching, and negatively on the gap in non-wage amenities between the two. The responsiveness of teachers to wage changes depends on where the expected relative preference for the outside option versus teaching lies with respect to the point of indifference between staying and leaving, and on how dispersed individual preferences are. When comparing two groups with the same equilibrium exit rate (that is, the same mass on both sides of the indifference point), the group with more similar preferences (less dispersion) must, on average, lie closer to this indifference threshold. In this case, more teachers of that group are close to being indifferent between teaching and the outside option, so a small change in the wage gap shifts a larger mass of these teachers across the indifference threshold. I demonstrate that when more productive teachers have a higher exit semi-elasticity than low-productive ones, a uniform increase in teacher wages unambiguously increases aggregate teaching quality through the change in composition induced by this exit channel.

I show that one can estimate the exogenous framework parameters, and, thus, predict the quantitative effect of changes in the wage gap for aggregate quality, using five empirical moments: the two type-specific exit elasticities, the two type-specific equilibrium exit rates, and the weighted average of the type-specific wage gaps between outside options and teaching. I measure the two type-specific exit elasticities using the ITT baseline estimates, and the equilibrium exit rates using the type-specific exit rates observed in the data prior to the reform. The latter entail approximately the same equilibrium exit rates for both groups, and therefore a distribution of relative preferences for the high-TVA type that lies closer to the indifference threshold. Finally, I use the pre-reform average wage gap between outside options and teaching that I observe from the data to estimate the type-specific wage gaps, by assuming that the observed average is weighted by the equilibrium exit probabilities of each type. Specifically, I construct the pre-reform wage gap for REP+ teachers by combining nationally set seniority-specific teacher wages with counterfactual outside wages. The latter are proxied using weighted average wages in the same commuting zone–age–sex–year cell among occupations that teachers are observed to

[3]Note that assumption (iii) is in line with recent findings from the literature (e.g. Johnston in press), but evidence on differential preferences for non-wage amenities by TVA are mixed, and earlier research suggests that high-TVA teachers value non-pecuniary aspects of teaching more strongly (see, for instance, Bó, Finan, & Rossi 2013). I show qualitatively that relaxing assumption (iii) to allow high-TVA teachers to value non-pecuniary aspects of teaching more strongly would amplify the effect of an increase in the wage gap on aggregate productivity.

transition into, as identified from occupation transitions in exhaustive employer–employee panel data.[4]

Using the estimated framework parameters, I first consider a counterfactual uniform increase in teacher wages equivalent to the relative REP+ bonus, holding outside wages constant. The model predicts only modest gains in aggregate quality—about 1.6 percent of a standard deviation in TVA over five years—rising to 3.3 percent when the amount is targeted to high-TVA teachers. These limited effects reflect that, at low equilibrium exit rates, few teachers are marginal to a wage increase. Under the Gumbel assumption, teachers' relative preferences for the outside option versus teaching follow a logistic distribution—they are concentrated around the middle of the distribution and become sparse in the tails. Because equilibrium exit rates are low, most teachers are located on the left of the indifference threshold—where teaching is much more attractive, such that there are predominantly "always stayers". Consequently, an increase in the teacher wage affects few teachers, as it shifts the type-specific distributions of relative preferences further away from the indifference threshold.

Conversely, an increase in outside wages—which increases the wage gap—shifts the distributions of relative preferences closer to the indifference threshold, making a larger share of teachers marginal and producing a sharper rise in exits. As the relative preference distribution of the high-TVA type lies closer to the indifference threshold, this translates into a larger decline in aggregate quality. For an increase in the outside wage of an equivalent amount as in the previous counterfactual, I show that aggregate quality decreases by 4 percent of a SD following a uniform increase in outside wages over five years, growing to 14.3 percent when concentrated to only the high-TVA type. This asymmetry arises under thin-tailed, unimodal distributions of relative preferences.

These simulations suggest three main conclusions. First, uniform increases in wages, all else equal, increase aggregate teacher quality through the exit channel, since high-productivity teachers' exit decisions are more elastic to wages than those of low-productivity teachers. Second, targeted wage increases for high-productivity teachers are more cost-effective, since high-productivity teachers' exit decisions are elastic to wages. Third, in a setting where outside-option wages are increasing, failing to adjust teacher wages to match the rising opportunity cost of teaching can lead to substantial losses in aggregate teacher quality.

I propose a mechanism consistent with the theoretical framework which can explain the higher elasticity of high-TVA teachers, given that their distribution of relative preferences for the outside option lies closer to the indifference threshold. Specifically, teachers with larger expected outside-option wages—and therefore larger wage

[4] Linking teacher wages and the estimated outside option wages to individual teachers, I find that the estimated wage gap for the average 2013 REP+ teacher stands at 4,000 euro—that is, the outside option wage is 15% higher than the average REP+ teacher wage.

gaps—optimally allocate more attention to learning about outside opportunities, which leads to greater precision when making a decision (reducing the dispersion of their distribution of relative preferences), and in turn, moves a larger mass of teachers in response to a change in the wage gap.

I find evidence consistent with this mechanism. Teachers with higher pre-reform wage gaps reduce their exit rate significantly less in response to the teacher wage increase, relative to the same difference observed in the control group. Splitting the wage gap into its two components—the outside option wage and the teacher wage—further shows that the higher responsiveness of teachers with a larger wage gap is driven by the teachers with higher outside option wages. Likewise, groups characterized by larger estimated wage gaps—specifically older teachers—exhibit a stronger decline in exit rates in response to the wage increase.

Related Literature This paper contributes to three main strands of the literature.

First, it relates to work on the effects of uniform changes in teacher wages on teacher quality.[5] Growing quasi-experimental evidence on the intensive margin, studying whether teachers change the effort they exert, is mixed. Ree, Muralidharan, Pradhan, and Rogers (2017) and Bobba et al. (2021) show that uniform increases in teacher wages do not lead to a within-teacher increase in effort, similar to Bates and Johnston (2025) who exploits an effective decrease in teacher wages around the pension-eligibility threshold, while Johnston, Rockoff, and Harrington (2025) finds that an effective decrease in teacher wages (driven by a policy which discourages early retirement) increases within-teacher effort.[6] Regarding the extensive margin, non-experimental studies suggest that larger wage gaps—arising either from higher outside wages or lower teacher wages—are associated with more exits and less entries of high-quality teachers, which suggests a lower aggregate teacher quality.[7] The most closely related paper to my

[5]A complementary stream of the literature focuses on asymmetric changes in teacher wages in certain schools, such as pay for performance or flexible wage schemes targeting high-quality teachers, overwhelmingly showing a positive effect on aggregate teacher quality in these schools (with the exception of Willén 2021), through both increased effort (Atkinson et al. 2009; Biasi 2021; Biasi, Fu, & Stromme 2021; Cowan & Goldhaber 2018; Dee & Keys 2004; Dee & Wyckoff 2015; Duflo, Hanna, & Ryan 2012; Figlio & Kenny 2007; Glewwe, Ilias, & Kremer 2010; Leaver, Ozier, Serneels, & Zeitlin 2021; Muralidharan & Sundararaman 2011), and the reallocation of high-quality teachers from other schools (Biasi 2021; Biasi et al. 2021; Brown & Andrabi 2025; Burgess, Greaves, & Murphy 2022; Cowan & Goldhaber 2018; Johnston in press; Morgan, Nguyen, Hanushek, Ost, & Rivkin 2023). Relatedly, papers that focus on bonuses based on school-wide performance, rather than teacher-specific performance, find little to no effect (Barrera-Osorio & Raju 2017; Fryer 2013; Goodman & Turner 2013). By design, such policies cannot speak to differential elasticities to wages across the productivity distribution as they target only high-productivity teachers.

[6]A finding which sides with models that predict that a longer work horizon encourages skill investment and discourages shirking (Becker 1962; Ben-Porath 1967; Gibbons & Murphy 1992; Lazear 1979).

[7]These papers use geographical and time-series variation in teacher wages (Bacolod 2007; Dolton & Marcenaro-Gutierrez 2011; Figlio 1997, 2002), in outside wages (Britton & Propper 2016, in both (Hoxby & Leigh 2004), or more broadly in economic conditions, such as unemployment (e.g., Deneault 2023; Falch, Johansen, & Strøm 2009; Fraenkel 2022; Nagler, Piopiunik, & West 2020), typically proxying for teacher quality using teachers' education, or students' standardized test scores.

quasi-experimental analysis is Bates and Johnston (2025). The authors find no differential retention in teaching for high-productivity teachers compared to low-productivity ones, given an effective decrease in future compensation around the pension-eligibility threshold. As the effective outside option for teachers at retirement is exiting the labor force altogether rather than transitioning to another job, that setting provides valuable insights into late-career retention. However, the differential elasticities obtained in such settings are unlikely to be informative about elasticities for younger teachers who are more likely to transition to another job, as pensions of teachers are not productivity-specific, but outside options are likely to be.[8] My paper is able to address this gap. Overall, my contribution in this strand is to provide, to my knowledge, the first causal estimates of productivity-specific exit elasticities to wages, showing that high-productivity teachers are more responsive than low-productivity ones in their choice of whether to stay in the educational sector or exit for another job.

Second, a related stream of the literature focuses on the effects of uniform bonuses in disadvantaged schools for the across-school inequality in teacher quality. Studies generally find that such bonuses are effective in reallocating certified or more experienced teachers to disadvantaged schools from advantaged to disadvantaged schools (Cabrera & Webbink 2020; Clotfelter et al. 2008; Pugatch & Schroeder 2014, 2018; Silhol & Wilner 2022). As certification and experience have been shown in the literature to be, at best, only weakly associated with teacher productivity (see Jackson (2014) for a review of the literature), these studies may only provide suggestive evidence of a decrease in the educational gap. Using data from Peru, Bobba et al. (2021) provides evidence that such policies indeed attract better teachers to disadvantaged schools, using both a measure of teacher value-added and data on applications for mobility within the teaching sector to plausibly isolate labor supply responses. My paper contributes to this literature by showing that policies that increase wages at disadvantaged schools can reduce the educational gap with advantaged schools not only by reallocating teachers across schools, but also through inducing disproportionately more good teachers in these schools not to exit the educational sector.

Third, the paper contributes to the broader literature on labor supply elasticities and imperfect competition in labor markets. A large body of work highlights monopsonistic features of labor markets (for instance, Hirsch, Jahn, Manning, & Oberfichtner 2022; Manning 2003, 2011; Naidu & Carr 2022), with related evidence in the public sector (e.g., Bó et al. 2013; Staiger & Rockoff 2010), where rigid pay scales and centralized allocation create limited mobility. In teaching, evidence on exits comes either from structural models (e.g., Dolton & van der Klaauw 1995), or from quasi-experimental settings often related

[8] In addition, because of the nature of the variation used in Bates and Johnston (2025), comparison is made only between individuals who are towards the end of their career and thus may respond to wage incentives differently than the average teacher. I am instead able to study the entire teacher workforce at disadvantaged schools.

to wage incentives at disadvantaged schools (e.g., Benhenda & Sims 2022; Clotfelter et al. 2008; Cowan & Goldhaber 2018; Falch 2011, 2017; Feng & Sass 2018; Hendricks 2014). Such studies typically find turnover elasticities to the teacher wage of between -3 to -6. The size of these elasticities is increasing in the persistence of the wage variation and in the salience of the reform. As turnover encompasses both mobility to other schools and exit from the educational sector, these elasticities are not directly comparable to the ones in this paper. Specifically, mobility across schools is often rigid as positions are scarce, compared to the entirety of outside options. My quasi-experimental evidence from the French teacher labor market complements existing estimates by plausibly isolating an average exit elasticity to teacher wages of -7.[9]

The paper proceeds as follows. Section 1 presents the French institutional setting, and explains the 2014 reform, which introduced the wage bonus at disadvantaged schools. Section 2 introduces the data used in the study. Section 3 presents the network methodology for TVA estimation derived in Tartova (2023). Section 4 details the empirical strategy and results. Section 5 outlines the theoretical framework and explores counterfactual policy counterfactuals. Section 6 provides evidence on the mechanism implied by the framework. Finally, Section 7 concludes.

1. Institutional setting

This section provides details on the French institutional setting and discusses in detail the 2014 reform that led to the differential increase in wages for REP+ and REP teachers.

1.1. The teaching profession

The teaching workforce in French public middle schools consists mainly of tenured teachers (*professeurs titulaires*) and a smaller share (between 5% and 10% of the teacher workforce) of contract teachers (*enseignants contractuels*).

Tenured teachers Tenured teachers form the core of the teaching workforce and are civil servants in the State's public education service. As such, they are recruited via competitive national examinations, allocated to schools based on a centralized system operated at a national level, hold permanent positions, progress automatically through a national pay scale based on seniority, and cannot be dismissed for performance reasons except in extreme cases involving disciplinary proceedings.

Entry into a tenured position requires at least a Master's degree and passing the relevant national exam. There are two main national exams which provide the certification level needed to become a middle school teacher—*Capes* (*Certificat*

[9]For comparability, I show that when taking into account both exits and transfers to other schools, I obtain a turnover elasticity to the teacher wage of -3.

d'aptitude au professorat de l'enseignement du second degré), which provides a basic certification level, and *Agrégation*, which provides an advanced certification level. The latter is much more difficult to obtain, as it is meant for individuals who target teaching at high schools or universities. As a result, less than 5% of middle school teachers are Agrégation-certified.

Newly recruited teachers serve a one-year probationary period with reduced teaching loads before receiving full tenure. After tenure, a standard teaching load for a Capes-certified teacher is 18 hours per week, whereas that of an Agrégation-certified teacher is 15 hours per week.

Tenured teachers are not directly recruited by schools, but are instead assigned to a school based on a centralized point-based system called *SIAM* (*Systeme d'information et d'aide aux mutations*). Points (and therefore priority in allocation preferences) are given to teachers for their seniority, years of experience in their current school, the need for spousal reunification, and existing disability, as well as years of experience at disadvantaged schools. Consequently, entrants into teaching have little say over their initial allocation, as they are primarily allocated based on residual needs. For this reason, I do not focus on studying the impact of the reform on the probability to enter teaching, as the first allocations of teachers may not be representative of their preferences.

Teacher wages are set on a national wage scale and depend primarily on the certification level and seniority of the teacher.[10] In addition to the base wage, teachers may receive small bonuses due to overtime work, area of residence, number of children, and working at a disadvantaged school. In addition, having top scores on evaluations based on pedagogical classroom inspections and completion of administrative obligations helps reduce the number of years of experience needed to progress in seniority.[11] Finally, teachers with at least 3 years of experience at disadvantaged schools also need fewer years of experience in order to progress on the seniority scale.

Contractual teachers In France's public middle schools (*collèges*), contractual teachers are non-tenured teachers hired on fixed-term or, more rarely, permanent contracts. Unlike civil servant teachers, they are not recruited through the national competitive exams and do not hold the civil service status. They are primarily hired to cover teacher shortages or to replace absent civil servants. As such, they are more prevalent in highly disadvantaged schools, which are considered hard-to-staff.

[10] Agrégation-certified teachers are paid according to a higher pay scale than Capes-certified teachers.

[11] More specifically, teachers are ranked on a list for promotions (*tableau d'avancement*) according to their pedagogical grade (with weight of 60%) and administrative grade (with weight of 40%). Teachers ranked in the highest level of the promotion list (*grand choix*) need relatively less years of teaching experience in order to progress on the wage scale, compared to teachers in the middle (*choix*) or bottom (*ancienneté*) of the list. This channel is very small, as the points received for performance are proportionally marginal.

In principle, contract teachers are required to hold at least a three-year higher education degree (Bachelor's level) in the relevant teaching subject.[12] Their salaries are set on a pay scale specific to non-tenured teachers, determined by qualifications and experience. Unlike tenured teachers, they do not progress automatically through salary steps.

I exclude contract teachers from my analysis because their employment conditions, and most notably dismissal possibilities, differ markedly from those of tenured staff. Specifically, most are employed on fixed-term contracts, often for one school year or shorter, which may simply not be renewed without justification. After six continuous years of service in the public education system, contractual teachers may obtain an open-ended contract, which offers greater protection but termination is still possible.

1.2. History of priority education and the 2014 reform

Brief history The French State established the first priority education system called ZEP (*Zone d'education prioritaire*) in 1982, with the goal of reducing the inequalities linked to social origins in academic success. Originally, 503 middle schools were classified as belonging to a priority education zone.[13] By 1999, the number of middle schools part of the system had increased to 1,000—such that every 1 in 5 middle school students was part of the system. As part of this program, a yearly teacher bonus of 300 euros was first introduced in 1990. By 2006, the compensation had reached 1,156 euros per year.

In 2006, the priority networks were split in two tiers, based on the intensity of social and academic difficulties. The most disadvantaged middle schools (254) were given a RAR label whereas the rest of the disadvantaged middle schools (782) were relabelled as RRS. Importantly, while some additional resources were provided for RAR schools, such as the employment of additional teachers and educational assistants, the teacher bonus was held constant across tiers.

Introduction of the 2014 reform By 2013, multiple evaluations had concluded that the priority education policies to date had largely failed to achieve its objectives. Despite successive reclassifications of schools and increases in resources, large socio-economic and academic achievement gaps persisted between students in priority and non-priority schools. Previous research finds little to no effect of the policy on student outcomes (Bénabou, Kramarz, & Prost 2009), and in some cases negative effects—through increased disparities in teacher quality across schools (Beffy & Davezies 2013) and greater social segregation as advantaged students move to private schools (Davezies & Garrouste 2020). Moreover, disadvantaged schools continued to face chronic staffing difficulties, with higher teacher turnover, difficulty attracting experienced teachers, and over-representation of novice or contractual teachers. These

[12] In shortage areas or for hard-to-fill subjects, this requirement may be relaxed, and candidates with lower-level qualifications and relevant experience may be hired.

[13] Primary schools were also part of the program, as well as few high schools. As my analysis focuses on middle schools, I will focus on the evolution of the priority education system for middle schools.

findings prompted the government audit (Modernisation de l'action publique) published in July 2013 that recommended a complete redesign of the policy, outlining the introduction of two new tiers of priority education: REP (Réseaux d'éducation prioritaire) REP and REP+, to replace the RRS and RAR tiers, respectively. The document included a detailed outline of the packages of policies envisioned for REP and REP+ schools, importantly including the precise amounts of staff bonuses.

During the same school year 2013-2014, specifically in January 2014, the French Ministry of Education announced formally the overhaul of the priority education system, redesigning the map of disadvantaged secondary schools. The new status classifications into REP and REP+ school were based on an index created by the Ministry of Education which was based on four indicators: the shares of very disadvantaged students within the school, need-based scholarship holders, students who repeated a grade before entering middle school, and students living in or very close to a disadvantaged urban neighborhood. Self-selection of schools into the program was not possible and there was little scope for manipulation, as historic shares already reported to the Ministry were used.

During the 2014-2015 school year, priority was given to the 100 networks in the highest difficulty, which already entered the new system. Each network is centered on a middle school and its primary schools. During the same school year, the rest of the list of schools which would become part of the system in the 2015-2016 school year was revealed, for a total of 365 REP+ and 732 REP middle schools.

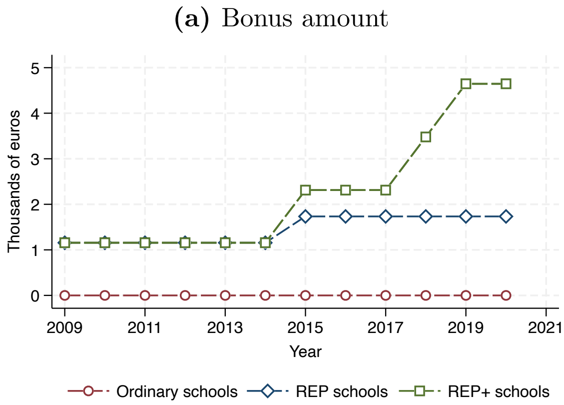
During the 2015-2016 school year, all announced priority schools entered the program, and the announced measures, including the staff bonus, were put in place.

Scope of the 2014 reform The 2014 reform of priority education introduced a common set of measures across both REP and REP+ networks, many of which were directly targeted at improving student learning conditions. At the middle school level, the pedagogical strand of the reform emphasized differentiated instruction, supported by tools such as a digital platform (D'Col) which offered sixth-grade students targeted exercises in core subjects, and structured homework assistance programs for 2 hours per week. Schools also benefited from the deployment of additional non-teaching staff—including education assistants, social workers, and school nurses—intended to strengthen the support system around students in disadvantaged areas.

For teachers, the reform sought to reinforce professional development. Dedicated time of 1.5 hours per week was introduced for collective work with field experts. Novice teachers in REP and REP+ schools benefited from reinforced mentoring and closer supervision for 3 days per year (increased to 18 half-days in 2015).

Two additional measures were intended for REP+ middle school teachers, and not for REP teachers, thus introducing a differential treatment I exploit in my empirical strategy.

First, the main aspect of the differential reform was that REP+ teachers received a significantly higher financial bonus than REP teachers. Figure 1 presents the unveiling of the bonus introduced in REP and REP+ schools starting from the 2015-2016 school year. As seen in Panel (a), in 2015-2016 the yearly bonus for teachers in REP+ schools more than doubled compared to the pre-reform amount (from 1,156 to 2,312 euros). For REP schools, the increase was smaller (to 1,734 euros, or about a 2 percent increase compared to pre-reform levels). Post-2015, the bonus in REP schools was not increased further. For REP+ schools, the yearly bonus was further increased to 3,479 euros in 2018-2019 (announced in the summer of 2018 prior to the start of the school year), then to 4,646 euros in 2019-2020. In sum, on average, teachers at REP+ schools received an average increase of 5 percent over the first four years, and an average increase of 7 percent over the first six years (Panel (b). For REP teachers, the increase for the average teacher was less than 2 percent.



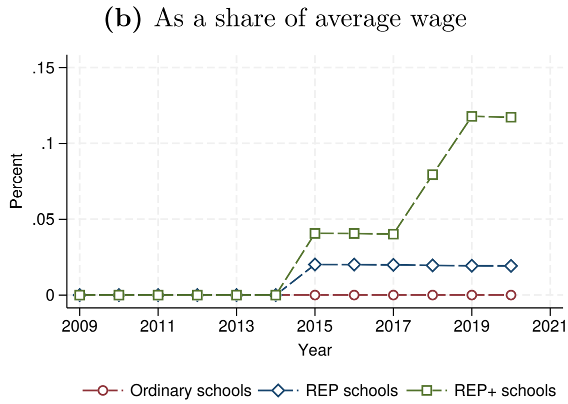


Figure 1: The yearly bonus scheme after the 2014 reform

Note: The figures show the yearly bonus introduced for teachers at REP and REP+ schools after the reform announced in 2014. Panel (a) shows the gross yearly bonus amount, not adjusted for inflation. Panel (b) shows the change in the bonus in real terms, as a share of the average wage in the given type of school. Underlying data source: Bases Relais, `https://lucaschancel.com/enseignants/`, 2009-2010 to 2020-2021.

Second, teaching service hours were re-weighted: each classroom hour started counting as 1.1 hours in the official calculation of teaching load. For a full-time Capes-certified (Agrégation-certified) teacher with a statutory teaching obligation of 18 (15) hours per week (out of a mandated 35-hour working week), this translated into an envisioned effective teaching load of approximately 16.4 (13.6) hours. The remaining 1.6 (1.4) hours per week were supposed to be reallocated to student follow-up and liaison with families. This treatment was not meant to decrease working hours, but to reallocate a small portion of working hours. In practice, official reports and union communications indicate that the reform rarely translated into the intended reduction in teaching hours, but instead led to an overall increase in working hours due to the additional time spent in liaison with families and students.[14] Consistent with this, the data show that average teaching hours fell by 0.7 hours per week, far less than the 1.6-hour reduction implied by the reform's design (see Section 4.1.4). This confirms that the reform led to an overall increase in hours. On the one-hand, if teachers respond negatively to increases in working hours, this treatment may have led to an increase probability of exit, thus offsetting the effect of the bonus. On the other hand, to the extent that the increase in working hours was converted into overtime pay that is marginally better remunerated compare to before the reform, it may have served as an additional small increase in wages.[15] In Section 4.1.4 I discuss the implications of this for analysis in more detail and show that under conservative assumptions, this does not quantitatively affect the baseline results.

2. Data and measurement

Teacher-school panel data The main data I use in this paper is teacher panel data from the Statistical Office of the French Ministry of Education (DEPP).

I construct a national-level panel of all middle-school teachers in Metropolitan France between the 2007-2008 and 2021-2022 school years, using the Bases Relais datasets. The database is a collection of yearly snapshots of active teachers in the school registry of all public schools in France. It provides personal information on all teachers, such as their date of birth, gender, qualification, seniority, tenure, specialization, experience, as well as position in the wage scale. It also provides information on the schools, subjects, and classes taught by each teacher every year.

The exhaustive information allows me to identify the year, school and commuting zone from which a teacher exits the profession. I define *exit* from the profession as 1 if a teacher in school year t—$t+1$ is not observed in secondary school teaching during at least school year $t+1$—$t+2$ and school year $t+2$—$t+3$, such that exit share in school year t—$t+1$ is defined as the teachers last observed in school year t—$t+1$, as a share of all teachers at t—$t+1$. This definition therefore abstracts from short-term interruptions from work (e.g. due to sick leave, maternity leave, or certification training) which may otherwise be mistaken for exits. The definition allows for individuals to re-enter teaching in year $t+3$—$t+4$. As teachers have a grace period of 5 years in which they can exit teaching and come back without losing their public servant status, this allows me to identify short-term exits, as well as long-term ones. I restrict the sample of teachers to tenured teachers, in order to isolate pure labor supply decisions, given that the employment of contractual teachers

[14] Independent public reports later confirmed that the reform was "often converted into annual overtime hours", undermining its goal (Cour des comptes 2018). Similarly, two major French teachers' unions, SGEN-CFDT and SNEP-FSU, noted that many REP+ teachers were still required to perform mandatory overtime, in addition to the introduced weekly hours for follow-ups with students and parents, contradicting the original purpose of the reform (SGEN-CFDT 2020, SNEP-FSU 2018).

[15] See Appendix C.4 for precise information on the calculation in overtime pay.

is not immune to dismissals. Furthermore, I take the subsample of exits for individuals under 50 to avoid capturing exits due to (early) retirement, which would add noise to the estimation.

The resulting aggregate exit rates by experience group are plotted in Figure B.7 for the sample of all tenured teachers. The sample is restricted to the school year 2019-2020, which is the last year for which I can apply my definition of exit. Exit rates have consistently been increasing across experience groups and subject taught for the last decade, but most strikingly so for teachers at the beginning of their career, whose exit rates have nearly doubled since 2012.

School panel data I compute average school-by-year student characteristics with the use of the Scolarité database (DEPP), in order to compare the composition of students across types of schools. The database collects student personal information from yearly school registries for all public schools in France.

The characteristics of the average tenured teacher below 50 years old teaching at a public middle school in 2013-2014, as well as the average student characteristics within each type of school for the school year 2013-2014 (i.e. the year of the announcement of the reform) are depicted in Table A.1.

REP+ school teachers represent 5% of the sample of all teachers in 2013, while REP teachers—roughly 20%. REP+ teachers are less likely to be female (60%, compared to 63% at REP and 67% at ordinary schools), they are less experienced overall (8 years on average, compared to 9 at REP and 11 at ordinary), younger (36 years old on average, compared to 37 at REP and 39 at ordinary), and stay at a school slightly less time (3.5 years, compared to 3.7 at REP and 3.9 at ordinary). They are equally likely to have a high qualification ("Agrégation"), but have a higher ex-ante exit rate (2.8%), compared to REP (2%) and ordinary schools (1.7%).

As Table A.1 also shows, there is substantial segregation between students across types of schools. REP+ students are significantly more likely to be labeled as "very disadvantaged" (74%, compared to 55% at REP and 36% at ordinary), and less likely to be labeled as "very advantaged (3%, compared to 10% at REP and 22% at ordinary), where the level of "advantageness" is defined on the basis of parental occupation. They are also sugnificantly more likely to receive a need-based scholarship (39%, compared to 25% at REP and 14% at ordinary), and be foreign nationals (13%, compared to 7% at REP and 2% at ordinary).

Matched student-teacher data In order to study the heterogeneous effect of the change in the wage for different levels of TVA, I match the teacher panel data to cross-sectional student data for the universe of students, provided by the DEPP. Specifically, as students sit standardized exams in Math and French at the end of the 9th grade (called the DNB exams), I match the subsample of 9th-grade Math and French teachers to their students, based on classroom identifiers, between 2007-2008 and 2020-2021.[16] This would allow to obtain TVA estimates for all Math and French teachers who have taught the 9th grade at least once during this period—about 80% of Math and French teachers, on average (see Table A.2).

I construct my student database relying on three databases—Scolarité, Sysca and DNB. Scolarité and Sysca contain very similar personal information on students (e.g. birth date, gender, parental occupation, need-based scholarship, nationality), but come with a caveat. Scolarité does not have student identifiers but has reliable class identifiers, whereas Sysca has student identifiers but does not have class identifiers prior to 2014. While I do not use student panel data and therefore do not need to be able to track students across years, I require student identifiers to merge the data to the database DNB, which contains students' scores in the Math and French DNB exams. I use observable student characteristics to perform a fuzzy match between the two datasets and obtain a panel of students covering about 95% of all students, which I merge with exam scores.

The merge between student and teacher data leads to the loss of, on average, roughly 10% of 9th-grade Math and French teachers, due to missing or wrongly encoded classroom identifiers on either the teacher- or the student-side of the data. Finally, the identification requirements of the method for TVA estimation (specifically, the requirement of belonging to the closed network of teachers, see Section 3 for more details), leads to a further reduction of the sample of Math and French teachers with a TVA measure, as 95% of the sample of teachers merged to student test scores belongs to the complete network. This leads to having a TVA measure for, on average, 70% of all Math and French middle school teachers (Table A.2). These numbers are slightly lower at REP+ (57% of Math teachers and 58% of French teachers) and REP (63% of Math or French teachers), due to the fact these teachers are observed in comparatively less classrooms (they have lower experience level).

Table A.2 further shows that the average Math and French teachers in the last pre-reform year are very similar to the average teacher, with two exceptions. First, Math teachers are much less likely to be female—both on average (56%) and at REP+ (50%) and REP (52%) schools. By contrast, French teachers are more likely to be female compared to the average teacher (80% on average). However, they are less likely to be female at REP+ (74%) and REP (75%) schools. Second, the average ex-ante exit rate for both Math and French teachers is higher than that of the average teacher. At REP+ schools, Math and French teachers exit at a rate of 3.3% and 3.8%, respectively, while their counterparts at REP schools exit at a rate of 2.5% and 2.7%, respectively, and their counterparts at ordinary schools—at a rate of 2% and 2.1%. These higher ex-ante unconditional probabilities may reflect several factors. One likely explanation is the additional pressure faced by Math

[16] Note that there were no standardized exams for the 2019-2020 academic year due to Covid, which entails that TVA cannot be estimated for teachers who taught only in the academic year 2019-2020. For all other teachers who taught in that year and other academic years, their TVA can be inferred based on their classroom observations excluding those in 2019-2020.

and French teachers, since the government views these as the most critical subjects.

Teacher salary data To compute the exit elasticities to the teacher wage and construct the wage gap necessary for the theoretical framework, I use publicly available data on teacher wages which vary by teachers' position on the wage scale (echelon) and over time.[17] The panel data exists for the sample of teachers with a basic (Capes) qualification, which comprises of the large majority of teachers in my sample (more than 90%).

Matched employer-employee data To construct the outside-option-wage side of the wage gap, I rely on exhaustive French employee–employer matched data from tax returns (BTS-Postes, previously called DADS Postes), provided by the National Institute of Statistics and Economic Studies (INSEE), covering the period 2009–2021. To extend the BTS-Postes into a longitudinal dataset covering the full 2009–2021 period, I apply the methodology proposed by Godechot, Palladino, and Babet (2023). Details on the methodology are available in Appendix C.1.

This extended panel allows me to observe the full trajectory over the period of 2009-2021 of individuals who were teachers at some point over this period. Using 4-digit occupation codes and NAF industry classifications, I am able to identify secondary school teachers in general education, and using the legal status of each employer, I am able to narrow the sample of interest to public school teachers. I use the national registry of businesses and establishments (Sirene) in order to obtain the names of secondary education establishments and identify those marked as middle schools (collèges), from the set of secondary schools (which include high schools). This allows to focus specifically on teachers in middle schools.

To define the set of outside option occupations to which middle school teachers transition, I restrict the sample of individuals to those identified for at least a year as teachers in public middle schools at any year t between 2009 and 2021, who are below 50 years old, who work full-time, and for whom teaching is a main job, based on the number of hours spent in each occupation at a given year. I identify as "exits" observations of individuals who are public middle school teachers in year t and who switch out of teaching in year $t+k$ (i.e. are not observed as either pre-primary, primary or secondary school teacher, and are not observed as a headmaster of a school, which can be considered as promotion rather than exit). I consider $k = \{1, 3, 5\}$ in order show how the target teachers switch across occupations (defined at an occupation-level 4) over time, allowing for the possibility that the job they take at $t + 1$ is a temporary job.[18]

Figure B.2 reports the top 15 broad occupations attracting exiting teachers for $k = \{1, 3, 5\}$, based on national exit shares to a given occupation. The extended panel data allows to confirm that transitions to an occupation generally tend to persist, with

[17] The data can be found at: `https://lucaschancel.com/enseignants/`.

[18] The method resembles the identification of outside options for teachers in Tsao (2025).

little support for the idea of observing transitory jobs. Teachers transition into a wide array of sectors and skill levels. The largest shares move into administrative jobs, either education-related (≈20% by $t + 3$ post-exit) or not (≈20%). Within administration, they take either public-sector positions (executive or management roles in government, financial services staff, or technical work for the State) or private-sector roles (general or sales administration, litigation, or personnel management). Within the category educational support staff, many become senior education counselors, supervising students without teaching duties in public or private institutions. About 10% move into higher education teaching or research, 5% into engineering, 5% into lower-level personnel management/secretarial/HR jobs, and 5% into adult training, typically in corporate settings. The remaining exits are dispersed across technician roles, managerial and other professional jobs, skilled labor, social work, domestic support, and other occupations.

The analysis reveals that a large share of exits (≈60%) are into other public-sector jobs. This pattern reflects the civil-servant status of French teachers, which facilitates mobility within the public sector subject to openings. Many transitions are therefore to occupations that, like teaching, are governed by rigid pay scales.

3. Estimation of teacher productivity as teacher value-added

To study heterogeneous elasticities to the wage gap by teacher productivity, I proxy productivity with teacher value-added (TVA), estimated using the Tartova (2023) network method. This approach is necessary because in France standardized testing occurs only at the end of 9th grade, making it impossible to use prior-year ($t - 1$) student scores to control for unobserved ability—a requirement for unbiased TVA estimation in standard approaches (see, e.g., Rockoff 2004; Kane & Staiger 2008; Chetty et al. 2014a). In what follows, I summarize the method.

Tartova (2023) develops a method for producing plausibly unbiased TVA estimates by exploiting cross-sectional variation in student test scores across subjects—applicable in my context, where students take standardized Math and French exams at the end of 9th grade—and by leveraging "networks" of teachers. The definition of a network is provided below:

DEFINITION (Teacher networks): *For a subset of classes $\mathcal{C}^n \subseteq \mathcal{C}$, a network of teachers is a subset of teachers $\mathcal{J}^n \subseteq \mathcal{J}$, for which for each class $c \in \mathcal{C}^n$, there exists at least one class $c' \in \mathcal{C}^n \setminus \{c\}$ such that the teacher for subject z in class c also teaches in class c':*

$$\exists c' \in \mathcal{C} \setminus \{c\},\ j(c, z) = j(c', z).$$

To illustrate the intuition behind teacher networks, I borrow Figure 3 from Tartova

(2023). Blue nodes represent Math teachers and yellow nodes—Literature teachers. An edge between a Math and a Literature teacher indicates that they have taught the same class at least once. Teachers belong to a network if each can be connected to each other through such classroom observations. For example, Math teachers M_1 and M_2 are both observed with Literature teacher L_1; comparing the Math–French score differences of their students isolates the relative TVA of M_1 and M_2, since L_1's contribution cancels out. Although M_2 and M_6 never share the same Literature teacher, they are connected through M_1 via L_1 and L_4, illustrating connectivity by transitivity. In this example, the number of classroom observations equals the number of teachers, allowing all teacher effects to be identified. Tartova (2023) shows formally that within any connected set of teachers (a "network"), the dimensionality problem is resolved and the OLS estimator is unique and unbiased in finite samples.

Estimation To estimate the TVA of Math teachers, I restrict the sample of teachers to the subset $\mathcal{J}^n$ and use the following OLS specification over the period 2007-2008 to 2021-2022:

$$\Delta A_i^{*MF} = \alpha + \mathbf{X}_i\beta + \sum_{m\in\mathcal{J}_M^n} \mu_m \mathbb{1}(j(M) = m) + \sum_{l\in\mathcal{J}_F^n} \mu_f \mathbb{1}(j(F) = f) + \varepsilon_i \tag{1}$$

In this equation, ΔA_i^{*MF} is the difference between the Math and French scores for student i. Importantly, μ_m are Math teacher fixed effects for teacher m, μ_f are French teacher fixed effects for teacher f. The intercept α captures the fact that students may be better in one of the two subjects. $\mathbf{X}_i$ are observable student and classroom characteristics. Specifically, I control for the socio-economic status (SES) of student i, their age, gender, and means-based scholarship, advanced classes taken, and exam repetitions. Classroom-level controls include all student-level controls averaged at the classroom-level, e.g. number of peers, average SES, average age, and percentage of female students. I include the same controls aggregated at the school level, and further add school-level lagged average test scores in Math and French. Finally, ε_i is the idiosyncratic error term. To reduce the noise in the estimates, I restrict the sample of classrooms to those with at least 5 students.

As teachers move across schools, the network of teachers $\mathcal{J}^n \subseteq \mathcal{J}$ spans across schools, connecting within-school networks with at least one mobile teacher. In the context of France where teachers often move across schools, more than 95% of teachers are connected in the same network and can therefore be compared. This leads me to a sample of roughly 16,600 Math and 21,200 French teachers under the age of 50.

The unbiasedness of the TVA estimates relies on the assumption that Math (French) teachers are not sorted to classrooms based on the relative Math-specific (French-specific) ability in these classrooms. While the assumption cannot be tested for unobservable student characteristics such as ability, Tartova (2023) performs tests of the assumption based on observable characteristics and shows evidence that while there seems to be sorting between teachers and students on common ability, there is little evidence of sorting on subject-specific ability at the middle school level.

As teachers, and especially less experiences ones, often teach few classrooms, I shrink TVA estimates using the Empirical Bayes shrinkage method outlined in Tartova (2023). The shrinkage method is outlined in Appendix C.2.

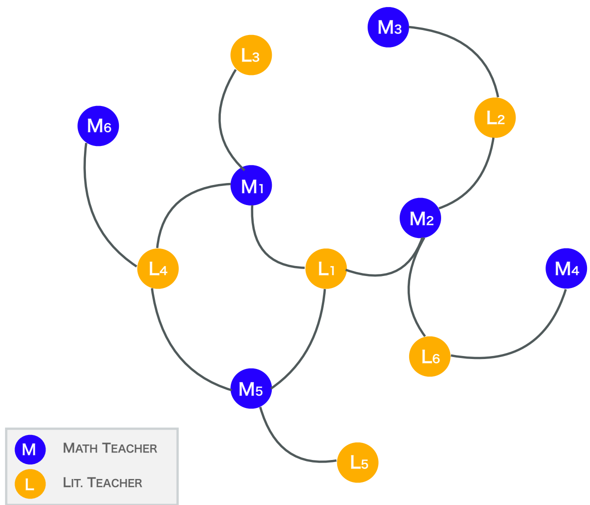


Figure 3: Example of a teacher network

Note: The figures represent an example of a network of teachers, for a set of Math and Literature teachers. The sequence of teachers is not specific to a single year. Each node represents a teacher - either in Math (blue) or in Literature (yellow). Each edge between two teachers signifies that the two teachers share at least one class. Source: Tartova (2023).

External validation of productivity proxy Tartova (2023) confirms that the method performs well in identifying the true teacher effect parameters in Monte Carlo simulations—in terms of standard deviation of the effect, the mean squared error from the true parameters, the rank correlation, and the statistical relationship with the true parameters, under a satisfied identification assumption of no sorting of students to teachers based on subject-specific student ability.

Using matched teacher-student data from New York City—where standard methods of TVA estimation can be applied as students are assessed with standardized exams annually—Tartova (2023) shows that the network method's estimates are comparable to the traditional methods' estimates (specifically, Kane & Staiger 2008) in realized data, too.

Standard deviation of TVA The estimated standard deviation (SD) of absolute TVA in France for the subsample of teachers below the age of 50, using the network method and exploiting the mobility of teachers across schools, closely resembles estimates found in the U.S. literature (see Table 1).

In Math, holding all else equal, a 1 SD better Math teacher leads to 0.13 SD higher test scores. This means that moving a student from a teacher who is at the 5th percentile of the teacher value-added distribution to someone who is at the 95th percentile would lead to a 0.44 SD increase in student exam scores. Given the distributions of Math test scores, this is equivalent to gaining 11 percentage points. Similarly, in French, a 1 SD better French teacher leads to 0.11 SD higher test scores, or a gain of 0.36 SD if we were to move a student from a teacher who is at the 5th percentile of the TVA distribution to one at the 95th percentile. Given the distributions of French test scores, this is equivalent to gaining 6.5 percentage points. By comparison, Chetty et al. (2014a) finds teacher effects in lower-secondary education in New York City in the magnitude of 0.13 SD in Math and 0.12 SD in Reading.

Table 1: SD of absolute TVA by subject

Note: This table reports the estimated standard deviation (SD) of TVA using the Network estimator in the French setting, measuring the absolute TVA of Math and French teachers teaching in the 9th grade at middle school prior to the age of 50. The SD is computed as the square root of σ_μ^2, where $\sigma_\mu^2 = Cov(\widehat{A}_{mt}, \widehat{A}_{mt'}) - \sigma_l^2$, such that $\sigma_l^2 = Cov(\widehat{A}_{lt}, \widehat{A}_{lt'})$. Underlying data source: DEPP, Bases Relais, Scolarité, Sysca, DNB, 2007-2008 to 2021-2022.

	Math	French
SD in TVA	0.133	0.108

Observable characteristics associated with the TVA estimates To get a better idea of what the TVA estimates represent, I run regressions of the TVA estimates on standardized observable teacher characteristics and school fixed effects. The coefficients of these regressions (shown in Figure B.3) provide two broad conclusions.

First, even though observables do not explain a lot of the TVA estimates (the R^2 in the individual regressions is small), there is a non-zero correlation with multiple characteristics. Specifically, Math TVA is positively correlated with experience, having a high qualification (Agrégation), being female, and having a higher pedagogical score during classroom observations by external examinators. For French, having more experience is not significantly correlated with TVA. Having a high qualification is positively associated with having a high TVA for French, by an order of magnitude similar to that for Math teachers. Being female and having a higher pedagogical score are also positively associated with French TVA. These results are broadly in line with existing literature, which generally finds either small or no correlation between TVA and having a higher qualification, and a positive correlation with experience at the beginning of one's career (see, for instance, Rockoff 2004; Rivkin, Hanushek, & Kain 2005; Kane, Rockoff, & Staiger 2008; Wiswall 2013; Jackson 2014; Bold et al. 2016; Bietenbeck, Piopiunik, & Wiederhold 2018; Bau & Das 2020).

Second, the level of correlation between TVA and the observables is different across Math and French. Particularly, Math TVA seems to be more correlated with every examined observable characteristic. This means that it is harder to pinpoint what makes a productive French teacher based on observable characteristics.

4. Teacher exit elasticity with respect to the wage

In this section, I present the empirical strategy used to identify the causal effect of teacher wages on exit from the profession and show both that the average teacher is highly responsive to changes in the wage, and that high-productivity teachers are three times more elastic to wages than low-productivity teachers.

4.1. Average exit elasticity with respect to the wage

4.1.1. Empirical strategy

To identify the causal effect of the change in the wage, I compare the exit rate of tenured teachers at REP+ to the exit rate of tenured teachers at REP schools in a difference-in-differences framework. My main empirical approach relies on estimating an Intention-to-Treat (ITT) effect, comparing exit rates of teachers assigned to treated schools (REP+) just before the reform to exit rates of teachers assigned to control schools (REP) just before the reform.

The advantage of this approach is that, given that treatment is defined before the reform, the identified effect obtained from the difference-in-differences framework is free of possible endogenous relative selection into treatment after the reform. Specifically, teachers whose labor supply is more elastic to wages would be more likely to select into treated schools because of the reform, for instance by moving from other type of schools after the reform. For that reason, if the relative composition in the treated schools changes towards more elastic teachers, the effect identified from a difference-in-differences framework that allows treatment to depend dynamically on the effective school a teacher teaches at in the post-period may be biased upwards reflecting the relative change in composition of teachers in treated schools toward more elastic teachers.

The ITT approach has two limitations stemming from the fact that it focuses only on incumbent teachers who were already in the profession in 2013. First, restricting the analysis to this subsample reduces statistical power. Second, the estimated coefficients may not reflect the responsiveness of the broader teacher population if responses to wage changes vary with experience or age. This is because the ITT approach focuses on incumbent teachers and ignores novice entrants that are, for instance, replacing retiring or exiting teachers. This is evident from the fact that post-reform teacher cohorts in the ITT sample exhibit a heavier concentration of older and more experienced teachers than the corresponding cohorts of teachers actually present in REP+ or REP schools (see Figure B.5).

Therefore, as robustness, after confirming that composition does not evolve differentially based on observable teacher characteristics over the first four years directly following the reform, I also estimate the Average Treatment Effect on the Treated (ATT), allowing for teachers' treatment status to vary over time as they move across school types. In other words, I take the real-time teacher allocation to a school in defining the treatment. This specification captures the effect of the wage increase among all teachers actually exposed to the reform in a given year. I show that the ATT estimates are in line with the ITT estimates presented below.

To identify the Intention-to-Treat (ITT) effect for the subsample of teachers who were already teaching in 2013–2014—the year the reform was announced, I define a time-invariant dummy variable, $\mathbb{1}(type_j = \text{REP+})$, equal to one if teacher j was employed in 2013–2014 at a school s that would later be classified by the reform as REP+ after 2014, and zero if teacher j was employed in 2013–2014 at a school that would later be classified as REP.[19] In other words, even if an incumbent teacher subsequently moved to a different school, they are considered treated according to their pre-reform assignment for the entire sample period. I exclude teachers who were teaching at ordinary schools in 2013-2014. The choice to define the treatment and control groups based on the 2013-2014 allocation of teachers, rather than based on earlier pre-reform years allocation, is to ensure the largest rate of compliance among teachers in the treated and control groups, as teachers may move across schools prior to the reform.

Formally, for the subsample of teachers j who are teaching at school s in commuting zone z in year in 2013-2014, I estimate the following event-study difference-in-differences specification for each year t:

$$exit_{jst} = \sum_{2013}^{2019} \eta_k \mathbb{1}(t = k) \times \mathbb{1}(type_j = REP+) + \gamma \mathbf{X}_{jt} + \theta_s + \lambda_{zt} + \varepsilon_{jst} \tag{2}$$

The outcome variable $exit_{jst}$ is a dummy equal to 1 in the last year before j exits (to reflect exit probability), and 0 otherwise, such that if a teacher is dropped out of the panel in the year they exit. The indicator $\mathbb{1}(t = k)$ equals 1 if the observation is in the academic year k—$k + 1$, and $\mathbb{1}(type_j = \text{REP+})$ is the treatment variable defined and described in the previous paragraph. The coefficients of interest, η_k, capture the causal effect of having been assigned before the reform to a school that later became REP+, and therefore received the disproportionate increase in wages, on a teacher's probability of exiting the profession, relative to teachers assigned to schools that became REP. I restrict the analysis to the 2019-2020 school year, which is the last year for which I can apply my definition of exit in my data, which finishes in 2021-2022.

[19] A school is considered as REP+ or REP, if it entered the REP+ or REP program in 2014-2015 or any subsequent academic year.

To account for unconditional differences in exit rates across teacher characteristics, I include a rich set of teacher-level controls $\mathbf{X}_{jt}$. Specifically, I include dummies for being a female teacher, having a high qualification ("Agrégation"), being in a specific experience tercile, being in a specific age tercile, and dummies for being either a Math teacher or a French teacher. I also account for time-invariant differences in exit probabilities across schools by including school fixed effects θ_s, which absorb average differences in exit rates between institutions. Finally, I include commuting zone-by-year fixed effects λ_{zt}, which control for shocks to local labor markets (such as changes in outside opportunities or demographics) that could influence teachers' exit decisions. As an alternative for local labor markets, I use regional-educational-authority-by-year (i.e. *académie*-by-year) fixed effects λ_{rt}. This implies that my specification is comparing teachers in REP+ and REP schools that are in the same geographical location. I report differences in relative exit rates relative to the base 2013-2014 school year, as the year during which the reform (and wage increase) was announced.

Identification assumption The validity of the difference-in-differences design relies on the parallel trends assumption, which requires that, absent the reform, exit rates for REP+ and REP teachers would have followed similar trajectories after 2013-2014. The parallel trends assumption cannot be tested in the pre-period within the ITT specification, because my differences-in-differences estimation sample starts in the last year before the reform—the base year, in which treatment is defined. This is because, by construction, any teacher observed in a treated or control school in 2013-2014 cannot have in previous years—treatment is selected based on survival until 2013-2014, thus leaving the dependent variable to always be zero in both treated and control groups prior to 2013-2014.[20]

To test the identifying assumption, I re-define the definition of the treatment variable by using each teacher's last observed assignment in the pre-reform period, irrespective of the school year of that last observation between 2009 and 2013. This alternative definition of treatment allows to estimate equation 2 over the period 2009-2019, as the observed pre-reform relative exit rates reflect the relative annual exit rates of the population of teachers at REP+ schools at a given year, relative to the population of teachers at REP schools in that year.

To probe the identifying assumption, I re-define the treatment variable using each teacher's last observed pre-reform assignment (2009–2013). This static grouping

[20] In fact, given the definition of my exit variable which entails that a teacher has exited from the profession for at least two years, the treated and control groups would both have a dependent variable of zero for the two academic years preceding 2013-2014, but may have a non-zero exit dummies prior to that. This implies mechanically zero exit rates in 2011–2012 and 2012–2013 for both treatment and control groups. Including these pre-periods would therefore mechanically generate negative coefficients for 2011–2012 and 2012–2013—since exit rates are measured relative to the 2013–2014 base year. Although some teachers may have been employed prior to 2011–2012, exit observations before that period would only reflect the exit rate of a small subset of teachers who exited and later re-entered the profession, rather than consistent attrition patterns.

("last-pre-reform REP+" vs "last-pre-reform REP") allows me to estimate event-study coefficients from 2009 to 2019 and to assess pre-trends without the mechanical zeros induced by conditioning on presence in 2013–2014. I also report robustness checks using assignment in 2011–2012.

Importantly, pre-reform estimates compare exit rates between cohorts defined by eventual pre-reform assignment, not the contemporaneous populations teaching in REP+/REP in each year. In the robustness test that identifies the ATT, for which one can compare the exit rates for teachers at REP+ schools against that of REP schools, I also estimate the relative pre-trend in exits given the contemporaneous assignment of teachers to the treatment and control groups.

Across these approaches, I find a small linearly increasing differential pre-trend, which would indicate that teachers at REP+ schools were exiting disproportionately more and more compared to teachers at REP schools prior to the reform (see Figure B.6 for the alternative ITT specifications). This structural break in the relative probability of exit at the time of the reform is consistent with a causal impact of the policy, rather than a continuation of pre-existing trends. Under the assumption that, absent the reform, the relative exit rate at REP+ versus REP schools would have continued to increase linearly as in the pre-period, my baseline difference-in-differences estimates can be interpreted as a lower bound on the true post-reform effect. I account for this more thoroughly in the robustness test with the ATT specification.

4.1.2. Results

I begin by estimating equation 2 for the sample of all incumbent tenured teachers at treated and control schools in the year prior to the reform without controls and fixed effects, before progressively introducing controls and fixed effects to demonstrate that the results are robust to different specifications. As shown in Figure 4, moving progressively from specifications with (i) treatment period dummy and treatment status dummy but no controls and no fixed effects, (ii) including school and year fixed effects, (iii) further adding teacher controls, (iv) adding regional-educational-authority-by-year fixed effects, or as an alternative to those, (v) adding commuting-zone-by-year fixed effects, does not significantly impact the coefficients η_k.

Overall, the estimates across all specifications point to a persistent and sizable effect of the relative wage increase for REP+ teachers. The reform led to an average reduction in exits from the teaching profession in the post-period, compared to the last pre-reform year, of about 0.5 percentage points for REP+ teachers relative to their REP counterparts. This effect is substantial, corresponding to roughly a 25 percent decline relative to the counterfactual exit rate these teachers would have experienced in the absence of the

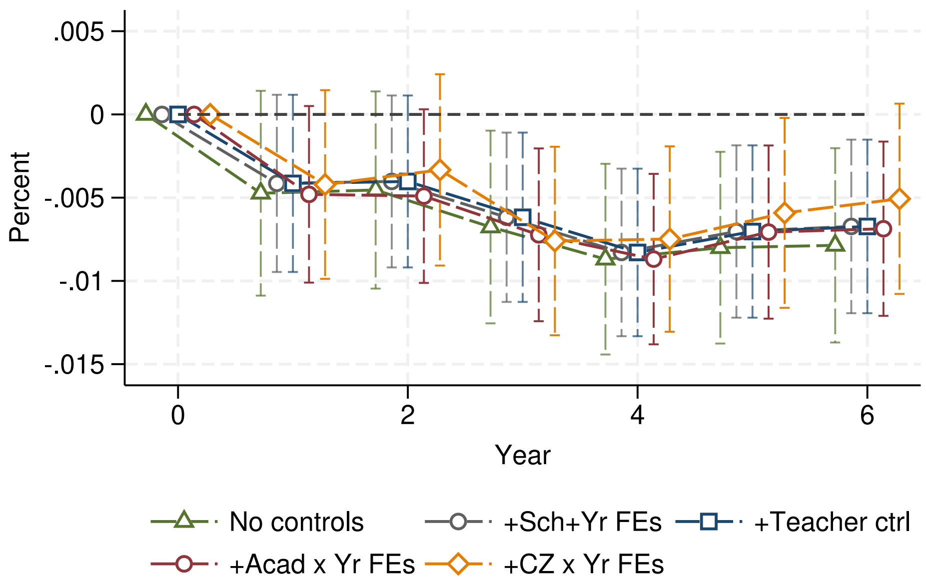


Figure 4: Differential exit probability for the average REP+ teacher compared to REP

Note: The figures present the η_k coefficients of the event study difference-in-difference regressions 2, respectively, sequentially adding control variables and fixed effects, for the sample of all REP+ and REP teachers. Year 0 represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, tercile of teaching experience, tercile of age, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2009-2010 to 2019-2020.

reform (see Table 2).[21]

From reduced-form estimates to exit elasticities with respect to the wage The reduced-form ITT estimates indicate that teachers initially assigned to treated schools, which are associated with a relative wage increase after the reform, respond by exiting the teaching sector less than teachers initially assigned to control schools.

To assess the magnitude of this relative response, I translate the reduced-form coefficients to exit elasticities with respect to the wage. To do so, I exploit the school-type-specific wage bonus amount generated by the reform as exogenous variation in the teaching wage under a Two-Stage Least Squares approach (2SLS).

To align the timing of exit decisions with wage expectations, I code the treatment in year t using the bonus schedule of year $t+1$ whenever the full scope and scale of that schedule (i.e. the list of targeted schools and the corresponding amounts) were known before the

[21]The counterfactual exit rate is defined as the average exit rate that initially assigned REP+ teachers would have experienced in the post-period, had they not experienced a differential increase in wages compared to initially assigned REP teachers, under the assumption of no differential trend in exit rates between the two groups after 2013. Specifically, the counterfactual exit rate (hereafter $\overline{exit}_{REP+}$) is constructed from the estimates of $exit_{jst} = \alpha + \beta_1 Post_t \times REP+_j + \beta_2 Post_t + \beta_3 REP+_j + \varepsilon_{jst}$, such that $\overline{exit}_{REP+} \equiv \alpha + \beta_2 + \beta_3$, where α represents the pre-reform exit rate for initially assigned REP teachers, β_2 the change in the exit rate for initially assigned REP teachers post-reform, and β_3 the pre-reform difference in exit rates between initially assigned REP+ and REP teachers.

interim period between t and $t+1$.[22] This is because, as explained above, I define an exit in school year t (e.g. 2014–2015) as equal to 1 if a teacher is observed for the last time in t and thus leaves the system between school years t and $t+1$. Because an effective exit between t and $t+1$ (recorded at t) is likely to be based on information revealed by or during the interim period between t and $t+1$ (e.g. the summer break) about wages in $t+1$, it is important to align the timing of wage information with this decision. Because each increase effective in $t+1$ was known to teachers by the interim period between t and $t+1$, this coding choice is equivalent to shifting the policy schedule one year earlier in the analysis. The resulting instrument is plotted in Figure B.4.

In the first stage of the 2SLS approach, I regress the school-level bonus $Bonus_{jst}$ for teacher j at school s in year t on the interaction of the post-reform indicator with the REP+ treatment, controlling for the same observable teacher characteristics as in the baseline reduced-form specification:

$$Bonus_{jt} = \beta Post_t \times \mathbb{1}(type_j = REP+) + \gamma \mathbf{X}_{jt} + \theta_s + \lambda_{zt} + \varepsilon_{jst} \tag{3}$$

The variable $Post_t$ takes the value of 0 in 2013-2014, and 1 between 2014-2015 and 2019-2020.

In the second stage of the 2SLS approach, I regress the teacher exit dummy on the fitted values of the bonus from the first stage, with the same set of controls:

$$exit_{jst} = \eta \widehat{Bonus}_{jt} + \gamma \mathbf{X}_{jt} + \theta_s + \lambda_{zt} + \varepsilon_{jst} \tag{4}$$

The η coefficient in regression 4 represents the average impact of the bonus on the probability of teacher j to exit after the reform. Because the interaction term in the first stage is based on an indicator for the post-reform period, my two-stage approach does not mechanically condition exits in year t on the realized bonus in year t, but instead conditions exits in the post-period to the average bonus in the post-period. This is done in order to reduce noise in the estimates and to avoid assuming a narrow dependence of exit on the contemporaneous effective bonus.

I compute the elasticity of exit with respect to the wage as:

$$\varepsilon_{exit} = \eta \cdot \frac{\overline{Wage}_{REP+,CF}}{P(exit)_{REP+,CF}} \tag{5}$$

where $\overline{Wage}_{REP+,CF}$ is the estimated average intended counterfactual wage for treated

[22] For example, the first differential bonus increase for REP+ teachers was implemented in 2015–2016. While the size of the wage increase was already known in 2013–2014, the complete list of treated schools only became available during 2014–2015. This implies that the full treatment was known to teachers in 2014–2015, so that the 2015–2016 bonus would affect exit decisions made at the end of the 2014–2015 school year.

teachers in the post-reform period, had it not been for the differential reform, and $P(exit)_{REP+,CF}$ is the average counterfactual exit rate of treated teachers.[23]

Table 2: Responsiveness to the wage increase for the average teacher

Note: The table presents the η coefficients from regressions 2 and 4, for the average teacher, as well as the first-stage regression results from regression 3. Column (1) presents the preferred reduced-form specification. Columns (2) presents the corresponding 2SLS specification. Column (3) shows the first stage of the 2SLS. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate for REP+ teachers, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4). The corresponding relative percentage decrease in exit rates and elasticity with respect to the wage (computed using equation 5) are reported below. The bonus is expressed in thousands of euros. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	Reduced-form	2SLS	First stage
Post × REP+	-0.00544**		1.19526***
	(0.00230)		(0.00300)
Bonus		-0.00453**	
		(0.00224)	
Constant	0.0196***		
	(0.00137)		
Teacher controls	✓	✓	✓
School FEs	✓	✓	✓
CZ x Year FEs	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148403	148403	148403
R-squared	0.050	0.001	0.836
First-stage F stat		171503	
Counterfactual exit	0.022	0.022	
As perc. of exit	-0.247		
Counterfactual wage		26.888	
Elasticity to wage		-5.533	

Table 2 presents the results. Using the observed $P(exit)_{REP+,CF}$ and $\overline{Wage}_{REP+,CF}$ (which are reported at the bottom of the regression table), I compute the elasticity ε_{exit} implied by the estimated η (equation 5). The estimated elasticity of –5.5 implies that a 1% increase in wages reduces teachers' exit probability by approximately 5.5%, indicating a strong negative responsiveness of exits to changes in wages.

As some teachers that are treated based on their 2013-2014 assignment may move out of treated schools after the reform is implemented, I use the compliance rate in order to identify the local average treatment effect (LATE) for teachers who were initially assigned

[23] Similarly to the counterfactual exit rate explained above, the counterfactual wage is defined as the average teacher wage that initially assigned REP+ teachers would have experienced in the post-period, had they not experienced a differential increase in wages compared to initially assigned REP teachers, under the assumption of no differential trend in wages between the two groups after 2013 for any other reason than the reform. Specifically, the counterfactual wage (hereafter $\overline{Wage}_{REP+,CF}$) is constructed from the estimates of $Wage_{jst} = \alpha + \beta_1 Post_t \times REP+_j + \beta_2 Post_t + \beta_3 REP+_j + \varepsilon_{jst}$, such that $\overline{Wage}_{REP+,CF} \equiv \alpha + \beta_2 + \beta_3$, where α represents the pre-reform average wage for initially assigned REP teachers, β_2 the change in the wage for initially assigned REP teachers post-reform, and β_3 the pre-reform difference in wages between initially assigned REP+ and REP teachers.

to the REP+ schools and who stayed in these schools after the reform was implemented (see Table A.3). The average compliance rate of 77.5% over the period 2014–2015 to 2019–2020 implies an exit elasticity of –7 for teachers who remained in REP+ schools after the reform was introduced, and thus actually received treatment.[24]

Although not perfectly comparable, my estimates are broadly consistent with turnover elasticities in the literature—typically ranging between -3 and -6 (e.g., Benhenda & Sims 2022; Clotfelter et al. 2008; Cowan & Goldhaber 2018; Falch 2011, 2017; Feng & Sass 2018; Hendricks 2014). Turnover combines both exits from the profession and school-to-school mobility; the latter is often more constrained by limited vacancies (i.e., by limited labor demand), so turnover elasticities are expected to be smaller in magnitude than exit elasticities.[25]

4.1.3. Robustness to an alternative definition of treatment: ATT approach

To address the limitations of the ITT approach—notably, statistical power and potential lack of representativeness due to focusing only on incumbent teachers as discussed above, I estimate the Average Treatment Effect on the Treated (ATT), comparing the exit rates of teachers effectively at REP+ to those of teachers effectively at REP schools before and after the reform. Formally, the treatment time-invariant dummy variable $\mathbb{1}(type_s = \text{REP+})$ is a time-invariant dummy equal to 1 for a school s that was eventually classified as REP+ after 2013-2014, and 0 if a school was classified as REP. A school is considered as REP+ or REP, if it entered the REP+ or REP program in 2014-2015 or any subsequent academic year. Treated teachers are therefore teachers who are in school for which $\mathbb{1}(type_s = \text{REP+}) = 1$ at a given year t. This specification captures the causal effect of the wage increase among all teachers actually exposed to the reform in a given year, rather than only incumbent teachers present prior to the reform as for the ITT case.

As discussed, while this approach captures the effective treatment, the estimated wage responsiveness of teachers may be biased upwards if more wage-sensitive teachers move into treated schools. To mitigate this concern, I test for structural breaks in the relative composition in treated versus control schools around the reform based on observable teacher characteristics. Figure B.7 depicts the coefficients of such regressions in an event-study difference-in-differences framework comparing the observable characteristics of teachers at treated schools with those of teachers at control schools. I find no evidence of differential compositional changes with respect to average years of experience, share of female teachers and average pedagogical score of teachers in the

[24] To get to this elasticity, I rescale the 2SLS coefficient by 1/compliance.

[25] Even though mobility decisions do not represent pure labor supply decisions in my framework, I aggregate exits and transfers out of treated schools to measure turnover for comparability to the literature. I obtain an ITT turnover elasticity of -2.3, which translates to a LATE elasticity of -3 (see Table A.4).

school between the two groups in the four years immediately following the reform. Because the average teacher experience at REP+ becomes significantly higher than that at REP schools from 2018, I restrict the analysis until 2017 and I further show robustness to restricting the sample until 2015, thereby further quantitatively limiting the concerns of endogenous selection into treatment.

Formally, for each teacher j teaching at school s in commuting zone z in year t I estimate the following event-study difference-in-differences specification:

$$exit_{jst} = \sum_{2009}^{2017} \eta_k \mathbb{1}(t = k) \times \mathbb{1}(type_s = REP+) + \gamma \mathbf{X}_{jt} + \theta_s + \lambda_{zt} + \varepsilon_{jst} \tag{6}$$

As in the ITT, the outcome variable $exit_{jst}$ is a dummy equal to 1 in the last year before j exits (to reflect exit probability), and 0 otherwise, such that if a teacher exits, she is dropped out of the panel. The indicator $\mathbb{1}(t = k)$ equals 1 if the observation is in the academic year k—$k + 1$, and $\mathbb{1}(type_s = \text{REP+})$ is the treatment variable. I exclude teachers at ordinary schools from the analysis. The coefficients of interest are the η_k, which trace the causal effect of being a teacher at a treated REP+ school relative to being a teacher at a REP school on the probability of exiting. I include the same set of teacher-level controls $\mathbf{X}_{jt}$, school fixed effects θ_s and commuting zone-by-year fixed effects λ_{zt} (or regional-educational-authority-by-year fixed effects λ_{rt}), as in the ITT analysis. The base year in this regression is once again the 2013-2014 school year. As treatment is not defined only on incumbent teachers in 2013, η_k can be estimated in the pre-reform period, too.

The results depict a strong decrease in exit rates for REP+ teachers relative to REP teachers following the introduction of the wage increase, after a clear structural break of the previously increasing relative average exit rate at REP+ schools at the time of the reform (Figure B.8). Specifically, an initial persistent linear increase in the relative exit rates at REP+ prior to the reform is interrupted in the year the reform was announced, followed by a significant decrease in the relative exit rate at REP+ schools after the introduction of the reform.

As discussed in the previous subsection, the pre-period is characterized by a differential positive trend in the relative exit rates at REP+ (as seen in Figure B.8), meaning the pre-trends are not parallel. While the structural break in the relative probability of exit around the time of the reform indicates a causal impact of the policy, one cannot rule out that this differential positive trend would have continued in the post-period, thus entailing that my estimates are a lower bound in absolute terms of the true effect.[26]

As a result, I propose two alternative assumptions. The first one is akin to the parallel trends assumption: I assume that the relative evolution of exits would have stopped in 2013-2014 had it not been for the reform. This assumption is conservative if one believes that school conditions at REP+ schools were likely on a deteriorating trend, prompting an increasing relative exit rate.

Second, I alternatively assume that the relative evolution of exits would have continued increasing linearly had it not been for the reform. This assumption prompts the use of a de-trended event-study difference-in-differences strategy, following Bhuller, Havnes, Leuven, and Mogstad (2013), Goodman-Bacon (2018, 2021), and Dobkin, Finkelstein, Kluender, and Notowidigdo (2018). For details on the estimation strategy, see Section C.3. Figure C.15 also provides a graphical representation of the η_k^{DT} coefficients, the coefficients from the de-trended regressions, plotted against the η_k coefficients from regression 6 for the three samples of interest. As one would expect, the de-trended coefficients suggest an even stronger average impact of the treatment, albeit not statistically different from the non-de-trended coefficients η_k.

Under the assumption that the ATT pre-trend would have stopped in 2013-2014, the ATT estimates indicate a relative decrease in exit rates for the average treated teacher at a magnitude of 0.7 percentage points on average, which, given the higher counterfactual exit rate for this treated group, also translates to a reduction of 24% (see Table A.5). Instead, under the assumption that the relative pre-trend would have continued in the post-period, the ATT estimates indicate a 1pp relative decrease in exit rates (30% of the counterfactual exit rate for the treated group).

The corresponding 2SLS estimates imply an elasticity to the wage of between -8 and -10, under the two different identification assumptions. It is important to note that these estimated elasticities cannot be directly compared to the ones shown in Table 2 for the ITT, as they are estimated over the short- or medium-run, given the compositional changes identified after year four after the introduction of the bonus. Thus, I compare the ATT elasticities to the ITT elasticities, adjusted for compliance, estimated over the same time period (column (6) of Table A.5). The implied exit elasticity to the wage from the ITT over the same time period, adjusted for compliance, is of a comparable magnitude of -9.8. Overall, the ATT results are consistent in magnitude with the ITT results.

[26] As a first attempt to account for these differences, I include a rich set of pre-reform school characteristics, averaged over 2009–2010 to 2013-2014, and interacted with year dummies, $\sum_{k=2009}^{2017} \omega_k \mathbb{1}(t = k) \times \overline{\mathbf{X}}_{s,2009-2013}$. These controls are intended to capture the possibility that heterogeneous pre-reform school environments (specifically, student socio-economic composition, scholarship incidence, share of repeaters) drive differential exit dynamics over time. However, including these controls does not eliminate the diverging pre-trends, suggesting that unobserved factors correlated with REP+ assignment remain (Figure B.11).

4.1.4. Ruling out the role of the differential change in statutory teaching hours

Since the organizational and pedagogical aspects of the reforms were largely identical across REP and REP+ schools, the difference in the wage increase treatment after 2015 remains the key treatment between the treated and control groups.

As discussed in Section 1, an additional intended treatment for REP+ teachers was the re-weighting of teaching hours. The intended treatment was to reduce the number of teaching hours by 1.6 hours for a full-time regular-qualification (Capes) teacher with a teaching load of 18 hours per week, and reallocate the 1.6 hours to student follow-up and liaison with families. For a high-qualification (Agrégation) teacher with a teaching load of 15 hours, this reallocation was intended to be 1.4 hours.

In Section 1, I provide anecdotal evidence from official reports and teachers' unions communications that the intended reduction in teaching hours was not realized due to a lack of teaching staff to cover the pre-reform teaching needs under the new minimum statutory hours. Instead, the reform led to an effective increase in working hours, due to the additional obligations after the reform.

Consistent with that, I show that the number of hours that teachers spent actually teaching decreased by far less than suggested by the reform (see Table A.6). Specifically, I show that the average REP+ Capes teacher prior to the reform teaches on average more than the mandated 18 hours—teaching instead 18.7 hours, and that these teaching hours decrease by only 0.7 (instead of 1.6) hours after the reform was introduced.[27] The average REP+ high-qualification (Agrégation) teacher also teaches on average more than the mandated 15 hours prior to the reform—teaching instead 16.1 hours with overtime work, which decrease by only 0.5 hours post-reform.

It follows that this leg of the reform actually increased overall working hours (by approximately 0.9 hours), instead of simply reallocating teaching hours, given the additional obligation of hours spent in liaison with students and parents.

If incentives to work overtime remained unchanged at the time of the reform, these additional hours can be regarded as involuntary, since teachers who preferred to work longer hours could already do so before the reform by voluntarily taking on overtime. Hence, the post-reform wage increase due to this overtime pay could not be viewed as a pure or unconditional bonus—it required teachers to work beyond their pre-reform workload. This interpretation implies that the estimated exit elasticity with respect to the wage increase may represent a lower bound (in absolute value) of the true elasticity, as some teachers may have been pushed out of REP+ schools due to the increased number of hours they had to work.

[27]Note that the results are very similar for the average high- and low-TVA teacher in the analysis in the next subsection 4.2.

If incentives to work overtime changed at the time of the reform, one cannot rule out teachers may have decided to work more hours voluntarily. Detailed computation of the overtime pay rule (showed in Appendix C.4) points that the reform led to a small increase in monetary incentives for the average Capes teacher, of approximately 100 euros annually. Assuming that teachers voluntarily decided to work more to obtain these 100 euros, this amount can be considered an additional effective bonus for REP+ teachers. Under this assumption, the baseline ITT exit elasticity to the wage of -5.5 computed using the bonus would instead be -5.1.[28] This lower bound of the elasticity in absolute terms, under the assumption the average teacher decides to work more rather than being obligated to, is quantitatively very close to the baseline elasticity.

Lastly, I rule out the importance of the residual concerns that the reform-led small reallocation of teaching hours towards other forms of school service may have impacted teacher exit decisions, for example by attracting better students and therefore improving the school environment. I show that adding school-by-year or one- or two-year lagged school-by-year controls for student composition to the baseline regression does not quantitatively change the baseline ITT estimates (see Table A.7). While these controls may be considered "bad controls" as their variation in the post-reform period is endogenous to the reform, the stability of the estimates across these alternative specifications provides evidence that my results are not driven by observable changes in the time-varying student composition.

4.1.5. Robustness tests

I show the baseline ITT results are robust to different sample restrictions. Specifically, the results are robust to alternative definitions of the exit variable—restricting to no re-entry for at least three, rather than two, years, excluding schools which opened after 2013-2014 or shut down prior to 2019-2020, excluding teachers who are novice in 2013-2014, and extending the sample of teachers to a less conservative retirement age cut—58 years old, instead of 50 years old (see Table A.8).

In addition, I show that the baseline ITT results are robust to different definitions of the treatment schools (see Table A.9). More precisely, the results are robust to defining as treated teachers only the sample of teachers who were at a school in 2013 that became a REP+ school in 2014 (i.e. the first 77 REP+ schools in Metropolitan France), or only in 2014 or 2015. In addition, they are robust to excluding from treatment the teachers in schools that became REP+ in 2014 specifically (the first announced treated group).

Finally, I also show that the results are robust to taking different subsamples of teachers—Capes-certified only, or Agregation-certified only, though the latter result is economically larger but statistically insignificant due to small sample size (Table A.10).

[28] To get to this back-of-the-envelope computation, note that the first stage coefficient in Table 2 would change to 1.19526+0.1=1.29526, leading to a second-stage η coefficient of -0.00544/1.29526=-0.0042.

4.2. Heterogeneous elasticities to the wage by TVA

4.2.1. Empirical strategy

To study the heterogeneous impact of the wage bonus on teachers with different levels of productivity, I characterize j as a high-TVA teacher in a subject m (Math or French, for which TVA measures can be estimated), if their TVA is above the subject-specific median of the TVA distribution across all teachers with estimated TVA who are observed in the pre-period, $HighTVA_{jm}$.

My empirical strategy relies on estimating the effects of the reform predicted by the ITT specifications 2 and 4 for the subsamples of high- and low-TVA teachers, in order to identify which subsample responds more strongly to the wage increase. In robustness, I also present the corresponding ATT results and show they produce comparable qualitative and quantitative conclusions. In robustness tests, I furthermore confirm the stability of the differential effects identified for high- and low-TVA teachers after accounting for potential differential treatment effects across other observable dimensions which are correlated with TVA.

Table 3: Responsiveness to the wage increase by TVA group

Note: The table presents the η coefficients from regressions 2 and 4 by TVA group. Columns (1) and (3) present the preferred reduced-form specification for the high- and low-TVA subsamples, respectively. Columns (2) and (4) present the corresponding 2SLS specification for the high- and low-TVA subsamples, respectively. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate for REP+ teachers, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4). The corresponding relative percentage decrease in exit rates and elasticity with respect to the wage (computed using equation 5) are reported below. The bonus is expressed in thousands of euros. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	High TVA		Low TVA	
	Reduced-form	2SLS	Reduced-form	2SLS
Post × REP+	-0.0127**		-0.00284	
	(0.00537)		(0.00485)	
Bonus		-0.0107**		-0.00245
		(0.00503)		(0.00528)
Constant	0.0127**		0.0326***	
	(0.00603)		(0.00664)	
Teacher controls	✓	✓	✓	✓
School FEs	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	24,059	24,059	21,619	21,619
R-squared	0.121	0.001	0.143	0.002
First-stage F stat		20,301		16,067
Counterfactual exit	0.022	0.022	0.018	0.018
As perc. of exit	-0.568		-0.162	
Counterfactual wage		26.813		26.786
Elasticity to wage		-12.870		-3.740

4.2.2. Results

The reduced-form ITT results for high- and low-TVA teachers are plotted in Figure 5. Panel (a) shows a strong and significant decrease in the exit probability of high-TVA teachers initially assigned to REP+ schools, compared to high-TVA teachers initially assigned to REP schools, after the introduction of the reform. Specifically, the relative decrease in exit rates for high-TVA teachers is at an average magnitude of 1.3pp over the first six years following the introduction of the reform—a reduction of 57% compared to the counterfactual exit level for this subgroup (see Table 3).

As shown in Panel (b), the exit rates of low-TVA teachers did not significantly decrease with the introduction of the reform (with the exception of 2017-2018). The average effect is smaller in magnitude and insignificant: the relative decrease in exit rates for low-TVA teachers is at a magnitude of 0.3pp—a reduction of roughly 16% compared to the counterfactual exit level for this subgroup (see Table 3).

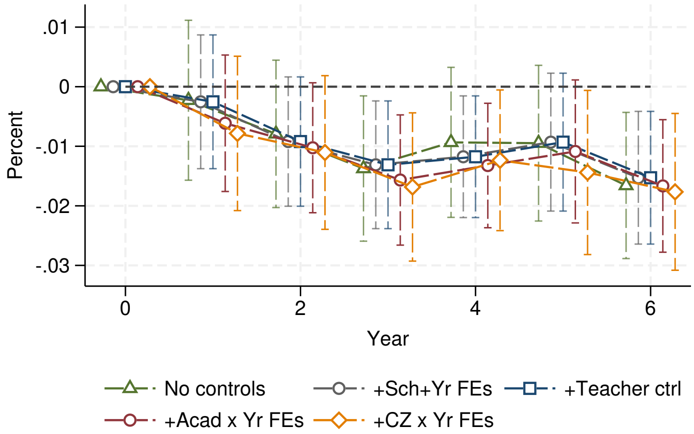
From reduced-form estimates to exit elasticities with respect to the wage Using the same 2SLS approach that leverages the differential intended bonus for treated teachers, as outlined in Section 4.1, I examine the heterogeneity in exit elasticities with respect to the wage by teachers' level of productivity. Columns (2) and (4) of Table 3 presents the results for the subsamples of high-TVA and low-TVA teachers. [29] The results indicate that high-TVA teachers are significantly more responsive to the wage than their low-TVA counterparts. Specifically, the estimated elasticity for high-TVA teachers stands at -13 (statistically significant), compared to -4 for low-TVA teachers (not statistically significant). In other words, high-productivity teachers are more than 3 times more responsive to wage changes than their lower-productivity counterparts. This relative result remains after accounting for the compliance rate in each subgroup (see Table A.3). Specifically, the compliance rates of the high-TVA group (79.6%) and the low-TVA group (76.9%) imply elasticities of -16.4 and -4.7, respectively, for high- and low-TVA teachers who remained in REP+ schools after the reform was introduced.

The larger elasticity (in absolute value) for high-TVA teachers suggests that a uniform wage increase is particularly effective at retaining more productive teachers and thus at increasing the aggregate quality of the teaching force through the exit channel. I confirm this in my theoretical framework in Section 5.

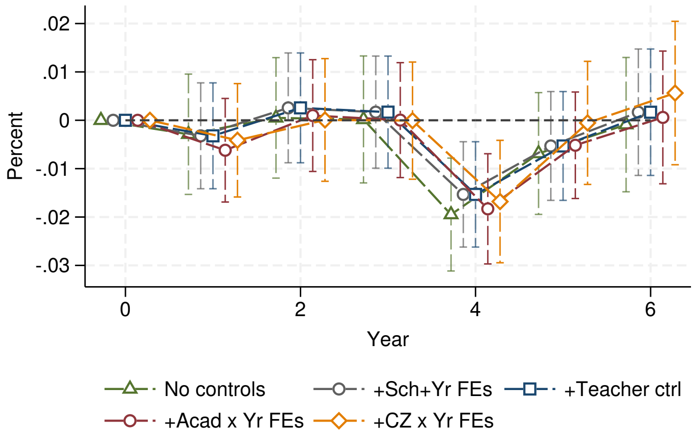
4.2.3. Robustness to an alternative definition of treatment: ATT approach

I further implement the ATT empirical strategy outlined in the previous subsection, comparing the exit rates of teachers effectively at REP+ to those of teachers effectively at REP before and after the reform. While ITT and ATT estimates are consistent for

[29] For reference, Table A.12 shows that the exit elasticity using the ITT 2SLS approach for the average Math and French teachers (not conditioning on TVA) is -8.5.



(a) High TVA



(b) Low TVA

Figure 5: Differential exit probability for REP+ teachers compared to REP by TVA

Note: The figures present the η_k coefficients of the event study difference-in-difference regression 2, sequentially adding control variables and fixed effects, for the sample of (a) high TVA and (b) low TVA REP+ and REP teachers. Year 0 represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

the average teacher, the limitations of the ITT design—such as the reduced sample of incumbents and their greater average experience—may be more pronounced within subsamples defined by teacher productivity. Estimating the ATT for high- and low-TVA teachers therefore serves as a robustness check, allowing me to verify that the observed heterogeneity is not driven by differential representativeness of incumbent teachers across productivity groups.

The reduced-form ATT results for high- and low-TVA teachers (plotted in Figure B.9) show that the identified ATT effects are consistent with the ITT results. Panel (a) shows that there is an increasing linear pre-trend for the subsample of high-TVA teachers, indicating that high-TVA teachers were exiting REP+ schools more and more, relatively to REP teachers' exit from REP schools, prior to the reform. This indicates that the estimated effects may be considered a lower bound of the true effects in absolute terms for high-TVA teachers if the increasing differential pre-trend would have continued, had it not been for the introduction of the bonus. By contrast, there is no differential pre-trend identified for low-TVA teachers.

The ATT estimates point that treated high-TVA teachers responded to the differential wage increase by decreasing their exit from the profession at a magnitude of between 1.5pp and 2pp compared to the control group—a reduction of between 64% and 69% in the medium-run compared to the counterfactual exit rate (Table A.11). For low-TVA teachers, the estimated reduction in exit rates relative to the control group ranges between 10% and 17% (not statistically significant).

The corresponding 2SLS estimates point to a medium-run exit elasticity to the wage of between -21 and -23 for high-TVA teachers (and -24 in the short-run), compared to the ITT medium-run exit elasticity for compliers of -23. For low-TVA teachers, the ATT medium-run exit elasticity lies between -3 and -6 (with an estimate of -7 in the short-run). To put this in perspective, the medium-run elasticity estimated with the ITT specification, adjusted for compliance, is -7, showing that both estimation strategies lead to the same conclusions. Irrespective of the specification, high-TVA teachers are economically and statistically significantly more responsive to the wage increase than low-TVA teachers, at a magnitude of between 2 and 7 times that of low-TVA teachers.

4.2.4. Robustness of the heterogeneous effect to additional fixed effects

I show the results are robust to the inclusion of more granular fixed effects. To summarize the differential effects in a parsimonious way, I estimate the augmented triple-interaction model by interacting the treatment-by-post term with the high-TVA indicator $HighTVA_{jm}$, and by allowing all controls and fixed effects to vary flexibly with this indicator:

$$\begin{aligned} exit_{jst} = \beta_1 \, \text{Post}_t \times \mathbb{1}(type_j = \text{REP+}) \times HighTVA_{jm} + \dots \\ + (\mathbf{X}_{jt} \times HighTVA_{jm})\delta + (\theta_s \times HighTVA_{jm}) + (\lambda_{zt} \times HighTVA_{jm}) + \varepsilon_{jst} \end{aligned} \tag{7}$$

where all lower-level interactions are included in the regression (and are excluded here for readability).

I then sequentially add subject-by-high-TVA fixed effects, treatment-by-year fixed effects, and school-by-year fixed effects. I repeat the same exercise for the ATT specification. The results, presented in Table A.13, yields similar conclusions. Both ITT and ATT estimates of the differential responsiveness of high-TVA teachers remain economically large and stable, even though the ITT coefficient is less precisely estimated due to the additional noise introduced by the smaller sample size and the temporary dip in low-TVA exits in 2017–2018.

This is visible from the event-study coefficients corresponding to the column (5) in Table A.13—the specification which includes teacher controls, interacted with the high-TVA dummy, school-by-high-TVA fixed effects, and school-by-year fixed effects. The coefficients of interest, η_k, are displayed in Figure B.14. They capture the dynamic causal effect of the relative wage increase for high-TVA teachers compared to their low-TVA counterparts and point to a statistically significant disproportionate reduction in exit rates for high-TVA, compared to low-TVA treated teachers, relative to the control group, for all years between 2015-2016 and 2019-2020 (with the exception of 2017-2018).

4.2.5. By how much can observables explain this heterogeneous effect?

To assess whether the heterogeneous treatment effects by teacher productivity are driven by specific observable characteristics that are correlated with TVA, I estimate an extended version of the ITT specification allowing for differential treatment effects across these dimensions. This exercise tests whether the observed heterogeneity by TVA reflects a broader productivity gradient rather than being driven by one particular characteristic, such as teacher experience or age, thereby supporting the generalizability of the results to contexts where the correlates of productivity may differ.

Specifically, in the version of regression 7 with the richest fixed effects, I sequentially introduce analogous interactions with other observable characteristics, such as terciles of experience, terciles of age, certification type (a dummy for high qualification), and gender. For instance, I estimate:

$$\begin{aligned} exit_{jst} = {} & \beta_1 \, \text{Post}_t \times \mathbb{1}(type_j = \text{REP+}) \times HighTVA_{jm} + \dots \\ & + \beta_2 \, \text{Post}_t \times \mathbb{1}(type_j = \text{REP+}) \times Experience_{jt} + \dots \\ & + (\mathbf{X}_{jt} \times HighTVA_{jm})\delta + (\mathbf{X}_{jt} \times Experience_{jt})\gamma \\ & + (\theta_s \times HighTVA_{jm}) + (\lambda_{zt} \times HighTVA_{jm}) \\ & + (\zeta_m \times HighTVA_{jm}) + (\phi_{st} \times HighTVA_{jm}) + \varepsilon_{jst}. \end{aligned} \tag{8}$$

where once again all lower-level interactions are included in the regression (and are excluded here for readability).

Table A.14 depicts the coefficients from these alternative specifications, while Table A.15 shows the corresponding analysis for the ATT specification. The estimated heterogeneous ITT treatment effects by TVA remain stable across specifications, indicating that the stronger responsiveness of high-TVA teachers is not driven by correlated characteristics such as experience, age, certification, or gender.

4.2.6. Other robustness tests

Finally, I also show that the results are robust to alternative definitions of the $HighTVA_{jm}$ dummy. Specifically, I redefine the $HighTVA_{jm}$ around the median within commuting zone z, or alternatively around the median within school, in the pre-2013 period (see Table A.16). The results for both the high- and low-TVA groups are qualitatively and quantitatively unchanged.

5. Discrete-choice framework of teacher labor supply

In this section, I rationalize the empirical result that high-productivity teachers are more elastic to wages in their decision to leave the profession in a simple discrete-choice framework of labor supply. I quantify the aggregate teacher quality effect of different counterfactual policies that impact the wage gap—the gap between wages in outside options for teachers and their teaching wage. I show that while uniform changes in teacher wages can increase aggregate teacher quality, targeted increases have a larger impact at a lower fiscal cost. I further demonstrate the asymmetry of results on teacher quality from counterfactual policies that increase teachers' outside options, which arises under thin-tailed, unimodal distributions of relative preferences for the outside option.

5.1. Setup

Discrete choice over two options I consider a simple static discrete choice model in which individuals choose between two options—teaching ($d = 0$) or an outside option ($d = 1$).

Heterogeneity by productivity Each individual is characterized by their teaching productivity c (which could be proxied, for instance, by TVA for individuals who are already in teaching). For simplicity, I assume that productivity is discrete, such that there are only two types of productivity—high and low, $c \in \{H, L\}$. Each individual observes their own idiosyncratic preferences and makes a choice that maximizes their utility.

Indirect utility The indirect utility associated with each option $d \in \{0, 1\}$ is given by:

$$U_d^c = V_d^c + e_d^c,$$

where V_d^c is the deterministic part of utility, and e_d^c is an idiosyncratic taste shock.

Distributional asssumption for e_d^c The idiosyncratic taste shocks e_d^c are i.i.d. and follow a Gumbel distribution:

$$e_d^c \sim \text{Gumbel}(\mu^c, \sigma^c), \quad c \in \{H, L\},\ d \in \{0, 1\}.$$

Without a loss of generality, I assume $\mu^c = 0$.

Linear utility in wages and amenities I assume utility from teaching ($d = 0$) is a linear function of the teaching wage w_0 and preferences for non-wage amenities q_0^c:

$$V_0^c = w_0 + q_0^c,$$

and utility from the outside option ($d = 1$) as:

$$V_1^c = w_1^c + q_1^c,$$

where w_1^c is the wage in sector $d = 1$, and q_1^c is the non-wage preference for the outside option $d = 1$. Importantly, w_0 is not a function of productivity—teachers are not paid according to their productivity. The non-wage preference for teaching q_0^c can be thought of as pro-sociality—a teachers' preferences for the non-wage amenities associated with teaching, and q_1^c can be thought of as one's motivation for being in outside option $d = 1$. Without a loss of generality, I assume $q_1^c = 0$.

5.2. Predictions

Exit probabilities Define the wage gap between an individual's outside option and teaching as:

$$\Delta w^c = w_1^c - w_0,$$

An individual chooses the outside option ($d = 1$) if:

$$U_1^c > U_0^c \quad \Leftrightarrow \quad \Delta w^c - q_0^c > e_0^c - e_1^c.$$

For the i.i.d. Gumbel-distributed e_0^c and e_1^c with common scale σ_e^c, the difference $\varepsilon^c \equiv e_1^c - e_0^c$ follows a logistic distribution with mean 0 and scale σ_e^c. Hence, the probabilities of exit for H and L, respectively, are given by:

$$P_1^H = \frac{1}{1 + \exp\left(\frac{q_0^H - \Delta w^H}{\sigma_e^H}\right)} \text{ and } P_1^L = \frac{1}{1 + \exp\left(\frac{q_0^L - \Delta w^L}{\sigma_e^L}\right)} \tag{9}$$

For each type, the exit probability P_1^c increases with the wage gap Δw^c, meaning that when outside wages rise relative to teaching wages, individuals are more likely to leave teaching. Conversely, the exit probability decreases with non-wage amenities associated with teaching q_0^c: individuals with greater pro-social motives are less likely to exit unconditionally.

The level of the exit probability P_1^c depends on the scale parameter σ_e^c, which captures how heterogeneous teachers are in their relative preferences of the outside option. When σ_e^c is small, teachers of a given type value their outside options relatively similarly, so choices are driven mostly by systematic differences in wages and amenities: individuals with slightly better outside options almost surely leave, and those with slightly worse ones almost surely stay. By contrast, when σ_e^c is large, relative preferences are more dispersed, and idiosyncratic taste shocks dominate utility differences, making choices appear more random and pushing P_1^c toward 0.5 regardless of wages and amenities.

Exit semi-elasticities with respect to the wage gap For a change in the wage gap Δw^c, these baseline probabilities for H and L change by:

$$\frac{\partial P_1^H}{\partial \Delta w^H} = \frac{1}{\sigma_e^H} \cdot P_1^H \cdot (1 - P_1^H) \text{ and } \frac{\partial P_1^L}{\partial \Delta w^L} = \frac{1}{\sigma_e^L} \cdot P_1^L \cdot (1 - P_1^L) \tag{10}$$

The teacher-type-specific responsiveness to a change in the wage gap increases as the dispersion of idiosyncratic taste shocks (the scale parameter σ_e^c) decreases, holding the equilibrium exit probability constant. As discussed, when σ_e^c is small, teachers' relative preferences are tightly clustered around their mean, making their responses more coordinated. For two groups with the same exit probability, a smaller σ_e^c implies that the mean relative preference $(\Delta w^c - q_0^c)$ must lie closer to zero—the indifference threshold—so that a similar share of individuals remains at the margin. In this case, many teachers have relative preferences close to the cutoff, and even a small change in the wage gap moves a large fraction across the decision boundary, generating a strong aggregate response. Conversely, when σ_e^c is large, preferences are more dispersed and the density near the threshold is lower, so responsiveness to wage changes is weaker.

Teacher responsiveness to a change in the wage gap also depends on the equilibrium probability of exit, holding the dispersion of idiosyncratic taste shocks constant. The exit probability P_1^c reflects the share of individuals whose relative preference for the outside option, $\Delta w^c - q_0^c + \varepsilon^c$, exceeds zero. Because the logistic distribution of ε^c is symmetric and most dense around zero, the largest mass of individuals lies near the indifference threshold when P_1^c is at 0.5, that is when $\Delta w^c = q_0^c$.

When P_1^c is below 0.5—that is, when the expected relative preference for the outside option is negative $(\Delta w^c < q_0^c)$—most teachers prefer to stay, and only those with large positive idiosyncratic shocks exit. As the wage gap rises, the mean of the distribution $\Delta w^c - q_0^c + \varepsilon^c$ shifts rightward, bringing a growing share of individuals close to the indifference threshold. The density of "marginal" teachers therefore increases, and the responsiveness of exit probabilities to wage changes becomes larger as P_1^c approaches 0.5.

Once P_1^c exceeds 0.5—when the expected relative preference for the outside option becomes positive $(\Delta w^c > q_0^c)$—most teachers already prefer the outside option, and the mass of individuals near indifference begins to decline. Further increases in the wage gap then move fewer teachers across the threshold, as there are fewer and fewer "marginal" teachers, and responsiveness to additional increases in the wage gap falls. Hence, responsiveness is increasing in P_1^c for $P_1^c < 0.5$ and decreasing for $P_1^c > 0.5$, reaching its maximum when the largest share of teachers are "marginal".

Effect of changes to the wage gap for teaching quality Average teacher quality can be written as the weighted average of the two types in the workforce,

$$Q = \frac{\nu_H T_H + \nu_L T_L}{T_H + T_L},$$

where ν_c is the average TVA of teachers of type c, and T_c is the number of employed teachers of that type. This can be expressed more simply as

$$Q = \nu_L + (\nu_H - \nu_L)s_H,$$

where s_H is the share of high-type teachers among those who remain in teaching.

Because the number of employed teachers of each type equals the number of potential teachers of that type times their probability of staying, $T_c = m_c P_0^c$, the share of high types can be written as

$$s_c = \frac{m_c P_0^c}{m_H P_0^H + m_L P_0^L},$$

where P_0^c is the probability of staying in teaching (one minus the exit probability P_1^c).

Define the semi-elasticity of exit for type c as:

$$\eta_c \equiv \frac{1}{P_1^c}\frac{\partial P_1^c}{\partial \Delta w^c} = \frac{1}{\sigma_e^c}(1 - P_1^c) = \frac{P_0^c}{\sigma_e^c}$$

Differentiating average quality Q with respect to the wage gap yields

$$\frac{dQ}{d\Delta w} = (\nu_H - \nu_L)\frac{ds_H}{d\Delta w} = (\nu_H - \nu_L)s_H P_1^H s^L P_1^L(\eta_L - \eta_H)$$

As high-TVA teachers have by definition a higher average TVA than low-TVA teachers, $(\nu_H > \nu_L)$ the sign of the quality change depends on the sign of the relative exit responsiveness of the two types to the wage gap, $\eta_L - \eta_H$. Because an increase in the teaching wage reduces the wage gap ($d\Delta w = -dw_0$), the sign of the quality change following in the teacher wage depends on the sign of the $\eta_H - \eta_L$.

Hence, raising teacher pay increases average teacher quality if high-type teachers are more responsive to wages than low-type teachers.

5.3. Estimating the framework

Identification strategy Estimating the exogenous parameters from this discrete choice framework (Δw^H, Δw^L, q_0^H, q_0^L, σ_e^H, σ_e^L) allows me to run policy counterfactuals on changes in the wage gap and to simulate their effect on aggregate quality.

I show that these exogenous parameters can be estimated using moments from the data and the identified semi-elasticities with respect to wages obtained from the 2SLS ITT approach. Specifically, I use five empirical moments and one identifying restriction to recover the six exogenous parameters, as detailed below.

The first and second empirical moments correspond to the baseline probabilities of exiting the teaching sector for each teacher type, $c \in \{H, L\}$. I equate the equilibrium exit probability equation for each type to the ex-ante exit probabilities observed for that type in REP+ schools in 2013 (see P^H_{exit} and P^L_{exit} in Table 4).

The third and fourth empirical moments correspond to the responsiveness of the exit probability to wages for each teacher type, $c \in \{H, L\}$. I equate $\frac{\partial P^c_{\text{exit}}}{\partial \Delta w^c}$ to the estimated η coefficients for high- and low-TVA teachers obtained from regression 4 (see Table 3).

The fifth empirical moment corresponds to the type-specific wage gaps. To obtain these, I estimate the average wage gap directly from the data by combining information on teacher wages with the weighted average wages in the outside occupations into which teachers are observed to transition. These estimates are constructed using an exhaustive matched employer–employee panel, which allows me to identify teachers and track their occupational choices after leaving the teaching sector (for details on the computation of the wage gap, see Appendix C.5). While I cannot match these data to my main teacher panel and therefore cannot identify the outside-option wage separately for high- and low-TVA teachers, I can identify the average outside-option wage and thus the average wage gap for the representative teacher. The fifth moment equates this average wage gap, estimated for the average REP+ teacher in 2013, to the weighted linear combination of the exogenous type-specific wage gaps, where the weights are given by the relative exit probabilities of each teacher type. Formally,

$$\Delta \overline{w} \;=\; \omega^H \Delta w^H \;+\; \omega^L \Delta w^L \text{ where } \omega^c \;=\; \frac{P^c_{\text{exit}}}{P^H_{\text{exit}} + P^L_{\text{exit}}}, \quad c \in \{H, L\}.$$

To exactly identify the six parameters in the framework, I assume that the non-pecuniary preferences for teaching are identical across types. More formally,

$$q^H_0 \;=\; q^L_0 \;=\; q_0.$$

This assumption is in line with recent literature on teacher preferences (see Johnston in press), which underlines that high- and low-TVA teachers have similar preferences for non-wage amenities. Earlier research related to public-sector workers (Bó et al. 2013) provides alternative evidence highlighting there may be a positive correlation between TVA and non-pecuniary preferences for teaching. Note that relaxing this assumption to allow for a positive correlation between c and q_0^c would lead to a higher Δw^H and a lower Δw^L, compared to the ones produced by this assumption, thus predicting a stronger impact on aggregate teacher productivity. Therefore, the framework predictions reached with this assumption can be considered a lower bound.[30]

Under this assumption, the estimated η^H and η^L, together with the estimate of the average wage gap $\Delta\overline{w}$ and the ex-ante exit probabilities P^H_{exit} and P^L_{exit}, constitute sufficient statistics for recovering the impact of a change in the wage gap on aggregate teacher productivity induced by exit responses. The resulting parameter estimates from the framework are reported in Table 4.

Table 4: Framework parameters

	Symbol	**Value**
Inputs		
High TVA exit probability	P^H_{exit}	0.023
Low TVA exit probability	P^L_{exit}	0.021
η (High TVA)	$\frac{\partial P^H_{\text{exit}}}{\partial \Delta w^H}$	0.0107
η (Low TVA)	$\frac{\partial P^L_{\text{exit}}}{\partial \Delta w^L}$	0.00245
Average wage gap	$\Delta\overline{w}$	4.3
Outputs		
Non-pecuniary preferences	$q_0^H = q_0^L = q_0$	23.8
Scale parameter (High TVA)	σ_e^H	2.1
Scale parameter (Low TVA)	σ_e^L	8.39
Wage gap (High TVA)	Δw^H	15.9
Wage gap (Low TVA)	Δw^L	-8.4

Discussion of the identified parameters Table 4 reveals that, in the context of the model, the empirical patterns—the higher responsiveness of high-TVA teachers at REP+ schools to changes in the wage gap and the similar baseline exit probabilities for high- and low-TVA teachers—can only arise from a specific joint configuration of the underlying parameters. Because both groups exhibit similar exit probabilities, the pairs $(\Delta w^c - q_0^c, \sigma_e^c)$ must jointly generate the same position on the logistic curve.

This means that if one group (high-TVA teachers) has a smaller dispersion of idiosyncratic shocks—i.e., a smaller σ_e^c—its mean relative preference for the outside option must lie closer to the indifference threshold to reproduce the same observed exit rate.

High-TVA teachers indeed exhibit a much smaller dispersion of preferences ($\sigma_e^H = 2.1$ versus $\sigma_e^L = 8.39$), implying a steeper logistic slope because in the data they are more responsive at the same probability of exit. To maintain this same exit probability while having a smaller dispersion of preferences, their mean relative preference for the outside option ($\Delta w^H - q_0^H = -7.9$) must therefore lie closer to zero—the point of expected indifference—than that of low-TVA teachers (-32.2).

[30] To see this, let $q_0 \neq q_0^{c,true}$, such that $q_0^{L,true} < q_0 < q_0^{H,true}$. Then, a H-type chooses to leave teaching if $\Delta w^{H,true} > q_0 + (q_0^{H,true} - q_0) + e_0^H - e_1^H$, or $\Delta w^{H,true} - (q_0^{H,true} - q_0) > q_0 + e_0^H - e_1^H$. Under the assumption of equal non-wage preferences, $\Delta w^H = \Delta w^{H,true} - (q_0^{H,true} - q_0)$. It follows that $\Delta w^{H,true} > \Delta w^H$.

This configuration is illustrated in Figure 6. The distribution of relative preferences for high-TVA teachers (blue curve) is narrower and centered closer to the indifference threshold, implying a dense mass of "marginal" individuals who are most sensitive to wage changes. By contrast, the distribution for low-TVA teachers (red curve) is flatter and shifted further away, meaning fewer teachers lie near indifference and exit responses to wage changes are correspondingly weaker.

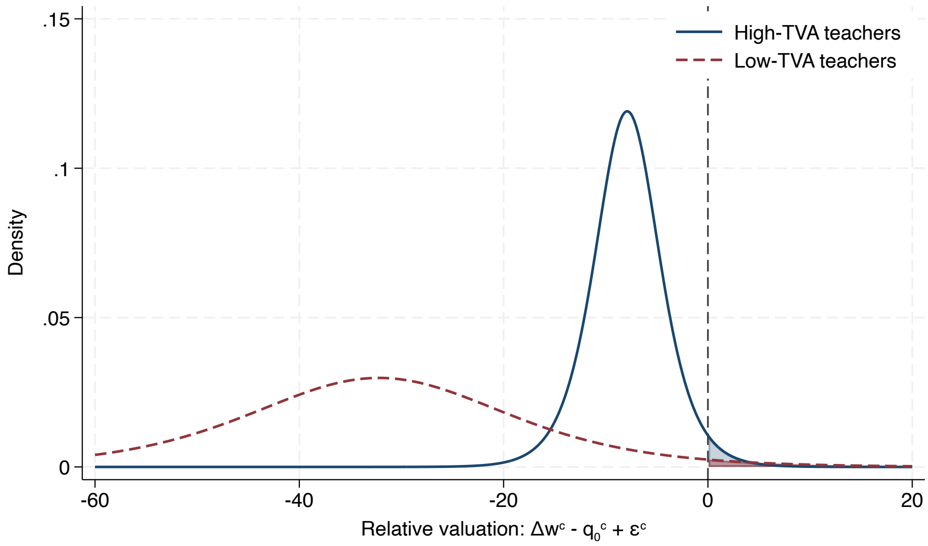


Figure 6: Equilibrium relative valuations for high- and low-TVA teachers

Note: The figure plots the implied distributions of the relative valuation of the outside option, $\Delta w^c - q_0^c + \varepsilon^c$, for high- and low-TVA teachers, assuming logistic idiosyncratic shocks with parameters reported in Table 4: $\mu_H = -7.9$, $\sigma_e^H = 2.1$, and $\mu_L = -32.2$, $\sigma_e^L = 8.39$. The vertical dotted line at zero represents the indifference threshold between staying in teaching and exiting. The shaded area corresponds to the fraction of teachers for whom $\Delta w^c - q_0^c + \varepsilon^c > 0$, i.e., those whose total relative valuation favors leaving the teaching sector.

5.4. Policy counterfactuals

Simulated effects of changes in the wage gap on average teaching quality Having estimated the parameters of the model, I can now use them to simulate policy counterfactuals and assess how changes in the wage gap Δw affect the aggregate quality of the teaching workforce.

To evaluate the implications of teacher labor supply choices for average teaching quality, I define average quality at time t as the weighted average of type-specific productivity:

$$Q_t = s_{H,t}\,\nu_H + s_{L,t}\,\nu_L,$$

where $s_{c,t}$ is the share of type-c teachers ($c \in \{H, L\}$) among those employed at time t, and ν_c denotes the average productivity (TVA) of type c.

Assuming no entry or replacement between periods, the share of each type in period $t+1$ equals its initial share multiplied by the probability of remaining in teaching, normalized by the total probability of remaining across types:

$$s_{H,t+1} = \frac{s_{H,t}\, P_0^H}{s_{H,t}\, P_0^H + s_{L,t}\, P_0^L}, \qquad s_{L,t+1} = \frac{s_{L,t}\, P_0^L}{s_{H,t}\, P_0^H + s_{L,t}\, P_0^L},$$

where $P_0^c = 1 - \widehat{P_{\text{exit}}^c}$ is the predicted probability that a type-c teacher remains in teaching for a given wage gap Δw^c.

Average quality at $t+1$ under a given policy scenario is then:

$$Q_{t+1}^{\text{policy}} = s_{H,t+1}^{\text{policy}}\, \nu_H + s_{L,t+1}^{\text{policy}}\, \nu_L,$$

and under the baseline (no policy change):

$$Q_{t+1}^{\text{baseline}} = s_{H,t+1}^{\text{baseline}}\, \nu_H + s_{L,t+1}^{\text{baseline}}\, \nu_L.$$

The effect of the policy on aggregate teacher quality is therefore measured as the difference in the change in average quality across the two scenarios:

$$\Delta Q_{t+1} \equiv (Q_{t+1}^{\text{policy}} - Q_t) - (Q_{t+1}^{\text{baseline}} - Q_t) \;=\; Q_{t+1}^{\text{policy}} - Q_{t+1}^{\text{baseline}}$$

Because $\nu_H > \nu_L$, this policy effect depends on how the relative retention of high- and low-type teachers responds to the change in the wage gap, that is, on the difference in their exit semi-elasticities $(\eta_H - \eta_L)$.

To simulate the dynamics of quality over multiple periods, I repeat this updating process iteratively. For each horizon $k > 1$, the new type shares are obtained by applying the same retention probabilities P_0^c from the corresponding scenario to the shares from the previous period. This recursion is repeated for both the policy and baseline scenarios, using the same set of estimated retention probabilities for each scenario. Aggregate quality at horizon $t+k$ is then computed under each scenario and the cumulative policy effect is given by the difference between Q_{t+k}^{policy} and $Q_{t+k}^{\text{baseline}}$.

Uniform vs. targeted increase in teacher wage I first consider a policy counterfactual in which teacher wages increase uniformly by the same magnitude as the reform, holding the outside-option wage constant. Specifically, REP+ teachers got on average 1,600 euros more annually as part of the policy, relative to REP teachers, and this increase was permanent. Under this treatment the exit probability of high-TVA teachers falls from 2.3% to 1.1% (a decrease of 53%), while that of low-TVA teachers falls from 2.1% to 1.7%—a decrease of 17% (see Table 5). This translates, assuming no replacement, into an increase in average teaching quality of 1.6% of a standard deviation in TVA (relative to 2013 levels) over 5 years, and 3.2% of a SD over 10 years, relative to a counterfactual with no change in the wage gap, assuming no replacement of exiting teachers.

Table 5: Counterfactual outcomes

Policy counterfactual	% Δ P_1^H	% Δ P_1^L	ΔQ_{t+5}	ΔQ_{t+10}
Uniform teacher wage ↑ (1.6k)	-53%	-17%	1.6% SD	3.2% SD
Targeted teacher wage ↑ (1.6k)	-53%	0%	2.3% SD	4.5% SD
Targeted teacher wage ↑ (3.2k)	-78%	0%	3.3% SD	6.6% SD
Uniform outside wage ↑ (1.6k)	+109%	+20%	-4% SD	-7.8% SD
Targeted outside wage ↑ (1.6k)	+109%	0%	-4.7% SD	-9.4% SD
Targeted outside wage ↑ (3.2k)	+324%	0%	-14.3% SD	-27.4% SD

I also consider counterfactuals where the bonus to teacher wages was directed only towards high-TVA teachers. In a counterfactual in which the whole sum dedicated to the uniform bonus was directed towards high-TVA (i.e., unchanged cost for the government but different allocation of resources), the gains in TVA are double those of the uniform bonus. Importantly, even if the government was to halve its costs and only distribute the 1.6k increase, the gains in TVA are larger than those observed with the uniform bonus—indicating an increase in TVA compared to a counterfactual with no change in the wage gap of 2.3% of a SD over 5 years and 4.5% over 10 years.

Thus, the framework predicts that introducing targeted wage bonuses to high-TVA teachers can increase average TVA by more than a uniform wage increase—and for half the price.

This is because targeted bonuses directly raise the retention incentives of the most productive teachers, who are also more responsive to wage changes. By contrast, a uniform wage increase raises pay for all teachers, including those with lower productivity and weaker responsiveness, thereby diluting its impact on the composition of the teaching workforce. In the targeted case, resources are concentrated where the marginal return in terms of retention and aggregate quality is highest: among high-TVA teachers who are most likely to adjust their labor supply decisions in response to pay. As a result, the same budget increase yields a larger improvement in average teaching quality.

The modest effects on average TVA suggest that only a small share of teachers are marginal in the context of a wage gap decrease—that is, teachers whose exit decisions would change in response to an increase in the teaching wage. Although high-productivity teachers are more responsive to wage incentives (as reflected in their higher exit elasticities), the overall low baseline exit rates for both groups imply that most teachers are either clear stayers or clear leavers. Consequently, "always leavers"—those who would exit regardless of wage changes—remain quantitatively more important than the relatively small group of marginal teachers whose decisions are sensitive to pay, which limits the aggregate impact of wage increases on teaching quality.

Uniform increase in teacher wage vs. uniform increase in outside-option wage As a further counterfactual, I simulate a uniform increase in the outside-option wage by the same amount of 1.6k. Under this scenario, the exit probability of high-TVA teachers rises from 2.3% to 4.8% (an increase of 109%), and the probability of low-TVA teachers rises from 2.1% to 2.5% (an increase of 20%). Assuming no replacement, this translates into a decrease in average teaching quality of 4% of a SD over five years and 7.8% over 10 years, relative to a counterfactual with no wage gap change.

It follows that the effect of an increase in the wage gap is asymmetric to the effect of a decrease in the wage gap, under logistic distributions of relative preferences. This follows directly from the low baseline probabilities of exit for both types (below 0.5). When the baseline exit probability P_1^c is low, most teachers already prefer to remain in teaching, and only a small share are close to indifference. A further decrease in the wage gap (i.e. a wage increase for teachers) therefore affects few individuals, as most teachers are already on the staying side of the distribution. By contrast, an increase in the wage gap (i.e. a rise in outside-option wages) shifts the distribution toward the region with a higher density of individuals near the indifference threshold, where small changes in relative preferences induce many teachers to switch from staying to leaving (i.e. to become marginal).

As a result, even modest increases in the outside-option wage produce disproportionately large increases in exit probabilities, while equivalent decreases in the wage gap yield comparatively small retention gains. This non-linearity is a direct implication of the logistic structure of choice probabilities and the low baseline exit rates observed in the data: when most teachers are far from indifference, wage decreases have little effect, but wage increases can push a larger mass of marginal teachers past the exit threshold. These conclusions follow for any thin-tailed, unimodal distribution of relative preferences.

Uniform vs. heterogeneous increase in outside-option wage for high TVA Finally, to reflect the idea that improvements in outside wages may not be uniform across the skill distribution, I consider a counterfactual where the outside-option wage increases only for high-TVA teachers.

I first consider a 1.6k increase in the wage gap for high-TVA teachers. In this case, the exit probability of high-TVA teachers rises from 2.3% to 4.8% (a 109% increase). The corresponding effect on quality is a decrease by 4.7% of a SD in five years and 9.4% over 10 years, assuming no replacement of exiting teachers.

A larger increase in the wage gap for high-TVA teachers, of 3.2k, leads to disproportionately larger effects on aggregate quality, as it shifts the distribution even closer toward the region with a higher density of individuals near the indifference threshold. Specifically, the predicted effect on aggregate quality is a decrease of 14.3% of a SD over five years, and 27.4% over 10 years, following a sharp increase in the exit probability of high-TVA teachers to 9.8% (a 324% increase).

Discussion of counterfactual results Taken at face value, these simulations yield three main insights about how wage policy shapes the composition and quality of the teaching workforce through the exit channel.

First, a uniform increase in teacher wages raises average TVA in the teaching sector because high-TVA teachers are more responsive to wage changes than their low-TVA counterparts. In the model, this difference in responsiveness arises from heterogeneous exit semi-elasticities: high-TVA teachers have higher η_H than η_L, meaning their retention probability increases more strongly following a rise in the teaching wage. As a result, a uniform pay increase shifts the composition of the teaching workforce toward high-productivity teachers, even though it raises wages equally for all.

Second, targeting wage increases toward high-TVA teachers is a substantially more cost-effective way to improve aggregate quality. Concentrating pay incentives where the marginal response is largest—among teachers whose exit decisions are most sensitive to wages—generates a stronger composition effect for a given budget. This is because the exit elasticity of low-TVA teachers is higher than zero in absolute terms. Thus, in equilibrium, the targeted policy raises average TVA by more than the uniform one, and even at half the fiscal cost.

Third, in an environment where outside-option wages are increasing relative to teacher wages, failing to adjust teacher wages accordingly can erode teaching quality over time. When labor market conditions improve outside of teaching, the opportunity cost of remaining in teaching increases disproportionately for high-TVA teachers. Without compensating adjustments in teacher wages, retention among the most productive teachers can fall substantially, leading to a gradual and potentially large decline in the average TVA of the teaching workforce. This insight helps rationalize why teacher quality may stagnate or decline in periods of rising economy-wide wages: even small and gradual increases in outside opportunities can shift a non-negligible mass of marginal high-TVA teachers out of teaching.

6. Mechanism

The framework in Section 5 can be micro-founded within a model of rational inattention (see e.g., Matějka & McKay 2015; Sims 1998, 2003). I propose that teachers face uncertainty about their outside-option wage w_1^c and choose how much "attention" θ^c to allocate to learning about it. Acquiring more precise information is costly, but it reduces the noise in decision-making when comparing teaching with outside options. Because the value of information rises with the stakes of the decision, teachers with a larger expected outside-option wage—that is, with higher wage gap Δw^c—optimally allocate more attention to learning about outside opportunities, which leads to greater precision when making a decision. This implies that the scale parameter σ_e^c is inversely related to attention, θ^c, which itself increases with the wage gap Δw^c. As a result, teachers with on average higher expected wage gaps optimally acquire more information about outside options, which leads to a lower σ_e^c for this group of teachers—they are able to take the decision to leave with more precision.

This micro-foundation yields a direct implication that I can test in the data. Because attention (and thus responsiveness) is increasing in the outside-option wage, teachers with larger ex-ante wage gaps—those whose outside options pay more relative to teaching—should exhibit stronger behavioral responses to a policy that narrows that gap.

Within the framework in Section 5, the higher responsiveness of high-TVA teachers could thus be rationalized by them facing a higher wage gap. This would be both consistent with the fact that the estimated framework implies a higher expected wage gap for high-TVA teachers and the broader idea that competitive labor markets reward better more productive workers.

While I cannot credibly isolate that this mechanism is causing the higher responsiveness of high-TVA teachers as I do not observe the wage gap per TVA type, I can test empirically whether, following the introduction of the bonus at REP+, exit probabilities fall more sharply among teachers observed to have higher baseline wage gaps.

I use the estimated average wage gaps defined within a cell of individuals of a given sex g, age group a, commuting zone z, year t, and position on the teacher wage scale ("echelon") e (for more details on computation, see Section C.5). The estimated wage gap is increasing in age, decreasing in the teacher's position on the wage scale (and relatedly, experience), and lower for female teachers (see Table A.17).

Specifically, for each teacher j at school s in year t, I run triple-interaction difference-in-differences regressions of the sort:

$$\begin{aligned} exit_{jst} = \beta\, \text{Post}_t \times \mathbb{1}(type_j = \text{REP+}) \times HighWageGap_j + \dots \\ + (\mathbf{X}_{jt} \times HighWageGap_j)\delta + (\theta_s \times HighWageGap_j) + \phi_{st} + \varepsilon_{jst} \end{aligned} \tag{11}$$

where all lower-level interactions are included in the regression (and are excluded here for readability). The dummy variable $HighWageGap_j$ equals 1 for teachers belonging to the top tercile of the wage-gap distribution in 2013, and 0 for teachers belonging to the bottom tercile.

The average REP+ teacher with a higher baseline wage gap is significantly more responsive to the increase in teacher wages than counterparts with a lower baseline wage gap, both economically and statistically (see column (1) in Table 6). Repeating the exercise for the subsamples of STEM and non-STEM teachers reveals that the magnitude of the triple-interaction coefficient β is even larger for the subsample of STEM teachers (see columns (3) and (5) in Table 6). One possible explanation is that STEM teachers are more likely to take the higher-paying outside options out of the set of identified outside occupations for the sample of all teachers.

I then decompose the baseline wage gap into its two components—the outside option wage and the teacher wage—to examine which dimension of heterogeneity better

Table 6: Triple difference-in-differences of exit on the wage gap

Note: The table presents the β coefficients from regressions 11 (columns 1, 3 and 5) and equivalent regressions with the two subcomponents of the wage gap (columns 2, 4 and 6). Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. All teacher controls are interacted with the the relevant wage dummy or dummies. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	All		STEM		Non-STEM	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × REP+ × High Wage Gap	-0.00531*		-0.0171***		-0.00653*	
	(0.00307)		(0.00639)		(0.00383)	
Post × REP+ × High Outside Wage		-0.00821**		-0.0239**		-0.0111***
		(0.00371)		(0.0101)		(0.00428)
Post × REP+ × High Wage		0.00199		0.0146		0.00247
		(0.00433)		(0.0116)		(0.00491)
Post × REP+	-0.000485	-0.00407	-0.0164	-0.0144	0.00314	-0.00767
	(0.00741)	(0.00911)	(0.0175)	(0.0307)	(0.00918)	(0.0115)
Teacher controls	✓	✓	✓	✓	✓	✓
School × High Wage Gap FEs	✓	–	✓	–	✓	–
School × High Outside Wage FEs	–	✓	–	✓	–	✓
School × High Wage FEs	–	✓	–	✓	–	✓
School × Year FEs	✓	✓	✓	✓	✓	✓
Observations	81,571	78,575	15,828	15,663	58,701	55,390
R-squared	0.121	0.142	0.301	0.386	0.146	0.176
Baseline exit (REP+ × High Wage Gap)	0.029		0.032		0.028	
Baseline exit (REP+ × Low Wage Gap)	0.029		0.019		0.030	
Baseline exit (REP+ × High Outside Wage)		0.031		0.031		0.029
Baseline exit (REP+ × Low Outside Wage)		0.026		0.015		0.026
Baseline exit (REP+ × High Wage)		0.041		0.040		0.047
Baseline exit (REP+ × Low Wage)		0.029		0.024		0.029

predicts responsiveness to the teacher wage increase. While the reform raised teacher wages uniformly across treated schools, teachers differed both in their baseline outside-option wage and in their baseline teacher wage. Splitting the wage gap into these two components therefore allows me to test whether individuals who had higher baseline outside option wages (high w_1) or those who had lower teacher wages (low w_0, i.e. received a higher percentage change in their wage) responded more strongly to the reform. This distinction is informative of the validity of the proposed microfoundation, which relies on the fact that a high outside-option wage w_1 reflects teachers who face greater potential returns to leaving and thus have stronger incentives to acquire information about outside options (higher attention θ, lower σ_e), leading to more wage-sensitive behavior.

The results presented in Table 6 are in line with this intuition: the triple-interaction term with the dummy of being in the top tercile of the 2013 outside wage distribution is statistically and economically significant for the average teacher (column 2) and the subsamples of STEM and non-STEM teachers (columns 4 and 6, respectively), indicating that treated teachers with a higher outside option wage experienced an even larger decrease in their exit probability. By contrast, the triple-interaction term with the dummy of being in the top tercile of the 2013 teacher wage distribution is positive but insignificant in both statistical and economic terms.[31]

Taken together, these suggestive evidence are in line with the framework's prediction that teachers facing higher-valued or better-known outside options (lower σ_e) react more sharply to changes in relative wages.

7. Conclusion

This paper provides the first causal evidence on the heterogeneous responsiveness of teacher exit decisions to wages by teacher productivity, and quantifies the implications of changes in the teacher wage gap for the aggregate quality of the teaching workforce implied by the exit channel. Exploiting a quasi-experimental wage increase introduced by the 2014 French REP+ reform, I find that teacher exit decisions are highly elastic to wages, and that high-productivity teachers—those with above-median value-added—are about three times as responsive as their lower-productivity counterparts.

Using a simple static model of labor supply with two productivity types, I show that as more productive teachers are more responsive to wage changes, a uniform increase in teacher pay raises the aggregate quality of the teaching force through its effect on exit decisions. Quantitatively, a uniform wage increase equivalent to the relative REP+ bonus yields modest but positive gains in average teacher value-added, while targeting the same amount toward high-productivity teachers produces larger improvements in aggregate quality. Conversely, I show that increases in outside-option wages—if not matched by commensurate increases in teacher pay—can substantially erode teacher quality. Taken together, these results highlight that teacher wage policies have important compositional effects that go beyond retention rates.

[31] This pattern is consistent with the complementary evidence from the triple-interaction regressions using other observables correlated with the wage gap, which show that older teachers (who have a larger wage gap on average)—respond more strongly to the reform (see Table A.18).

References

Allegretto, S. (2025, September). *The teacher pay penalty reached a record high in 2024: Three decades of leaving public school teachers behind.* https://www.epi.org/publication/the-teacher-pay-penalty-reached-a-record-high-in-2024-three-decades-of-leaving-public-school-teachers-behind/. (Economic Policy Institute)

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., & Wilson, D. (2009). Evaluating the impact of performance-related pay for teachers in england. *Labour Economics*, *16*, 251-261.

Bacolod, M. P. (2007). DO ALTERNATIVE OPPORTUNITIES MATTER? THE ROLE OF FEMALE LABOR MARKETS IN THE DECLINE OF TEACHER QUALITY. *The Review of Economics and Statistics*, *89*(4), 737–751.

Barrera-Osorio, F., & Raju, D. (2017). Teacher performance pay: Experimental evidence from pakistan. *Journal of Public Economics*, *148*, 75–91.

Bates, M. D., & Johnston, A. C. (2025, January). *Do pensions enhance teacher effort and selective retention?* (No. Working Paper 33986). Retrieved from https://www.nber.org/papers/w33986 doi: 10.3386/w33986

Bau, N., & Das, J. (2020). The impact of individual teachers on student achievement: Evidence from panel data. *Americal Economic Journal: Economic Policy*, *12*(1), 62–96.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, *70*(5, Part 2), 9–49.

Beffy, M., & Davezies, L. (2013). Has the "Ambition Success Networks" Educational Program Achieved its Ambition? *Annals of Economics and Statistics*, *111-112*, 271–293.

Benhenda, A., & Sims, S. (2022). The effect of financial incentives on the retention of shortage-subject teachers: evidence from England. *UCL Working paper No. 22-04*.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4, Part 1), 352–365.

Bhuller, M., Havnes, T., Leuven, E., & Mogstad, M. (2013). Broadband Internet: An Information Superhighway to Sex Crime? *The Review of Economic Studies*, *80*(4), 1237–1266.

Biasi, B. (2021). The labor market for teachers under different pay schemes. *American Economic Journal: Economic Policy*, *13*(2), 65–102.

Biasi, B., Fu, C., & Stromme, J. (2021). Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage. *NBER Working Papers 28530*.

Bietenbeck, J., Piopiunik, M., & Wiederhold, S. (2018). Africa's Skill Tragedy: Does Teachers' Lack of Knowledge Lead to Low Student Performance? *Journal of Human Resources*, *53*, 553–578.

Bobba, M., Ederer, T., Leon-Ciliotta, G., Neilson, C., , & Nieddu, M. G. (2021). Teacher Compensation and Structural Inequality: Evidence from Centralized Teacher School Choice in Perú. *NBER Working Paper No. 29068*.

Bold, T., Filmer, D. P., Martin, G., Molina, E., Christophe Rockmore, J. S., Brian W. Stacy, & Wane, W. (2016). What Do Teachers Know and Do? Does It Matter? Evidence from Primary Schools in Africa. *World Bank Working Paper 7956*.

Britton, J., & Propper, C. (2016). Teacher pay and school productivity: Exploiting wage regulation. *Journal of Public Economics*, *133*, 75–89.

Brown, C., & Andrabi, T. (2025). *Inducing positive sorting through performance pay: Experimental evidence from pakistani schools.* (University of California at Berkeley Working Paper)

Burgess, S. M., Greaves, E., & Murphy, R. J. (2022). Deregulating teacher labor markets. *Economics of Education Review*, *88*, 102253. doi: 10.1016/j.econedurev.2022.102253

Bénabou, R., Kramarz, F., & Prost, C. (2009). The French zones d'éducation prioritaire: Much ado about nothing? *Economics of Education Review*, *28*(3), 345–356.

Bó, E. D., Finan, F., & Rossi, M. A. (2013). Strengthening State Capabilities: the role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, *128*(3), 1169–1218.

Cabrera, J. M., & Webbink, D. (2020). Do higher salaries yield better teachers and better student outcomes? *Journal of Human Resources*, *55*(4), 1222–1257. doi: 10.3368/jhr.55.4.0717-8911R3

Chetty, R., Friedman, J., & Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, *104*(9), 2593–2632.

Chetty, R., Friedman, J., & Rockoff, J. E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, *104*(9), 2633–2679.

Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, *92*(5–6), 1352–1370.

Cour des comptes. (2018). L'ÉDUCATION PRIORITAIRE. *Rapport*.

Cowan, J., & Goldhaber, D. (2018). Do bonuses affect teacher staffing and student achievement in high poverty schools? evidence from an incentive for national board certified teachers in washington state. *Economics of Education Review*, *65*, 138–152. doi: 10.1016/j.econedurev.2018.06.010

Davezies, L., & Garrouste, M. (2020). More Harm than Good?: Sorting Effects in a Compensatory Education Program. *Journal of Human Resources*, *55*(1), 240–277.

Dee, T. S., & Keys, B. J. (2004, 6). Does merit pay reward good teachers? evidence from a randomized experiment. *Journal of Policy Analysis and Management*, *23*, 471-488.

Dee, T. S., & Wyckoff, J. (2015, 3). Incentives, selection, and teacher performance: Evidence from impact. *Journal of Policy Analysis and Management*, *34*, 267-297.

Deneault, C. (2023). Local Labor Markets and Selection into the Teaching Profession. *Working Paper*.

Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The Economic Consequences of Hospital Admissions. *American Economic Review*, *108*(2), 308–352.

Dolton, P., & Marcenaro-Gutierrez, O. (2011). If you pay peanuts do you get monkeys? a cross-country analysis of teacher pay and pupil performance. *Economic Policy*, *26*(65), 5–55. doi: 10.1111/j.1468-0327.2010.00257.x

Dolton, P., & van der Klaauw, W. (1995, March). Leaving teaching in the UK: A duration analysis. *The Economic Journal*, *105*(429), 431–444. doi: 10.2307/2235490

Duflo, E., Hanna, R., & Ryan, S. P. (2012, 6). Incentives work: Getting teachers to come to school. *American Economic Review*, *102*, 1241-78. doi: 10.1257/AER.102.4.1241

Falch, T. (2011). Teacher Mobility Responses to Wage Changes: Evidence from a Quasi-Natural Experiment. *American Economic Review: Papers Proceedings*, *101*(3), 460–465.

Falch, T. (2017). Wage and Recruitment: Evidence from External Wage Changes. *CESifo Working Paper, No. 4078*.

Falch, T., Johansen, K., & Strøm, B. (2009). Teacher shortages and the business cycle. *Labour Economics*, *16*(6), 648–658.

Feng, L., & Sass, T. R. (2018, 1). The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management*, *37*, 112-135.

Figlio, D. N. (1997). Teacher salaries and teacher quality. *Economics Letters*, *55*(2), 267–271. doi: 10.1016/S0165-1765(97)00088-X

Figlio, D. N. (2002). Can Public Schools Buy Better-Qualified Teachers? *ILR Review*, *55*(4), 686–699.

Figlio, D. N., & Kenny, L. W. (2007, 6). Individual teacher incentives and student performance. *Journal of Public Economics*, *91*, 901-914. doi: 10.1016/J.JPUBECO.2006.10.001

Fraenkel, R. C. (2022). Local labor markets and job match quality: Teachers. *Labour Economics*, *78*.

Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from new york city public schools. *Journal of Labor Economics*, *31*(2), 373–407.

Gibbons, R., & Murphy, K. J. (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy*, *100*(3), 468–505.

Glewwe, P., Ilias, N., & Kremer, M. (2010, 7). Teacher incentives. *American Economic Journal: Applied Economics*, *2*, 205-27. doi: 10.1257/APP.2.3.205

Godechot, O., Palladino, M. G., & Babet, D. (2023). In the Land of AKM: Explaining the Dynamics of Wage Inequality in France. *Insee Working Paper*.

Goodman, S. F., & Turner, L. J. (2013, 4). The design of teacher incentive pay and educational outcomes: Evidence from the new york city bonus program. *Journal of Labor Economics*, *31*, 409-420. doi: 10.1086/668676

Goodman-Bacon, A. (2018). Public Insurance and Mortality: Evidence from Medicaid Implementation. *Journal of Political Economy*, *126*, 216–262.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, *225*, 254–277.

Hendricks, M. D. (2014). Does it pay to pay teachers more? evidence from texas. *Journal of Public Economics*, *109*.

Hirsch, B., Jahn, E. J., Manning, A., & Oberfichtner, M. (2022, 11). The wage elasticity of recruitment. *IZA Discussion Paper No. 15675*. doi: 10.2139/SSRN.4268783

Hoxby, C., & Leigh, A. (2004). Pulled Away or Pushed out? Explaining the Decline of Teacher Aptitude in the United States. *The American Economic Review*, *94*(2), 236–240.

Jackson, C. K. (2014). Teacher Quality at the High School Level: The Importance of Accounting for Tracks. *Journal of Labor Economics*, *32*(4).

Johnston, A. C. (in press). Preferences, selection, and the structure of teacher pay. *American*

Economic Journal: Applied Economics. doi: 10.1257/APP.20210763

Johnston, A. C., Rockoff, J. E., & Harrington, J. R. (2025, April). *Pension reform and labor supply: Retention and productivity under a pension cut* (No. Working Paper 33673). doi: 10.3386/w33673

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615–631.

Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. *NBER Working Paper*(14607).

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy*, *87*(6), 1261–1284.

Leaver, C., Ozier, O., Serneels, P., & Zeitlin, A. (2021, July). Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from rwandan primary schools. *American Economic Review*, *111*(7), 2213–2246.

Manning, A. (2003). Monopsony in Motion: Imperfect Competition in Labor Markets. *Princeton University Press*.

Manning, A. (2011). Imperfect competition in the labor market. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4B, pp. 973–1041). Elsevier. doi: 10.1016/S0169-7218(11)02411-9

Matějka, F., & McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, *105*(1), 272–298. doi: 10.1257/aer.20130047

Morgan, A. J., Nguyen, M., Hanushek, E. A., Ost, B., & Rivkin, S. G. (2023, March). *Attracting and retaining highly effective educators in hard-to-staff schools* (No. Working Paper 31051). Retrieved from https://www.nber.org/papers/w31051 doi: 10.3386/w31051

Muralidharan, K., & Sundararaman, V. (2011, February). Teacher performance pay: Experimental evidence from india. *Journal of Political Economy*, *119*(1), 39–77.

Nagler, M., Piopiunik, M., & West, M. R. (2020). Weak Markets, Strong Teachers: Recession at Career Start and Teacher Effectiveness. *Journal of Labor Economics*, *38*(2).

Naidu, S., & Carr, M. (2022, 9). If you don't like your job, can you always quit? pervasive monopsony power and freedom in the labor market. *Journal of Law and Political Economy*, *3*. doi: 10.5070/LP63159035

OECD. (2023). *Education at a glance 2023: Oecd indicators*. Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/e13bef63-en doi: 10.1787/e13bef63-en

Pugatch, T., & Schroeder, E. (2014). Incentives for teacher relocation: Evidence from the gambian hardship allowance. *Economics of Education Review*, *41*, 120–136.

Pugatch, T., & Schroeder, E. (2018). Teacher pay and student performance: Evidence from the gambian hardship allowance. *Journal of Development Effectiveness*, *10*(2), 249–276.

Ree, J. J. D., Muralidharan, K., Pradhan, M. P., & Rogers, F. H. (2017). *Double for nothing? experimental evidence on an unconditional teacher salary increase in indonesia.*

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, *73*(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review Papers and Proceedings*, *94*(2), 247—252.

SGEN-CFDT. (2020). *Journées rep+ et pondération: une réponse possible pour travailler ensemble.* Retrieved from https://www.sgen-cfdt.fr/actu/journees-rep-et-ponderation-une-reponse-possible-pour-travailler-ensemble/ (Accessed October 2025)

Silhol, J., & Wilner, L. (2022, September). *Teachers' desired mobility to disadvantaged schools: Do financial incentives matter?* (Tech. Rep. No. 9906). CESifo Working Paper Series. Retrieved from https://ssrn.com/abstract=4212070 (CESifo Working Paper No. 9906)

Sims, C. A. (1998). Stickiness. *Carnegie-Rochester Conference Series on Public Policy*, *49*(1), 317–356.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.

SNEP-FSU. (2018). *Rep+ : faire vivre les acquis.* Retrieved from https://www.aix.snes.edu/rep-faire-vivre-les-acquis-que.html (Accessed October 2025)

Staiger, D. O., & Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives*, *24*(3), 97–118.

Tartova, D. (2023). Teacher Value-Added in the Absence of Annual Test Scores: Utilising Teacher Networks. *SSRN Working Paper*. Retrieved from https://ssrn.com/abstract=4395082

Tsao, C. (2025). It's not (just) about the money: Pay and the value of working conditions in teaching. *Working Paper*.

Willén, A. (2021). Decentralization of wage determination: Evidence from a national teacher reform. *Journal of Public Economics*, *198*, 104388. doi: 10.1016/j.jpubeco.2021.104388

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, *100*, 61–78.

Appendix

A. Tables

A.1. Descriptive tables

Table A.1: Descriptive statistics for the avg. teacher and student by school type in 2013-2014

Note: Means of each variable reported, with standard deviations in brackets. High qualification refers to "agrégation". The sample of teachers is based on tenure teachers below the age of 50, teaching in public middle schools in 2013-2014. Underlying data source: DEPP, Bases Relais, Scolarité, 2013-2014.

	Category of school			
	REP	REP+	Ordinary	Total
N	24,983 (20.30%)	6,197 (5.00%)	91,726 (74.60%)	122,906 (100.00%)
Female	0.63 (0.48)	0.60 (0.49)	0.67 (0.47)	0.65 (0.48)
Experience	9.23 (5.89)	8.41 (5.45)	11.38 (6.21)	10.80 (6.19)
Experience in school	3.66 (2.40)	3.48 (2.37)	3.88 (2.37)	3.81 (2.38)
Age	36.57 (6.36)	35.90 (6.11)	38.78 (6.14)	38.19 (6.27)
High qualification	0.05 (0.23)	0.05 (0.22)	0.06 (0.23)	0.06 (0.23)
Exit rate	0.020 (0.14)	0.028 (0.17)	0.017 (0.13)	0.018 (0.13)
Share very disadv. stud.	0.55 (0.12)	0.74 (0.10)	0.36 (0.13)	0.42 (0.17)
Share foreign stud.	0.07 (0.07)	0.13 (0.09)	0.02 (0.04)	0.04 (0.06)
Share scholarship stud.	0.25 (0.17)	0.39 (0.28)	0.14 (0.11)	0.18 (0.15)
Share very adv. stud.	0.10 (0.06)	0.03 (0.02)	0.22 (0.13)	0.18 (0.13)

Table A.2: Descriptive statistics for Math and French teachers by school type in 2013-2014

Note: Means of each variable reported, with standard deviations in brackets. High qualification refers to "agrégation". The sample of Math and French teachers is based on tenure teachers below the age of 50, teaching in public middle schools in 2013-2014. Underlying data source: DEPP, Bases Relais, Scolarité, Sysca, DNB, 2013-2014.

	Category of school			
	REP	REP+	Ordinary	Total
Math teachers				
N	3,687 (15.2%)	1,813 (7.5%)	18,729 (77.3%)	24,229 (100.0%)
Female	0.52 (0.50)	0.50 (0.50)	0.57 (0.50)	0.56 (0.50)
Experience	8.90 (5.80)	8.66 (5.51)	10.97 (6.14)	10.48 (6.11)
High qualification	0.04 (0.20)	0.05 (0.21)	0.05 (0.21)	0.05 (0.21)
Experience in school	3.57 (2.39)	3.49 (2.37)	3.74 (2.38)	3.69 (2.38)
Age	36.37 (6.50)	36.28 (6.21)	38.37 (6.06)	37.91 (6.20)
Exit rate	0.025 (0.16)	0.033 (0.18)	0.020 (0.14)	0.021 (0.15)
Teaches 9th grade Math	0.77 (0.42)	0.75 (0.43)	0.80 (0.40)	0.79 (0.41)
Has Math TVA	0.63 (0.48)	0.57 (0.50)	0.70 (0.46)	0.68 (0.47)
French teachers				
N	5,088 (16.2%)	2,435 (7.8%)	23,822 (76.0%)	31,345 (100.0%)
Female	0.75 (0.43)	0.74 (0.44)	0.82 (0.38)	0.80 (0.40)
Experience	9.27 (5.68)	8.69 (5.44)	11.30 (6.14)	10.77 (6.10)
High qualification	0.04 (0.21)	0.04 (0.19)	0.05 (0.22)	0.05 (0.22)
Experience in school	3.54 (2.38)	3.47 (2.36)	3.74 (2.39)	3.68 (2.39)
Age	36.79 (6.17)	36.33 (6.06)	38.58 (6.13)	38.12 (6.19)
Exit rate	0.027 (0.16)	0.038 (0.19)	0.021 (0.14)	0.023 (0.15)
Teaches 9th grade French	0.77 (0.42)	0.74 (0.44)	0.80 (0.40)	0.79 (0.41)
Has French TVA	0.63 (0.48)	0.58 (0.49)	0.71 (0.45)	0.69 (0.46)

A.2. Regression tables

Table A.3: Compliance rate between initial and current REP+ assignment

Note: The table reports the relationship between the initial treatment assignment, i.e. the ITT treatment variable (being in a REP+ school in 2013, the year before the bonus introduction), and the current treatment status (being in a REP+ school at time t versus being in any other school), in the post-reform period, for the subsamples of all teachers, high-TVA teachers, and low-TVA teachers. The dependent variable is the current school assignment, and the independent variable is the treatment variable based on 2013-2014 school assignment. The estimated coefficient represents the compliance rate—the proportion of teachers who remained in REP+ schools relative to their initial assignment. Columns (3), (6) and (9) report compliance by year for each of the three subsamples. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	All			High TVA			Low TVA		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × REP+	0.766***	0.775***		0.782***	0.796***		0.751***	0.769***	
	(0.00344)	(0.00334)		(0.00845)	(0.00817)		(0.00908)	(0.00869)	
REP+	-1.84e-13	-0.0516***	-0.0500***	8.50e-14	-0.0415***	-0.0401***	-4.54e-14	-0.0543***	-0.0527***
	(0.00312)	(0.00311)	(0.00307)	(0.00771)	(0.00766)	(0.00755)	(0.00825)	(0.00812)	(0.00801)
Year=2014 × REP+			0.913***			0.927***			0.901***
			(0.00432)			(0.0105)			(0.0112)
Year=2015 × REP+			0.840***			0.865***			0.834***
			(0.00441)			(0.0107)			(0.0115)
Year=2016 × REP+			0.781***			0.809***			0.768***
			(0.00449)			(0.0109)			(0.0117)
Year=2017 × REP+			0.727***			0.750***			0.720***
			(0.00457)			(0.0110)			(0.0119)
Year=2018 × REP+			0.681***			0.704***			0.679***
			(0.00464)			(0.0112)			(0.0120)
Year=2019 × REP+			0.652***			0.673***			0.652***
			(0.00471)			(0.0113)			(0.0122)
Constant	0.0185***	0.0323***	0.0320***	0.0184***	0.0272***	0.0270***	0.0161***	0.0288***	0.0286***
	(0.000734)	(0.000740)	(0.000729)	(0.00173)	(0.00174)	(0.00172)	(0.00194)	(0.00194)	(0.00191)
CZ FEs		✓		✓		✓		✓	
year	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148,861	148,860	148,860	24,420	24,415	24,415	22,053	22,048	22,048
R-squared	0.675	0.695	0.704	0.694	0.718	0.727	0.662	0.696	0.704

Table A.4: Turnover responsiveness to the wage increase for the average teacher

Note: The table presents the η coefficients from a regression of the type of 4 for the average teacher, whereby the dependent variable includes exits and transfers, for comparability to the literature. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual turnover rate for REP+ teachers, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4). The corresponding relative percentage decrease in turnover rates and elasticity with respect to the wage (computed using equation 5) are reported below. The bonus is expressed in thousands of euros. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	Outflow
Bonus	-0.00676
	(0.00422)
Teacher controls	✓
School FEs	✓
CZ $\times$ Year FEs	✓
Sample	[2013,2019]
Observations	148403
R-squared	0.006
First-stage F stat	5133
Counterfactual outflow	0.080
Counterfactual wage	26.888
Outflow elasticity to wage	-2.285

Table A.5: Responsiveness to the wage increase for the average teacher predicted by ATT

Note: The table presents the η coefficients from regressions 6 and 4, for the average teacher. Column (1) presents the preferred reduced-form specification, and column (2) shows the de-trended coefficients. Columns (3)-(5) present the corresponding 2SLS specification for the ATT. Column (6) presents the corresponding 2SLS specification for the ITT specification in the medium-run for comparison with the ATT. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate at REP+ schools, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4 and Section C.3). The corresponding relative percentage decrease in exit rates and elasticity with respect to the wage (computed using equation 5) are reported below. The bonus is expressed in thousands of euros. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.

	Reduced-form			2SLS			2SLS (ITT)
	Short-run	Medium-run	De-trended	Short-run	Medium-run	De-trended	Medium-run
Post × REP+	-0.00589***	-0.00746***	-0.0103***				
	(0.00202)	(0.00183)	(0.00149)				
Bonus				-0.0112**	-0.00950***	-0.0131***	-0.00793**
				(0.00467)	(0.00271)	(0.00211)	(0.00376)
Constant	0.0345***	0.0345***	0.000132				
	0.00131	(0.00109)	(0.000872)				
Teacher controls	✓	✓	✓	✓	✓	✓	✓
School FEs	✓	✓	✓	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓	✓	✓	✓
Sample	[2012,2015]	[2012,2017]	[2009,2017]	[2012,2015]	[2012,2017]	[2009,2017]	[2013,2017]
Observations	106,569	161,549	242,157	106,569	161,549	242,157	112,139
R-squared	0.019	0.017	0.013	0.003	0.003	0.000	0.001
First-stage F stat				2.83e+12	473,587	476,504	558,281
Counterfactual exit	0.031	0.031	0.035	0.032	0.031	0.035	0.023
As perc. of exit	-0.198	-0.241	-0.296				
Counterfactual wage				25.849	26.059	26.059	27.021
Elasticity to wage				-8.945	-7.995	-9.798	-9.458

Table A.6: Differential number of hours spent in teaching for REP+ teachers compared to REP

Note: The figures present the coefficients of an event study difference-in-difference regression where the effective number of hours taught by teacher j in year t at school s are regressed on a $Post_t$ dummy multiplied by the treatment variable $REP{+}_j$ (as defined under ITT). Teacher-by-year level controls include tercile of teaching experience and tercile of age. Fixed effects include teacher, school, and commuting-zone-by-year fixed effects. Column (1) focuses on all teachers, column (2) on Capes-certified teachers, column (3) on Agrégation-certified teachers, and columns (4) and (5) on the subsamples of high- and low-TVA teachers, respectively. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	All			High TVA	Low TVA
	(1) All	(2) Capes	(3) Agregation	(4)	(5)
Post × REP+	-0.773***	-0.741***	-0.517	-0.874***	-0.711***
	(0.135)	(0.135)	(0.394)	(0.215)	(0.228)
Teacher controls	✓	✓	✓	✓	✓
CZ FEs	✓	✓	✓	✓	✓
School FEs	✓	✓	✓	✓	✓
Teacher FEs	✓	✓	✓	✓	✓
Sample					
Observations	113291	106558	6132	18635	16380
R-squared	0.613	0.609	0.678	0.616	0.623
Mean hours for REP+ in 2013	18.579	18.698	16.126	18.562	18.711

Table A.7: Robustness checks for the sample of all teachers: adding school controls

Note: The table presents the η coefficients from regression 2, for the average teacher. Column (1) presents the preferred reduced-form specification. Columns (2)-(4) add different school controls—contemporaneous (2), as 1-year lags (3), and as 2-year lags (4). Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate for REP+ teachers as well as the corresponding relative percentage decrease in exit rates is reported at the bottom (for details, see Section 4). The set of school-level controls include the average share of disadvantaged students, need-based scholarship holders, female students, and primary-school repeaters. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	(1) Exit	(2) Exit	(3) Exit	(4) Exit
Post × REP+	-0.00544**	-0.00551**	-0.00553**	-0.00551**
	(0.00230)	(0.00231)	(0.00231)	(0.00232)
Constant	0.0196***	0.0157	0.0159	0.0353**
	(0.00137)	(0.0169)	(0.0171)	(0.0165)
Teacher controls	✓	✓	✓	✓
School controls	–	✓	–	–
School controls lag1	–	–	✓	–
School controls lag2	–	–	–	✓
School FEs	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148,403	148,372	148,262	148,081
R-squared	0.050	0.050	0.050	0.050
Counterfactual exit	0.022	0.022	0.022	0.022
As perc. of exit	-0.247	-0.250	-0.251	-0.250

Table A.8: Robustness checks for the sample of all teachers: different specification choices

Note: The table presents the η coefficients from regression 2, for the average teacher. Column (1) presents the preferred reduced-form specification. Column (2) uses a definition of exit which restricts to no re-entry for at least three, rather than two, years. Column (3) excludes schools which opened after 2013-2014 or shut down prior to 2019-2020. Column (4) excludes teachers who are novice (with less than 2 years of experience) in 2013-2014. Column (5) extends the sample of teachers to a less conservative retirement age cut—58 years old, instead of 50 years old. The mean counterfactual exit rate for REP+ teachers as well as the corresponding relative percentage decrease in exit rates is reported at the bottom (for details, see Section 4). The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	(1) Baseline	(2) Exit for 3y min.	(3) Balanced sch. panel	(4) No novice	(5) Exit below 58 yo
Post × REP+	-0.00545**	-0.00451**	-0.00549*	-0.00519**	-0.00582***
	(0.00230)	(0.00211)	(0.00230)	(0.00215)	(0.00219)
Constant	0.0191***	0.0160***	0.0192***	0.0160***	0.0186***
	(0.00125)	(0.00114)	(0.00126)	(0.00121)	(0.00123)
Teacher controls	✓	✓	✓	✓	✓
School FEs	✓	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148,403	148,403	147,728	148,050	172,327
R-squared	0.050	0.046	0.050	0.051	0.046
Counterfactual exit	0.022	0.020	0.022	0.022	0.023
As perc. of exit	-0.247	-0.230	-0.219	-0.235	-0.256

Table A.9: Reduced-form estimates for the average teacher given different definitions of treatment

Note: The table presents the η coefficients from regression 2, for the average teacher. Column (1) presents the preferred reduced-form specification. Columns (2) defines treatment as 1 if a teacher is assigned in 2013 to a school that becomes REP+ in or after 2015, and 0 if they are assigned in 2013 to a school that becomes REP in or after 2015. Columns (3) defines treatment as 1 if a teacher is assigned in 2013 to a school that becomes REP+ in 2014, and 0 if they are assigned in 2013 to a school that becomes REP in or after 2015. Columns (4) defines treatment as 1 if a teacher is assigned in 2013 to a school that becomes REP+ in 2014 or 2015, and 0 if they are assigned in 2013 to a school that becomes REP in 2015. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4). The corresponding relative percentage decrease in exit rates is reported below. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	(1) Baseline	(2) Excluding treated in 2014	(3) Treated in 2014	(4) Treated in 2014-2015
Post × REP+	-0.00544**	-0.00535**	-0.00628	-0.00502**
	(0.00230)	(0.00252)	(0.00384)	(0.00234)
Constant	0.0196***	0.0206***	0.0191***	0.0196***
	(0.00137)	(0.00142)	(0.00144)	(0.00142)
Teacher controls	✓	✓	✓	✓
School FEs	✓	✓	✓	✓
CZ x Year FEs	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148403	135528	114565	143154
R-squared	0.050	0.053	0.056	0.051
Counterfactual exit	0.022	0.021	0.025	0.022
As perc. of exit	-0.247	-0.255	-0.254	-0.227

Table A.10: Reduced form estimates given different subsamples of teachers

Note: The table presents the η coefficients from regression 2, for the average teacher. Column (1) presents the preferred reduced-form specification. Columns (2)-(4) add different school controls—contemporaneous (2), as 1-year lags (3), and as 2-year lags (4). Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate for REP+ teachers as well as the corresponding relative percentage decrease in exit rates is reported at the bottom (for details, see Section 4). The set of school-level controls include the average share of disadvantaged students, need-based scholarship holders, female students, and primary-school repeaters. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	Baseline (1)	Capes (2)	Agregation (3)
Post × REP+	-0.00544**	-0.00541**	-0.0125
	(0.00230)	(0.00230)	(0.0145)
Constant	0.0196***	0.0185***	0.0524***
	(0.00137)	(0.00139)	(0.0120)
Teacher controls	✓	✓	✓
School FEs	✓	✓	✓
CZ × Year FEs	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]
Observations	148403	139916	7981
R-squared	0.050	0.053	0.225
Counterfactual exit	0.022	0.022	0.021
As perc. of exit	-0.247	-0.245	-0.583

Table A.11: Responsiveness to the wage increase by TVA group predicted by ATT

Note: The table presents the η coefficients from regressions 6 and 4 by TVA group, with columns (1)-(6) depicting the estimates for the subsample of high-TVA teachers and columns (7)-(12) depicting the estimates for the subsample of low-TVA teachers. Columns (1) and (7) presents the preferred reduced-form specification for each group, and columns (2) and (8) shows the de-trended coefficients. Columns (3)-(5) and (9)-(11) present the corresponding 2SLS specification for the ATT. Columns (6) and (12) present the corresponding 2SLS specification for the ITT specification in the medium-run for comparison with the ATT. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate at REP+ schools, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4 and Section C.3). The corresponding relative percentage decrease in exit rates and elasticity with respect to the wage (computed using equation 5) are reported below. The bonus is expressed in thousands of euros. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.

	High TVA						**Low TVA**					
	Reduced-form		2SLS			2SLS (ITT)	Reduced-form		2SLS			2SLS (ITT)
	Medium-run	De-trended	Short-run	Medium-run	De-trended	Medium-run	Medium-run	De-trended	Short-run	Medium-run	De-trended	Medium-run
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post × REP+	-0.0152***	-0.0201***					-0.00330	-0.00200				
	(0.00387)	(0.00306)					(0.00389)	(0.00322)				
Bonus			-0.0223*	-0.0194***	-0.0257***	-0.0160*			-0.00524	-0.00425	-0.00257	-0.00536
			(0.0119)	(0.00556)	(0.00418)	(0.00817)			(0.0126)	(0.00590)	(0.00462)	(0.00878)
Constant	0.0288***	0.0118***					0.0299***	0.00406				
	(0.00539)	(0.00447)					(0.00553)	(0.00437)				
Teacher controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
School FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sample	[2012,2017]	[2009,2017]	[2012,2015]	[2012,2017]	[2009,2017]	[2013,2017]	[2012,2017]	[2009,2017]	[2012,2015]	[2012,2017]	[2009,2017]	[2013,2017]
Observations	24,957	37,415	12,446	24,957	37,415	17,958	23,328	35,161	11,663	23,328	35,161	16,292
R-squared	0.077	0.067	0.003	0.002	0.001	0.002	0.088	0.073	0.004	0.003	0.000	0.004
First-stage F stat			4.03e+11	63,419	68,333	68,412			3.56e+11	56,503	59,142	56,934
Counterfactual exit	0.024	0.029	0.024	0.024	0.029	0.022	0.019	0.020	0.020	0.019	0.020	0.017
As perc. of exit	-0.636	-0.686					-0.171	-0.099				
Counterfactual wage			25.934	26.169	26.169	26.897			25.808	26.121	26.121	26.946
Elasticity to wage			-23.751	-21.288	-22.913	-19.938			-6.728	-5.737	-3.312	-8.387

Table A.12: Responsiveness to the wage increase and implied elasticities to the teacher wage for the average Math and French teacher (ATT vs ITT)

Note: The table presents the η coefficients from regression 4, for the sample of Math and French teachers. Column (1) presents the results for the ITT specification. Columns (2)-(4) present the results for the ATT, using as dependent variable either $exit_{jst}$("Baseline") or $\widehat{exit}_{jst}^{DT}$ ("De-trended"). The corresponding sample period is reported below each regression column. The "De-trend" regressions use as pre-period all school yeas between 2009-2010 and 2013-2014, and perform a de-trending of the dependent variable prior to the 2SLS, as shown in regression C.1. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The mean counterfactual exit rate at REP+ schools, as well as the mean counterfactual wage for REP+ teachers are reported at the bottom for each subsample (for details, see Section 4). The bonus is expressed in thousands of euros. The implied elasticity to the wage computed at the bottom of the table uses equation 5. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2009-2010 to 2019-2020.

	ITT		ATT	
	(1)	(2)	(3)	(4)
	Baseline	ATT medium-run	ATT short-run	ATT De-trended
Bonus	-0.00899**	-0.0201***	-0.0255***	-0.0276***
	(0.00378)	(0.00443)	(0.00944)	(0.00339)
Teacher controls	✓	✓	✓	✓
School FEs	✓	✓	✓	✓
CZ x Year FEs	✓	✓	✓	✓
Sample	[2013,2019]	[2012,2017]	[2012,2015]	[2009,2017]
Observations	63049	68972	34303	103196
R-squared	0.003	0.009	0.009	0.001
First-stage F stat	65591	199169	1.25e+12	202963
Counterfactual exit	0.028	0.038	0.040	0.046
Counterfactual wage	26.849	26.125	25.884	26.125
Elasticity to wage	-8.483	-13.670	-16.616	-15.605

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: Reduced-form triple difference-in-differences comparing high- versus low-TVA (ATT vs ITT)

Note: The table presents the β coefficients from the regressions of the type 7, sequentially adding fixed effects and controls for the ITT specification (columns 1-5) and the ATT specification (columns 6-10), where in the latter case the treatment variable is $\mathbb{1}(type_s = \text{REP+})$. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. All teacher controls are interacted with the High TVA dummy. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2012-2013 to 2019-2020.

	ITT (1) Exit	ITT (2) Exit	ITT (3) Exit	ITT (4) Exit	ITT (5) Exit	ATT (6) Exit	ATT (7) Exit	ATT (8) Exit	ATT (9) Exit	ATT (10) Exit
Post × REP+ × High TVA	-0.0101 (0.00723)	-0.0101 (0.00722)	-0.0101 (0.00723)	-0.0101 (0.00723)	-0.0119 (0.00883)	-0.0122** (0.00545)	-0.0121** (0.00544)	-0.0122** (0.00544)	-0.0121** (0.00544)	-0.0129* (0.00663)
Post × REP+	-0.00268 (0.00484)	-0.00284 (0.00483)	-0.00284 (0.00483)		-0.00445 (0.00831)	-0.00332 (0.00389)	-0.00330 (0.00388)	-0.00330 (0.00388)		
Teacher controls	–	✓	✓	✓	✓	–	✓	✓	✓	✓
School x High TVA FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CZ x Year x High TVA FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Subject x High TVA FEs	–	–	✓	✓	✓	–	–	✓	✓	✓
REP+ x Year Fes	–	–	–	✓	–	–	–	–	✓	–
School x Year FEs	–	–	–	–	✓	–	–	–	–	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]
Observations	45613	45613	45613	45613	41168	48220	48220	48220	48220	48215
R-squared	0.131	0.132	0.132	0.132	0.223	0.080	0.083	0.083	0.083	0.161

Table A.14: Reduced-form triple difference-in-differences comparing high- versus low-TVA controlling for other heterogeneous effects (ITT)

Note: The table presents the β coefficients from regressions 7 (column 1) and 8 (columns 2-7). Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. All teacher controls are interacted with the High TVA dummy. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	(1) Exit	(2) Exit	(3) Exit	(4) Exit	(5) Exit	(6) Exit	(7) Exit
Post × REP+ × High TVA	-0.0119 (0.00884)	-0.0119 (0.00883)	-0.0121 (0.00880)	-0.0120 (0.00888)	-0.0114 (0.00891)	-0.0121 (0.00886)	-0.0120 (0.00895)
Post × REP+ × Experience terc.=2			0.0105 (0.0108)				0.00842 (0.0137)
Post × REP+ × Experience terc.=3			-0.00199 (0.0122)				-0.00723 (0.0173)
Post × REP+ × Age terc.=2				0.00382 (0.0106)			0.00220 (0.0134)
Post × REP+ × Age terc.=3				0.00111 (0.0111)			0.00721 (0.0162)
Post × REP+ × Female					-0.00324 (0.00896)		-0.00333 (0.00903)
Post × REP+ × High qualif.						-0.00459 (0.0264)	-0.00581 (0.0262)
Teacher controls	–	✓	✓	✓	✓	✓	✓
School x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
CZ x Year x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
Subject x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
School x Year FEs	✓	✓	✓	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	41168	41168	41168	41168	41168	41168	41168
R-squared	0.222	0.223	0.224	0.223	0.223	0.224	0.224

Table A.15: Reduced-form triple difference-in-differences comparing high- versus low-TVA controlling for other heterogeneous effects (ATT)

Note: The table presents the β coefficients from the regressions 7 (column 1) and 8 (columns 2-7), where the treatment variable is $\mathbb{1}(type_s = \text{REP+})$. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. All teacher controls are interacted with the High TVA dummy. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2012-2013 to 2017-2018.

	(1) Exit	(2) Exit	(3) Exit	(4) Exit	(5) Exit	(6) Exit	(7) Exit
Post × REP+ × High TVA	-0.0134**	-0.0129*	-0.0132**	-0.0131**	-0.0130*	-0.0129*	-0.0132**
	(0.00664)	(0.00663)	(0.00661)	(0.00664)	(0.00669)	(0.00664)	(0.00670)
Post × REP+ × Experience terc.=2			0.00386				0.00328
			(0.00762)				(0.00924)
Post × REP+ × Experience terc.=3			-0.000792				-0.00291
			(0.00827)				(0.0118)
Post × REP+ × Age terc.=2				0.00305			0.00216
				(0.00742)			(0.00909)
Post × REP+ × Age terc.=3				0.00147			0.00329
				(0.00794)			(0.0114)
Post × REP+ × Female					0.000624		-0.000300
					(0.00659)		(0.00668)
Post × REP+ × High qualif.						-0.0274	-0.0282
						(0.0202)	(0.0203)
Teacher controls	–	✓	✓	✓	✓	✓	✓
School x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
CZ x Year x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
Subject x High TVA FEs	✓	✓	✓	✓	✓	✓	✓
School x Year FEs	✓	✓	✓	✓	✓	✓	✓
Sample	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]	[2012,2017]
Observations	48215	48215	48215	48215	48215	48215	48215
R-squared	0.159	0.161	0.161	0.161	0.161	0.161	0.162

Table A.16: Reduced-form triple difference-in-differences comparing high- versus low-TVA for different type of specifications

Note: The table presents the η coefficients from regressions 2 by TVA group for different specifications of the TVA groups. Columns (1) and (4) present the preferred reduced-form specification for the high- and low-TVA subsamples, respectively, for comparability. Columns (2) and (5) present a reduced-form specification where the High- and Low-TVA groups are defined within commuting zone in the pre-2013 period. Columns (3) and (6) present a reduced-form specification where the High- and Low-TVA groups are defined within school in the pre-2013 period. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	High TVA			Low TVA		
	Baseline	TVA w/ CZ	TVA w/ school	Baseline	TVA w/ CZ	TVA w/ school
	(1)	(2)	(3)	(4)	(5)	(6)
Post × REP+	-0.0127**	-0.0128**	-0.0124***	-0.00284	-0.00340	-0.00304
	(0.00537)	(0.00518)	(0.00467)	(0.00485)	(0.00493)	(0.00563)
Teacher controls	✓	✓	✓	✓	✓	✓
CZ × Year FEs	✓	✓	✓	✓	✓	✓
School FEs	✓	✓	✓	✓	✓	✓
Sample	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]	[2013,2019]
Observations	24,059	23,710	25,896	21,619	21,874	19,748
R-squared	0.121	0.141	0.133	0.143	0.142	0.153

Table A.17: Wage gap on observable characteristics

Note: The table presents the coefficients from regressions which regress the wage gap on observable teacher characteristics associated with the wage gap in the subsample of teachers in either the treatment or control group, estimated over the pre-reform period. Teacher-by-year level controls are excluded from the regression to show unconditional correlations. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2009-2010 to 2013-2014.

	(1) Wage gap	(2) Wage gap	(3) Wage gap	(4) Wage gap
Age terc.=3	2.129***			
	(0.0341)			
Female		-5.104***		
		(0.0183)		
Echelon terc.=3			-0.495***	
			(0.0316)	
Experience terc.=3				-0.484***
				(0.0364)
Constant	2.434***	6.596***	3.479***	3.344***
	(0.0167)	(0.0142)	(0.0158)	(0.0166)
School × Year FEs	✓	✓	✓	✓
Observations	61,033	91,167	78,415	65,492
R-squared	0.553	0.745	0.514	0.528

Table A.18: Triple difference-in-differences of exit on observable teacher characteristics

Note: The table presents the coefficients from a triple-interaction regression identifying the effect of the reform for subgroups of teachers based on observable teacher characteristics. Teacher-by-year level controls include gender, educational qualifications, teaching experience tercile, age tercile, and subject specialization. All teacher controls are interacted with the relevant observable characteristic(s) in the regression. The confidence intervals correspond to 95% confidence levels. Standard errors, clustered by teacher and school-by-year, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

	(1) Exit	(2) Exit	(3) Exit	(4) Exit	(5) Exit
Post × REP+ × Age terc.=2					-0.000299
					(0.00698)
Post × REP+ × Age terc.=3	-0.0107*				-0.0141*
	(0.00556)				(0.00811)
Post × REP+ × Female		0.00269			0.00500
		(0.00409)			(0.00388)
Post × REP+ × Echelon terc.=2					-0.00849
					(0.00780)
Post × REP+ × Echelon terc.=3			-0.00829*		0.00889
			(0.00491)		(0.0113)
Post × REP+ × Experience terc.=2					0.00215
					(0.00773)
Post × REP+ × Experience terc.=3				-0.00150	-0.00976
				(0.00525)	(0.0128)
Teacher controls	✓	✓	✓	✓	✓
School × Year FEs	✓	✓	✓	✓	✓
School × Age terc. FEs	✓	–	–	–	✓
School × Female FEs	–	✓	–	–	✓
School × Echelon terc. FEs	–	–	✓	–	✓
School × Experience terc. FEs	–	–	–	✓	✓
Observations	85,426	141,097	87,709	93,952	108,855
R-squared	0.129	0.104	0.135	0.120	0.188

B. Figures

B.1. Descriptive figures

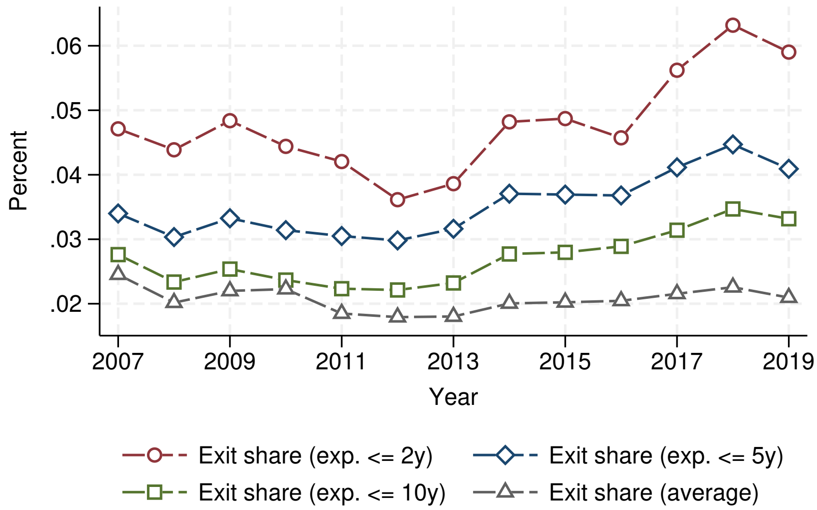


Figure B.1: Exit rates of teachers from the profession by experience groups and subject

Note: The figure plots the share of teachers in middle school exiting the profession every school year, restricted to teachers below the age of 50 in order to reduce noise from retirement decisions, for the sample of all tenured teachers. School year t—$t+1$ is denoted as t. Exit is defined as 1 if a teacher in school year t is not observed in secondary school teaching during at least school year $t+1$ and school year $t+2$, such that exit shares at t are defined as the teachers last observed in school year t, as a share of all teachers at t. The last school year in graph is 2019-2020, as it is the last year for which this horizon of two years is possible. Exit shares are split in different experience groups, indicated in the labels of each line. Underlying data source: DEPP, Bases Relais, 2007-2008/2019-2020.

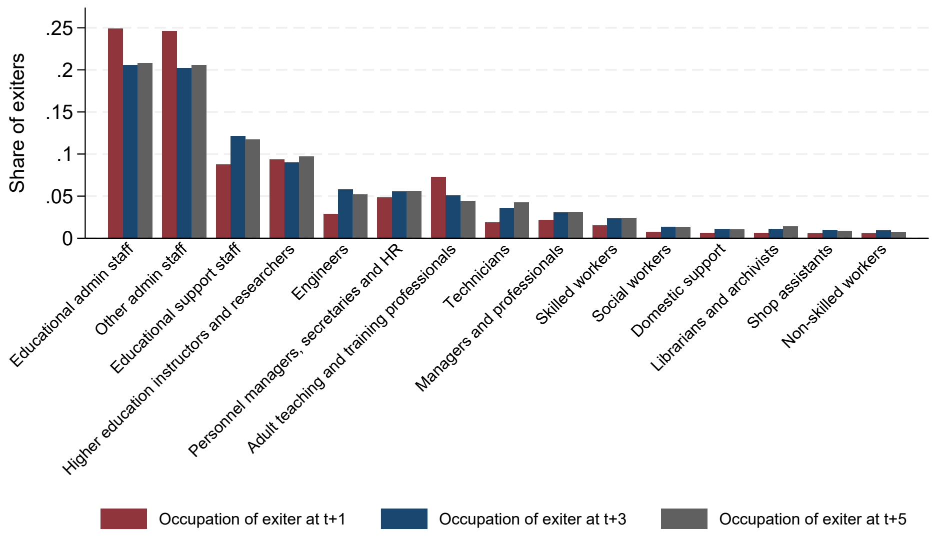


Figure B.2: Top occupations for teachers leaving the teachers profession at $t+1$, $t+3$, $t+5$

Note: The figure is constructed based on the 15 most popular occupations towards which teachers transition between 2009 and 2021. The shares of exiters represent national shares, based on the number of exiting teachers being found at a given occupation at, respectively, year $t+1$, $t+3$ or $t+5$ after an exit from teaching at year t. This aggregation is based on either occupation level 2 or level 3 and industry level 2 only to distinguish between the education industry and the rest. The aggregation is not used for creating the outside option salaries panel, and is only useful for providing an aggregate idea of outside options. Underlying data source: Insee, BTS Postes, 2009-2021.

B.2. Analysis

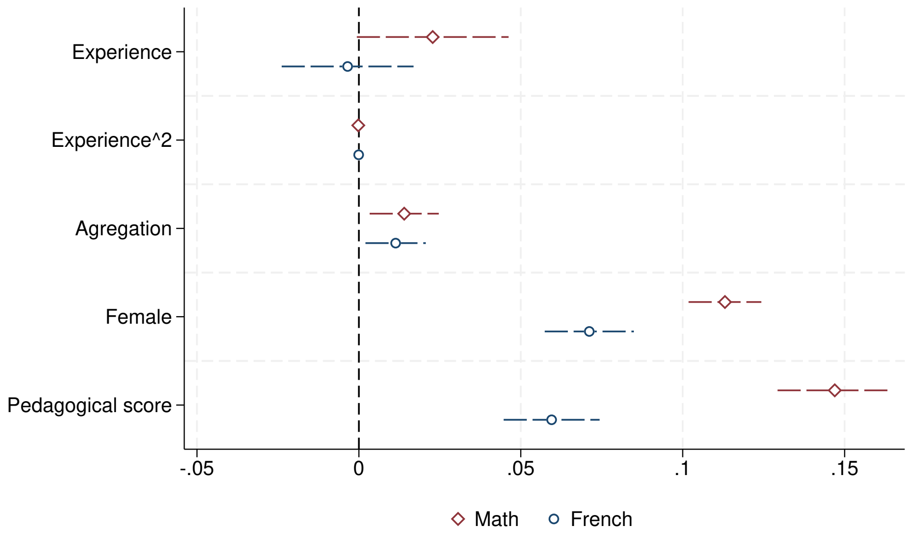


Figure B.3: Correlation of TVA with observable teacher characteristics

Note: The graph shows the estimates of individual regressions of the standardized TVA estimates in either Math or French on standardied observable teacher characteristics. In cases of non-dummy variables (such as experience), I average the variables at the teacher level, such that the dataset used has one observation per teacher. This choice reflects the fact that TVA is not measured within year. The dummy Agregation represents having a high qualification, while the dummy Capes represents having an ordinary qualification. The dummy contractual represents non-tenured non-qualified teachers. I then use this teacher-level dataset to run regressions with have on the left hand-side the TVA of the teacher, and on the right hand-side one of the observable characteristics. I plot the coefficients of each of these individual regressions. Underlying data source: Bases Relais, Scolarité, Sysca, 2007-2008 to 2021-2022.

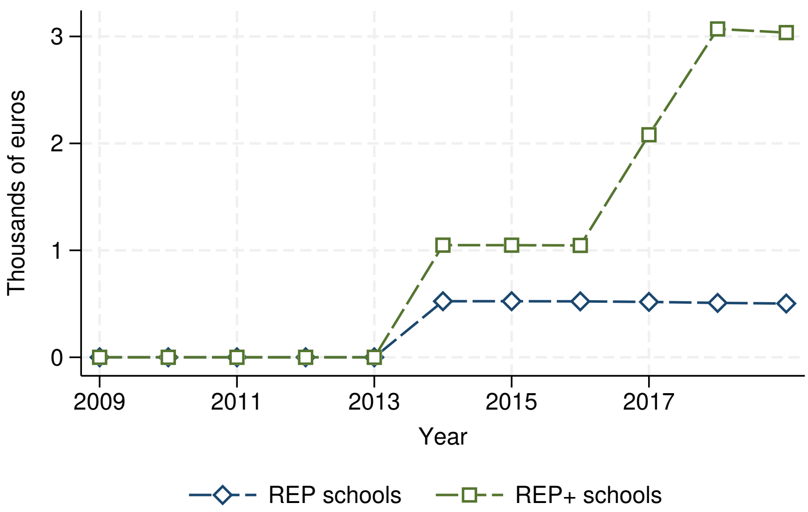
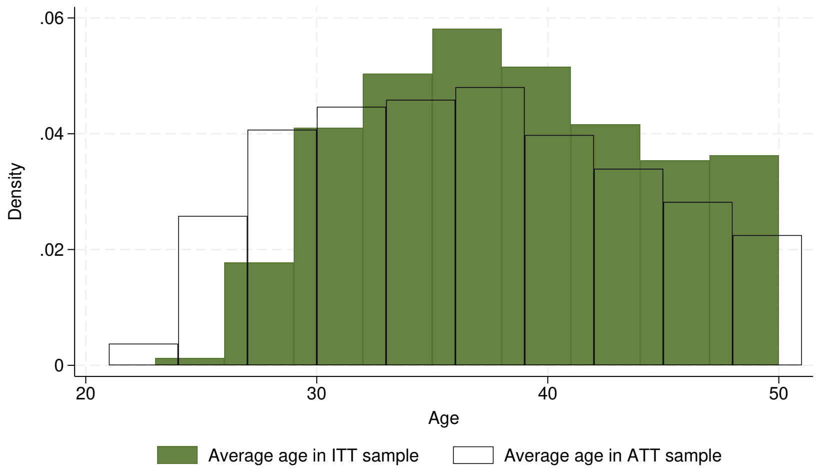
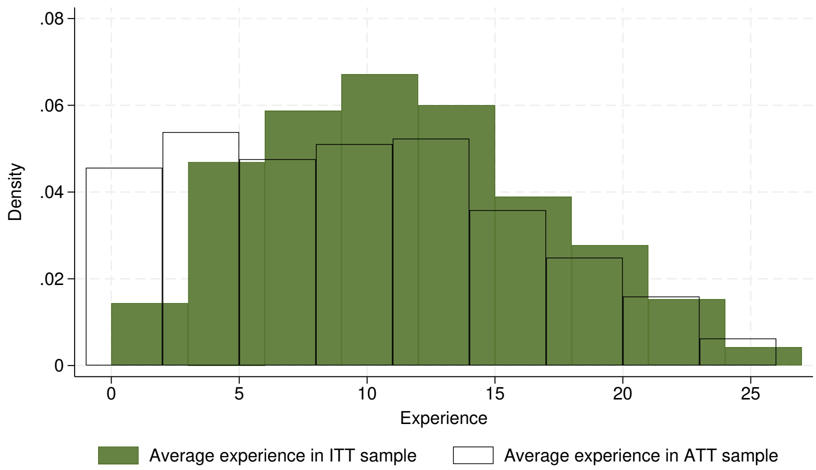


Figure B.4: Wage bonus as an instrument for the change in the wage gap

Note: The figures show the yearly bonus treatment variable for teachers at REP and REP+ schools implied by the reform. The bonus treatment variable shifts the bonus by a year in order to align the timing of the bonus with the definition of the exit variable. Underlying data source: Bases Relais, Scolarité, Sysca, DNB, 2009-2010 to 2019-2020.



(a) Age



(b) Experience

Figure B.5: Age and experience distributions under the ITT and ATT samples in the post-reform period

Note: The figures present histograms of the (a) age and (b) experience distributions in the post-reform period for the teachers identified as treated or control groups under the ITT approach, versus those identified as treated or control groups under the ATT approach. Specifically, the former are the groups of teachers assigned to REP+ or REP schools in 2013. The latter are the groups of teachers effectively at REP+ or REP schools. Underlying data source: Bases Relais, 2014-2015 to 2019-2020.

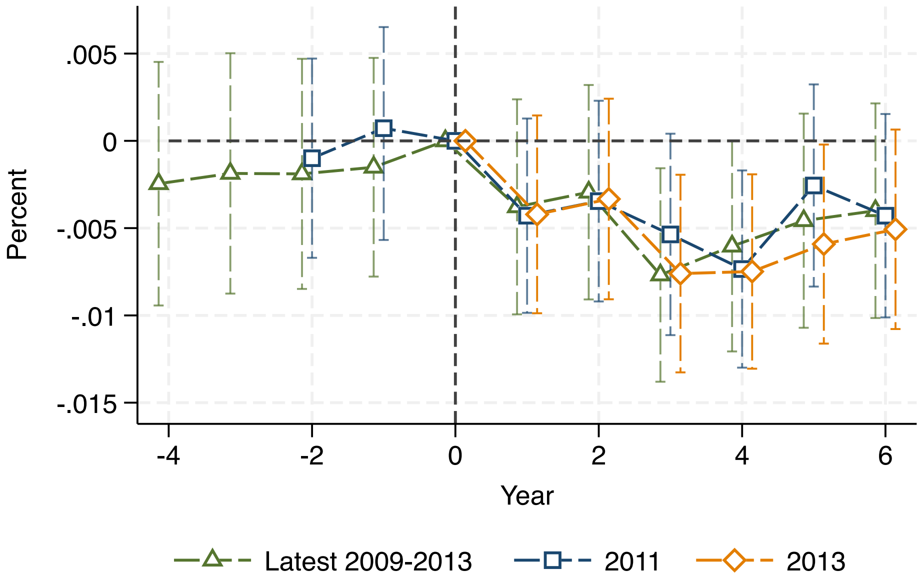
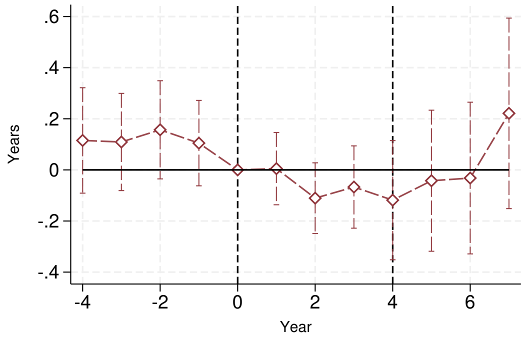
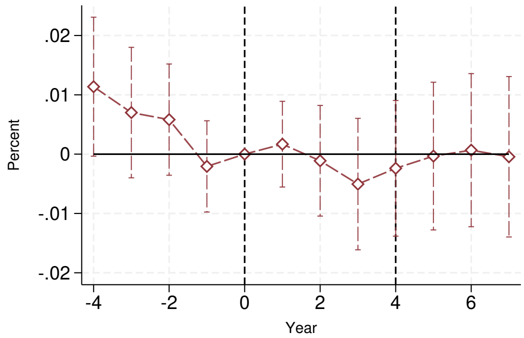


Figure B.6: Differential exit probability for the average REP+ teacher compared to REP under alternative definitions of treatment based on year of assignment

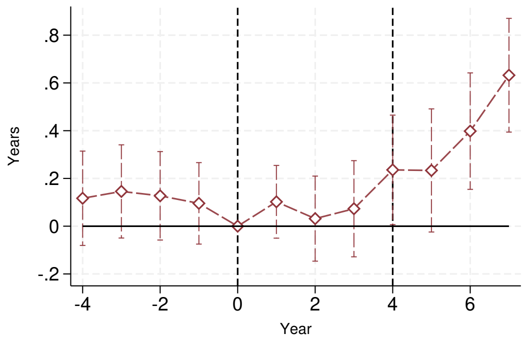
Note: The figures present the η_k coefficients of the event study difference-in-difference regressions 2, under different definitions of the treatment variable. Specifically, the "2013" specification presents the baseline ITT estimates when treatment is defined based on the 2013-2014 assignment of teachers. Similarly, the "2011" specification presents the ITT estimates when treatment is defined based on the 2011-2012 assignment of teachers. Finally, the "Latest 2009-2013" specification presents the ITT estimates when treatment is defined based on the assignment of a teacher in the last year in which they are observed in the pre-reform period. Year 0 represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, tercile of teaching experience, tercile of age, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2009-2010 to 2019-2020.



(a) Average pedagogical score



(b) Share female



(c) Average experience

Figure B.7: Compositional changes at REP+ schools relative to REP schools

Note: The figure presents the η_k coefficients from a regression $\overline{Y}_{st} = \sum_{k=2007}^{2021} \eta_k \mathbb{1}(t = k) \times \mathbb{1}(type_s = REP+) + \theta_s + \lambda_{zt} + \varepsilon_{st}$, weighted by the number of teachers in school s. The $\overline{Y}_{st}$ variable of each regression is specified below each graph. Year 0, which coincides with the first vertical dashed line, represents the year 2013, when the reform was first announced before it was gradually implemented from 2014 onwards. Year 4, which coincides with the second vertical dashed line, represents the year 2017—the year before the additional gradual increase in teacher salaries. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais.

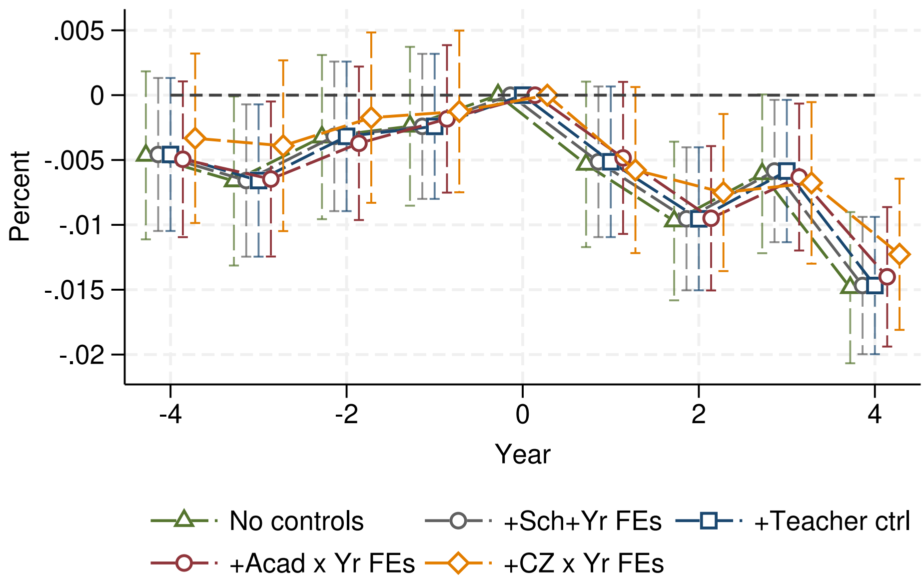
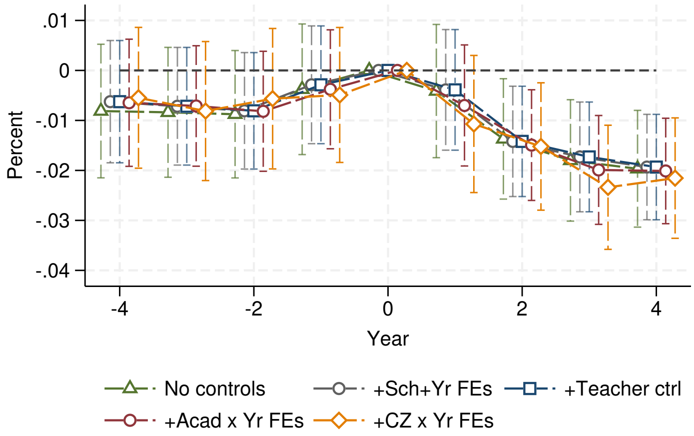
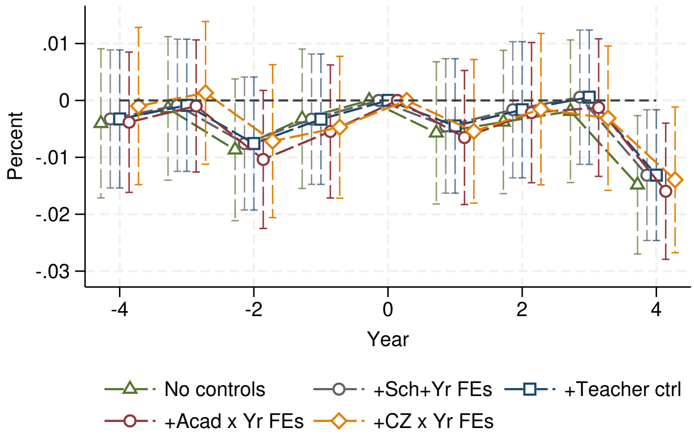


Figure B.8: Differential exit probability for the average REP+ teacher compared to REP

Note: The figures present the η_k coefficients of the event study difference-in-difference regressions 6, respectively, sequentially adding control variables and fixed effects, for the sample of all REP+ and REP teachers. Year 0 represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, tercile of teaching experience, tercile of age, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2009-2010 to 2019-2020.



(c) High TVA



(d) Low TVA

Figure B.9: Differential exit probability for REP+ teachers compared to REP by TVA

Note: The figures present the η_k coefficients of the event study difference-in-difference regression 6, sequentially adding control variables and fixed effects, for the sample of (a) high TVA and (b) low TVA REP+ and REP teachers. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience, and subject specialization. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.

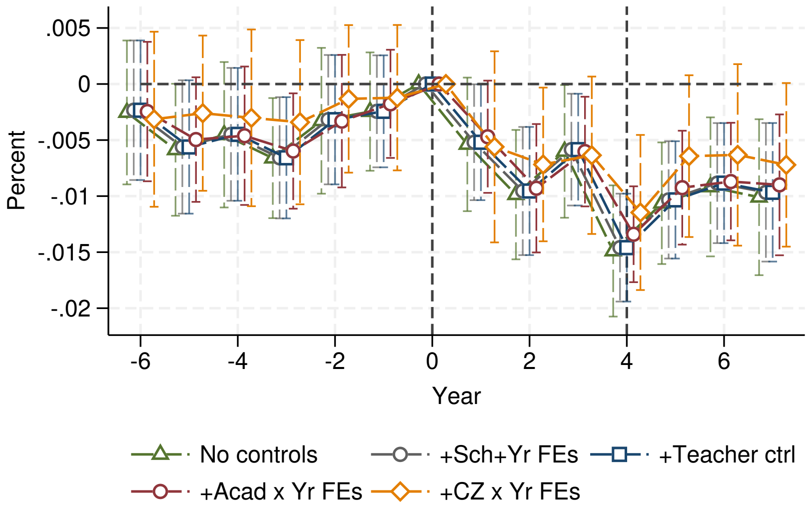


Figure B.10: Differential exit probability for REP+ compared to REP teachers, long horizon

Note: The figures present the η_k coefficients of the event study difference-in-difference regression 6, extended for $k \in \{2007, 2020\}$, sequentially adding control variables and fixed effects, for the sample of all REP+ and REP teachers. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience, and subject specialization. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais, 2007-2008 to 2020-2021.

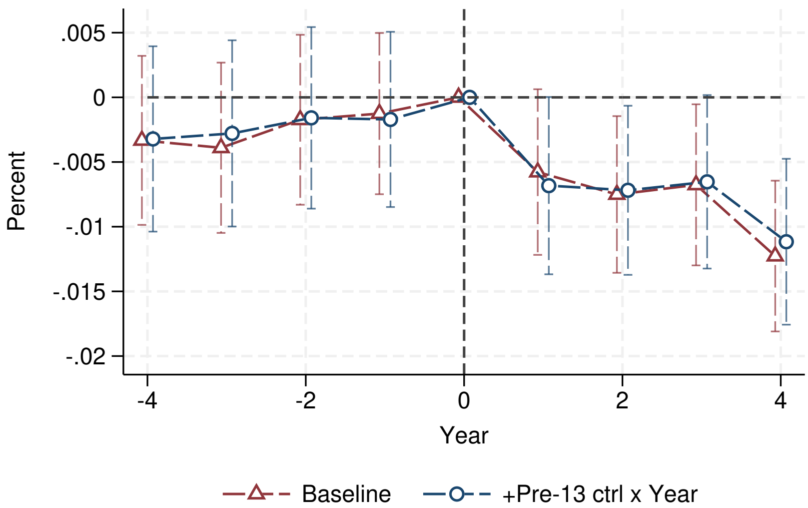
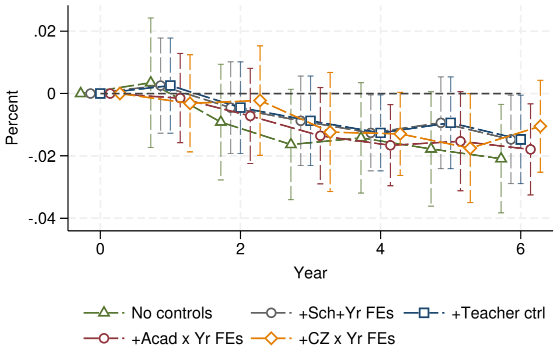
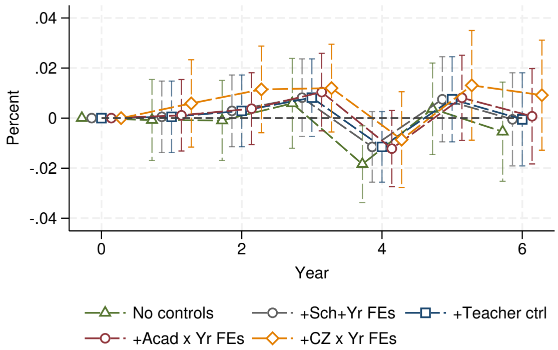


Figure B.11: Differential exit probability for REP+ teachers with pre-2013 school controls

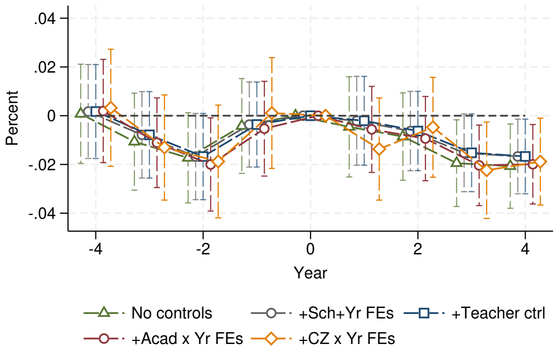
Note: The figure presents the results of the event study difference-in-difference regression 6, adding pre-2013 average school control variables multiplied by yearly dummies to the baseline specification, for the sample of all teachers. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience terciles, age terciles, and subject specialization. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.



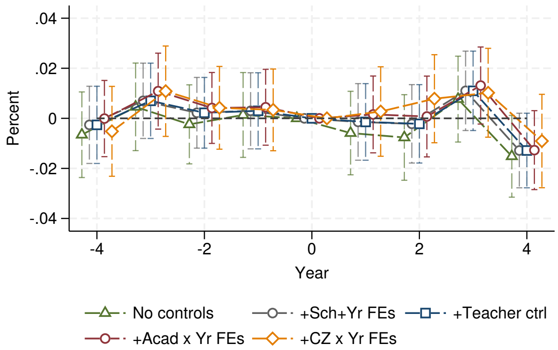
(a) High TVA (ITT)



(b) Low TVA (ITT)



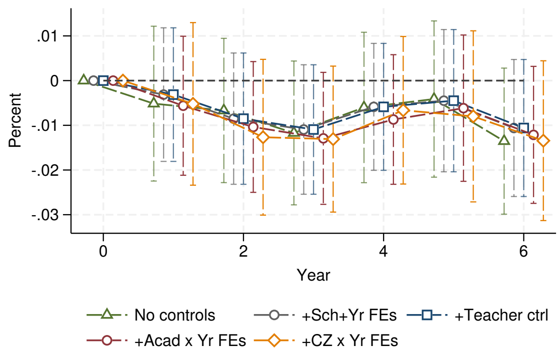
(c) High TVA (ATT)



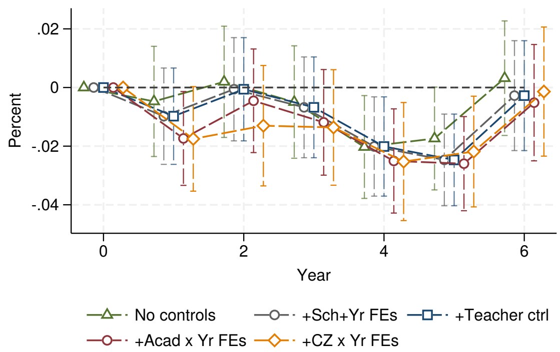
(d) Low TVA (ATT)

Figure B.12: Differential exit probability for REP+ compared to REP teachers by Math TVA

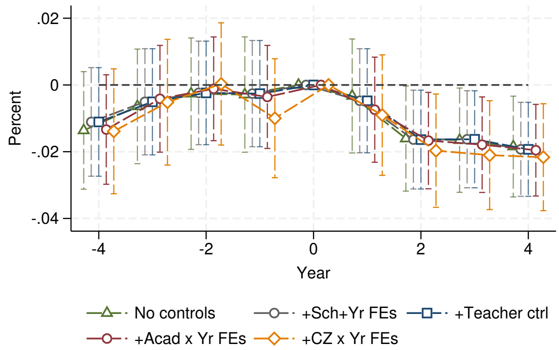
Note: The figures present the η_k coefficients of the event study difference-in-difference regressions 2 and 6, sequentially adding control variables and fixed effects, for the sample of (a) high TVA and (b) low TVA Math REP+ and REP teachers. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience, and subject specialization. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.



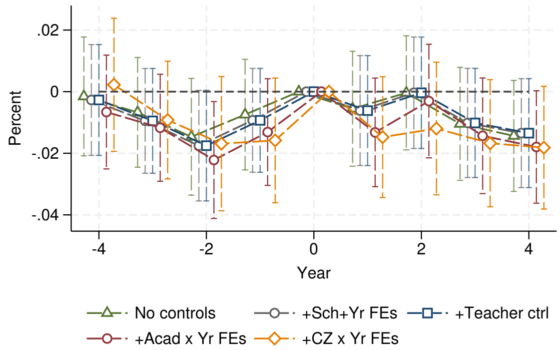
(a) High TVA (ITT)



(b) Low TVA (ITT)



(c) High TVA (ATT)



(d) Low TVA (ATT)

Figure B.13: Differential exit probability for REP+ compared to REP teachers by French TVA

Note: The figures present the η_k coefficients of the event study difference-in-difference regressions 2 and 6, sequentially adding control variables and fixed effects, for the sample of (a) high TVA and (b) low TVA French REP+ and REP teachers. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience, and subject specialization. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.

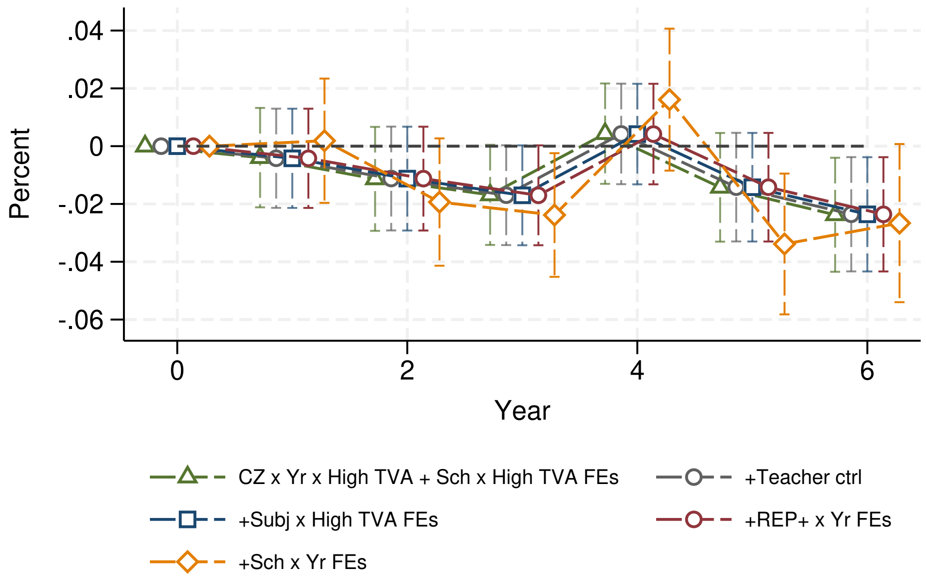


Figure B.14: Differential exit probability for high-TVA REP+ teachers

Note: The figures present the η_k coefficients of the event study triple-interaction difference-in-difference regression 7 with year dummies instead of a $Post_t$ dummy, sequentially adding control variables and fixed effects. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. Teacher-by-year level controls include gender, educational qualifications, teaching experience, and subject specialization. All controls and fixed effects are interacted by the $HighTVA_{jm}$ dummy. The confidence intervals correspond to 95% confidence levels. Standard errors are clustered by teacher and school-by-year. Underlying data source: Bases Relais, 2013-2014 to 2019-2020.

C. Technical appendix

C.1. The Godechot et al. (2023) method for creating an exhaustive panel of tax returns data

To extend the BTS-Postes into a longitudinal dataset covering the full 2009–2021 period, I apply the methodology proposed by Godechot et al. (2023).

The database has a built-in two-year panel structure: for each yearly file, information is reported not only for the current year but also retrospectively for the previous year. This design enables one to follow individuals over short horizons, but does not directly produce a long panel. Having a longer panel than two years is particularly useful in this context, since teachers who change careers may pass through temporary or transitional jobs before settling into more permanent employment. Focusing only on the first post-teaching job could, therefore, potentially misrepresent the types of occupations teachers ultimately move into.

To extend the BTS-Postes into a longitudinal dataset covering the full 2009–2021 period, I apply the methodology proposed by Godechot et al. (2023). Their approach exploits the overlap in variables reported across consecutive yearly files. Specifically, observations for a given individual in year t of file $y-1$ can be matched to the corresponding records in year $t-1$ of file y using a set of common identifiers, including establishment ID, demographic information (gender and age), and job characteristics (e.g., hours worked, job duration, start and end dates, municipality of residence and workplace, and salary). This matching procedure achieves a unique link for the vast majority of individuals—around 98% over my sample period, while unmatched cases typically reflect either data modifications (which reduced the number of matches to 92% for 2016 and 2017), or employment interruptions exceeding one year (as the dataset does not contain unemployment spells). The latter are kept in the data but recorded as distinct individuals (who then appear under multiple identifiers).

C.2. Bayesian shrinkage of TVA estimates

An Empirical Bayes shrinkage method outlined in Tartova (2023) is then needed to deal with the fact that teachers are often observed in few classrooms, making estimates likely to be noisy. The shrinkage method is inspired by the work of Kane and Staiger (2008) and Chetty et al. (2014a). Specifically, to compute the TVA estimates for Math teachers, I first take the residual from the regression 1 that purges the effect of the French teacher effects and observed covariates from ΔA_i^{*MF}, and denote it $\widehat{A}_i^{MF}$:

$$\widehat{A}_i^{MF} \equiv \widehat{\mu}_m + \widehat{\varepsilon}_i$$

To then get from these student-level estimates to TVA, first create teacher-year averages of $\widehat{A}_i^{MF}$ which are the minimum variance unbiased estimates of μ_m for each teacher m:

$$\widehat{A}_m = \sum_t w_{mt}\widehat{A}_{mt} \text{ where } w_{mt} = \frac{h_{mt}}{\sum_t h_{mt}} \text{ and } h_{mt} = \frac{1}{Var(\widehat{A}_{mt}|\mu_m)}$$

where $Var(\widehat{A}_{mt}|\mu_m)$ is the conditional variance of $\widehat{A}_{mt}$.

To estimate the variance of the teacher-specific fixed effect, σ_μ^2, Tartova (2023) first calculates the covariance of a teacher's residualized performance across two randomly chosen years, applying weights based on the number of students taught in each year:

$$Cov(\widehat{A}_{mt}, \widehat{A}_{mt'})$$

To isolate and subtract any residual variance attributable to the link teacher, the paper then computes the covariance of the link teacher's residual across the same two years:

$$\sigma_f^2 = Cov(\widehat{A}_{ft}, \widehat{A}_{ft'})$$

The variance of the fixed teacher effect is then identified as the difference between the two covariances:

$$\sigma_\mu^2 = Cov(\widehat{A}_{mt}, \widehat{A}_{mt'}) - \sigma_f^2$$

Finally, using $\widehat{A}_m$ and σ_μ^2, I construct an empirical Bayes estimator for each teacher's TVA by multiplying this weighted average residual by an estimate of its reliability (the signal-to-noise ratio):

$$\widehat{TVA}_m = \widehat{A}_m \left(\frac{\sigma_\mu^2}{\sigma_\mu^2 + 1/\sum_t h_{mt}} \right)$$

where σ_μ^2 is the variance of the Math teacher effect. The estimates $\widehat{TVA}_m$ are forecast unbiased predictors of the true Math teacher effects. The same method can be performed to estimate the TVA of French teachers.

C.3. De-trended difference-in-difference estimation

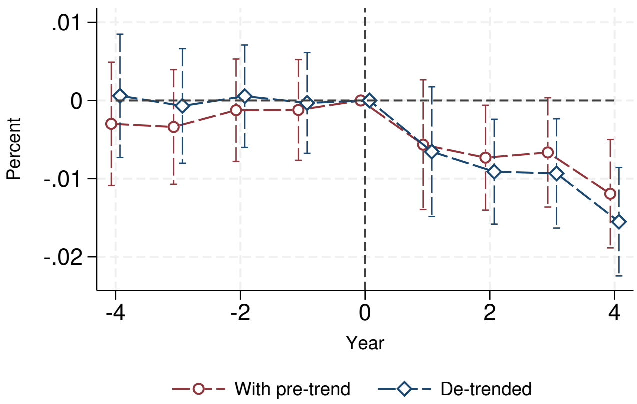
The assumption that the relative pre-trends observed in the data would have continued in the post-period in the absence of the reform prompts the use of a de-trended event-study difference-in-differences strategy, following Bhuller et al. (2013), Goodman-Bacon (2018, 2021), and Dobkin et al. (2018). Specifically, I estimate group-specific linear trends using only pre-2013 observations and subtract these from post-2013 outcomes. This approach relaxes the conventional parallel trends assumption by allowing REP+ and REP teachers to differ along a linear pre-trend, and interprets post-2013 coefficients as deviations from this extrapolated trajectory. Concretely, I first estimate a model which includes a linear differential trend in exit probabilities for REP+ relative to REP schools in the pre-reform sample period (2009–2013):

$$exit_{jst} = \alpha + \lambda\,\tilde{t} + \zeta\left(\tilde{t} \times \mathbb{1}\{type_s = REP+\}\right) + \mathbf{X}'_{jt}\gamma + \theta_s + \lambda_{zt} + \varepsilon_{jst}, \qquad t \in [2007, 2013] \tag{C.1}$$

in order to predict counterfactual exit rates in the post-period had this differential linear pre-trend continued (denoted as $\widehat{exit}^{T}_{jst}$), that is then used to define the de-trended exit variable as $\widehat{exit}^{DT}_{jst} \equiv exit_{jst} - \widehat{exit}^{T}_{jst}$. Using this definition, I run:

$$\widehat{exit}^{DT}_{jst} = \sum_{2009}^{2017} \eta^{DT}_{k}\mathbb{1}(t = k) \times \mathbb{1}(type_s = REP+) + \mathbf{X}_{jt}\gamma + \theta_s + \lambda_{zt} + \varepsilon_{jst} \tag{C.2}$$

Counterfactual exit rate The counterfactual exit rate in the de-trended difference-in-differences is defined as the average exit rate that REP+ teachers would have had over the period 2014-2017, had they not experienced a differential increase in wages compared to REP teachers, under the assumption that the differential trend in exit rates between REP and REP+ teachers would have continued on the same linear trend. Specifically, the counterfactual exit rate (hereafter $\overline{exit}^{CF2}_{REP+}$) under this assumption is estimated in three steps. First, I predict the counterfactual exit rate ($\widehat{exit}^{T}_{jst}$) had the differential pre-trend (obtained by running regression C.1 without controls or fixed effects) continued. Second, I recover the counterfactual post-reform for REP+ teachers had there been no reform, by running $\widehat{exit}^{T}_{jst} = \alpha + \beta_1 Post_t \times REP+_s + \beta_2 Post_t + \beta_3 REP+_s + \varepsilon_{jst}$ over the full sample period $t \in [2007, 2017]$, as $\overline{exit}^{P}_{REP+} \equiv \alpha + \beta_1 + \beta_2 + \beta_3$. Third, I run regression C.2 without controls and fixed effects in order to obtain the REP+ post level in exit rate one would expect after de-trending, if REP+ and REP evolved the same way post-reform ($\overline{exit}^{P,DT}_{REP+} \equiv \alpha' + \beta'_2 + \beta'_3$), and compute the counterfactual REP+ exit rate as $\overline{exit}^{CF2}_{REP+} \equiv \overline{exit}^{P}_{REP+} + \overline{exit}^{P,DT}_{REP+}$.



(a) All

Figure C.15: Differential exit probability for REP+ compared to REP teachers (de-trended DID)

Note: The figures compare the η_k coefficients of the baseline event study difference-in-difference regression 6 to the η_k^{DT} of the de-trended event study difference-in-difference regression C.2, for the sample of all REP+ and REP teachers. All regressions include school and commuting-zone-by-year fixed effects and teacher-by-year level controls. Teacher-by-year level controls include gender, educational qualifications, teaching experience terciles, age terciles, and subject specialization. Year 0, which coincides with the vertical dashed line, represents the school year 2013-2014, when the reform was first announced. The confidence intervals correspond to 95% confidence levels. Underlying data source: Bases Relais, 2009-2010 to 2017-2018.

C.4. Explaining the teaching hours leg of the reform

In this section, I explain in more detail how the part of the reform which aimed to reduce the number of mandatory teaching hours and redistribute them towards time spent with students and parents operates and how it translated into more working hours and therefore overtime pay (*Heure Supplémentaire Annuelle*, or HSA). I do this by providing two examples for a Capes-certified teacher with a statutory 18-hour teaching obligation.

A Capes teacher with 18 hours of baseline teaching obligation who teaches 18.7 hours pre-reform Prior to the reform, a teacher who taught 18.7 hours per week received annual overtime pay equivalent to 0.7 HSA. In addition to teaching hours, the 17 additional hours of work of the 35-hour working week were spent in other activities than teaching. It follows that the average REP+ Capes teacher was working on 35.7 hours per week.

After the reform, each hour of teaching in a REP+ school counted as 1.1 hours for the calculation of service. Thus, if the teacher continued to do the same amount of overtime, his hours would be counted as

$$18/1.1 + 0.7 = 17.06 \text{ hours.}$$

The 1.65 hours were to be redistributed to tasks such as liaison with families, which was one of the reform's goals. Such a redistribution would have entailed no additional working hours compared to before the reform, with a total of 35.7 hours per week worked.

If, in practice, the teacher's teaching load was reduced by only 0.74 hours instead of the full 1.65 (see Table A.6), if the teacher does not take any additional overtime work, they would in reality be teaching:

$$18/1.1 + (1.65 - 0.74) = 17.27 \text{ hours.}$$

which would entail 17.27+1.65+17 = 35.92 hours of work per week. It follows that the teacher is teaching more than their pre-reform preferred optimum of 35.7 hours.

For the number of teaching hours, his teaching hours would be counted as:

$$17.27x1.1 = 19 \text{ hours,}$$

which entails the teacher would be earning 1 HSA after the reform—their work would be counted as 36 hours in total, for the 35.92 hours of actual work they do.

If this teacher was to work the same amount of hours pre-reform, therefore teaching for 18.92 hours, they would have got 0.92 HSA in overtime pay.

It follows that the reform introduced a small monetary incentive to work more, equivalent to 36 (hours remunerated) - 35.92 (hours actually worked) = 0.08 HSA. Between 2014 and 2019, each first HSA was paid around 1,290, with each subsequent reduced to 1,070. For the average teacher, this would entail between

$$0.08 \times 1{,}290 = 103.2 \text{ euros more annually.}$$

This positive difference suggests that the reform slightly increased monetary incentives to take on additional teaching hours, even in the absence of institutional constraints requiring teachers to do so. Specifically, if a teacher could decide to work more in response to the reform, every overtime hour they would would be remunerated slightly more highly than it would have been pre-reform. In the case of someone who works 35.92 hours in total, they would get 103.2 euros more annually than they would have for the same 0.92 number of hours of overtime work prior to the reform.

C.5. Computing the wage gap

Using the set of all outside-option occupations, $\mathcal{O}$, I define the outside-option wage of a teacher of a given sex g, age group a, who teaches in a given commuting zone z in year t, as:

$$OutsideOptionWage_{gazt} = \frac{\overline{S_o} \cdot D_{ogazt}}{\sum_o \overline{S_o} \cdot D_{ogazt}} \cdot OutsideOptionWage_{ogazt} \tag{C.3}$$

where occupation $o \in \mathcal{O}$, $\overline{S_o}$ is the average share of teachers exiting to o over the period 2009-2021, and D_{ogazt} is the share of workers at o in a given cell $gazt$.

A concern for the estimation of the counterfactual outside-option wage for teachers might be that the wages actually attained by teachers transitioning to a given occupation may be lower than the those of incumbents. Figure C.16 provides evidence of a strong and close to 1-to-1 correlation between (i) the mean wage in a given occupation o for individuals of a given sex g, age group a, commuting zone z and year t, and (ii) the mean wage that teachers exiting to an occupation o obtain 3 years after exit, within the same sex g, age group a, commuting zone z and year t. This suggest that the average outside-option wage, $OutsideOptionWage_{gazt}$, provides a good estimate of the true average local outside-option wage for teachers of a given sex g and age group a. As exiting teachers may also move to a different labor market (and indeed, roughly 40% do), I also estimate a version of the outside-option wage which takes into account national, rather than local, weights of labor demand and wages:

$$OutsideOptionWage_{gat} = \frac{\overline{S_o} \cdot D_{ogat}}{\sum_o \overline{S_o} \cdot D_{ogat}} \cdot OutsideOptionWage_{ogat}$$

This alternative measure of the outside-option is highly correlated with the local one, but predicts a higher estimate of the teacher wage gap than that predicted from the local outside-option wage. As a result, I take the the local estimate of the outside wage, $OutsideOptionWage_{gazt}$, as my main, more conservative, estimate. For moving teachers, $OutsideOptionWage_{gazt}$ may be lower than their true counterfactual outside-option wage.

One caveat of the database is that it does not allow to distinguish between exiting teachers' taught subject. However, as the majority of occupations are general and can be performed by both Math and French teachers, this is less concerning.

Measurement of the wage gap I match the outside-option wage defined in equation C.3, $OutsideOptionWage_{gazt}$, and the teaching wage that varies by a teacher's position on the wage scale e, school type s (REP/REP+/ordinary), and year t, to my teacher panel data, in order to obtain teacher-specific wage gaps that vary by e, sex g,

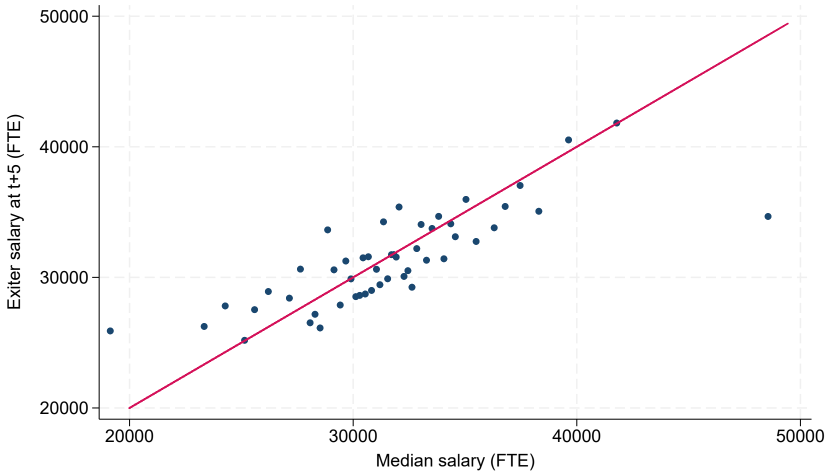


Figure C.16: Correlation between wages at $t+5$ of teachers exiting to an occupation and salaries of incumbents

Note: The figure is presents the binscatter between (i) the mean wage in a given occupation o for individuals of a given sex g, age group a, commuting zone z and year t, and (ii) the mean wage that teachers exiting to an occupation o obtain 5 years after exit, within the same sex g, age group a, commuting zone z and year t. The red line is the 45-degree line. Underlying data source: Insee, BTS Postes, 2009-2021.

age group a, commuting zone z and year t:

$$WageGap_{gazest} = OutsideOptionWage_{gazt} - TeachingWage_{est} \quad \text{(C.4)}$$

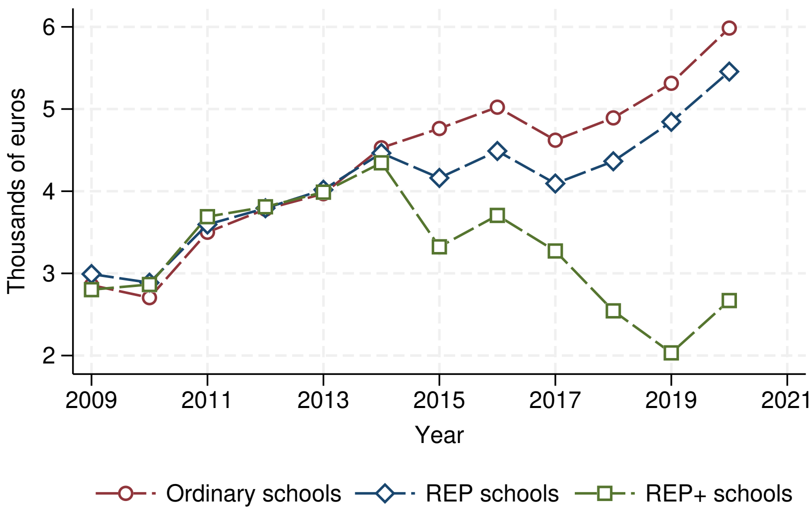


Figure C.17: Aggregated annual wage gap

Note: The figure presents the average annual wage gap in my sample per school type (computed in thousands of euros), computed as $WageGap_{gazest} = OutsideOptionWage_{gazt} - TeachingWage_{est}$, for each age group a, gender g, commuting zone z, echelon e, school type s and year t. School year t—$t+1$ is denoted as t. School year t—$t+1$ is denoted as t. Both teaching wages and outside option wages are adjusted for inflation (with 2007 as base year). Underlying data source: Bases Relais, BTS Postes, Insee CPI, https://lucaschancel.com/enseignants/.

Figure C.17 depicts the estimated average wage gap per type of school over time, underlining the increasing trend of the wage gap prior to the reform and the importance of the treatment introduced by it (see Figures C.18 and C.19 below for the trend in each of the two components of the wage gap by school type). The figure shows that during the school year 2013-2014, the average wage gap was roughly 4,000 euros. It continued to increase thereafter for ordinary-school teachers, despite a small increase in all novice teacher wages during the 2017-2018 school year, reaching 6,000 euros by 2020-2021. For REP-schools teachers, the wage gap also continued to increase, reaching 5,500 euros by 2020-2021. For REP+-school teachers, the average wage gap dropped to 2,800 euros by 2020-2021, representing a sizable drop in the wage gap for treated teachers.

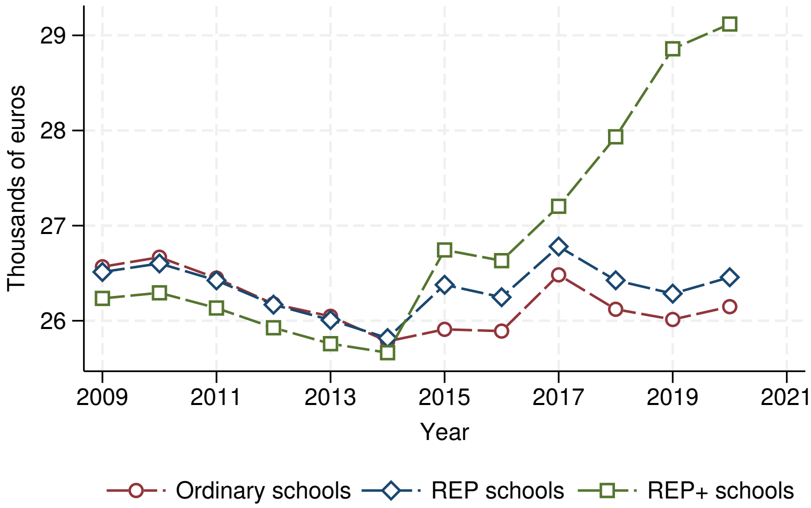


Figure C.18: Average annual teacher wage

Note: The figure presents the average teacher wage in my sample per school type (computed in thousands of euros), varying by echelon e, school type s and year t. School year t—$t+1$ is denoted as t. Wages and are adjusted for inflation (with 2007 as base year). Underlying data source: Bases Relais, Insee CPI, https://lucaschancel.com/enseignants/, 2009-2010 to 2020-2021.

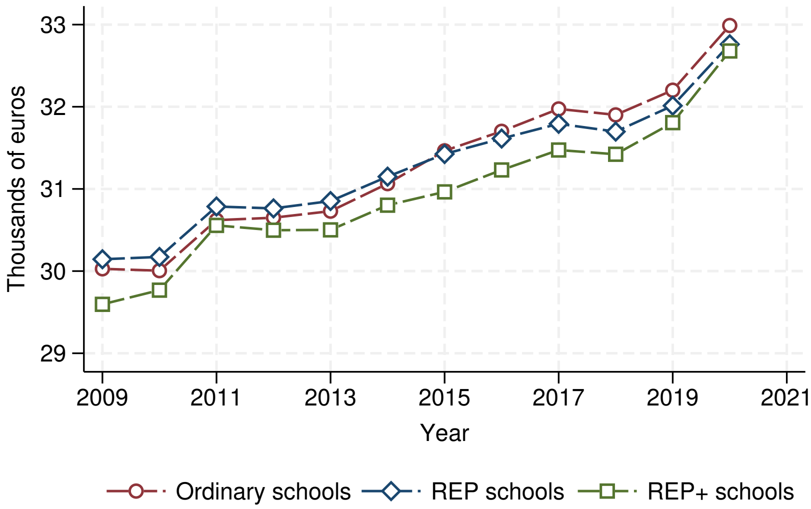


Figure C.19: Average annual outside-option wage

Note: The figure presents the average outside option wage in my sample per type of school (computed in thousands of euros), varying by age group a, sex g, commuting zone z and year t. School year t—$t+1$ is denoted as t. Wages and are adjusted for inflation (with 2007 as base year). Underlying data source: Bases Relais, BTS Postes, Insee CPI, 2009-2010 to 2020-2021.