Teaching Teachers

Danilo Kuzmanic[a,b,*] , Per Engzell[a] , John Jerrim[a]

[a] Social Research Institute, UCL Institute of Education, University College London, 20 Bedford Way, London WC1H 0AL, United Kingdom
[b] Centro de Investigación Avanzada en Educación (CIAE), Universidad de Chile, Periodista José Carrasco Tapia 75, Santiago, Chile

[*] Corresponding author: Danilo Kuzmanic
Social Research Institute, UCL Institute of Education, University College London, 20 Bedford Way, London WC1H 0AL, United Kingdom
Email: danilo.kuzmanic@uchile.cl

Acknowledgments

We thank the Chilean Ministry of Education, the Agencia de Calidad de la Educación, and the Departamento de Evaluación, Medición y Registro Educacional (DEMRE) for providing access to the data used in this study. The findings and conclusions reported in this paper are solely those of the authors and do not necessarily reflect the views of these institutions. Danilo Kuzmanic received funding from ANID (grant numbers FB0003 / Support 2024 AFB240004 and 72230237) during the preparation of this manuscript. The funding source had no involvement in the study design; data collection, analysis, or interpretation; writing of the manuscript; or the decision to submit the article for publication. The authors declare no competing interests.

Abstract

Using Chilean administrative data, we estimate how teacher value-added (VA) affects the VA of students who later become teachers. Exposure to a math teacher with 1 SD higher VA raises teacher VA in the next generation by 0.14 SD and student achievement one generation removed by 0.04 SD. Effects compound with longer exposure and appear driven by observable classroom practices. Under plausible assumptions, these spillovers raise cumulative returns to instructional quality by 28 percent after one generation and 39 percent in the long run. (*JEL:* I21, I28, J24.)
Keywords: Teacher value-added, spillovers, human capital, high school, Chile.

1. Introduction

Valuing teachers solely by their direct impact on students' learning might overlook a crucial channel of human capital formation. High-quality teachers not only influence students directly but also cultivate the next generation of teachers. This study examines the effect of high school teachers' value-added (VA) on the VA of their students who later become teachers. We refer to teachers who influence the next generation in this way as "grandteachers," and show that ignoring this spillover effect has led research and policy to underestimate the returns to teacher quality.

There are at least two ways grandteachers could influence the future teacher workforce. First, they might attract talented students into teaching, shaping the quality of recruited entrants. Second, they may enhance students' teaching effectiveness regardless of recruitment, through skill transfer and role modeling. The distinction matters: if recruitment were the main driver, effects might be zero-sum, diverting talent from other sectors. We find limited evidence of a recruitment effect; instead, the influence of grandteachers manifests through improved teaching effectiveness.

We test the robustness of this effect by exploiting variation in students' cumulative exposure to high-VA teachers. If unobserved factors such as student sorting drove the correlation, we would not expect it to scale smoothly with exposure. Instead, we find a continuous gradient: each additional year with an effective grandteacher increases future teacher VA and the performance of grandstudents (i.e., the students of teachers taught by grandteachers). A back-of-the-envelope calculation suggests spillovers raise cumulative returns to teacher quality by 39% in the long run.

To explore transmission mechanisms, we analyze evaluation data on instructional practices through written assessments and video recordings on two generations of teachers. The strongest correlations between grandteachers and their "descendant" teachers emerge in classroom management and lesson structure, both observable, practice-based skills that are central to instructional productivity. Correlations are weaker or non-existent in less visible domains, such as pedagogical reasoning or written planning. These patterns suggest that productivity predominantly transmits through direct observation and experience, consistent with a role modeling explanation.

Our findings bridge several literatures. We connect research on teacher VA (Aaronson et al., 2007; Bacher-Hicks and Koedel, 2023; Chetty et al., 2014a; Jackson, 2018; Mansfield, 2015) to studies on workplace spillovers and peer learning, where exposure to skilled colleagues raises future productivity (Cornelissen et al., 2017; De Grip and Sauermann, 2012;

Mas and Moretti, 2009), including among teachers (Goldhaber et al., 2020; Jackson and Bruegmann, 2009; Kraft et al., 2018; Papay et al., 2020). Ample research documents intergenerational spillovers in education (Björklund and Salvanes, 2011; Braun and Stuhler, 2018; Holmlund et al., 2011), while a newer literature explores teachers' role in driving cultural and technological change (Borowiecki, 2022; Koschnick, 2025). Relatedly, evidence on peer effects in education emphasizes classroom climate and behavior (Bietenbeck, 2025; Carrell et al., 2018; Carrell and Hoekstra, 2010; Feng et al., 2025; Golsteyn et al., 2021; Kraft et al., 2016; Shure, 2021), which emerge as key drivers in our study. Building on these insights, we document an overlooked intergenerational channel of human capital formation, where teacher VA propagates across generations of teachers.

The remainder of the paper is organized as follows. In Section 2 we define the key parameters of interest to this study. Section 3 describes the institutional context and data, and Section 4 outlines the empirical strategy. Section 5 presents the main results, beginning with baseline VA estimates, evidence of small or absent recruitment effects, and a substantial influence on teacher VA. We follow this with placebo and robustness checks, including a continuous dose–response analysis. Section 6 provides a back-of-the-envelope calculation of the social multiplier. In Section 7, we exploit the multigenerational data structure to propose an adjustment for measurement error. Section 8 explores mechanisms using teacher evaluation data, suggesting role modeling as a key channel. Section 9 concludes.

2. Key parameters, identification, and estimation

The aim of this paper is to assess by how much conventional estimates of teacher VA underestimate the extended impact of teachers considering students indirectly exposed. To formalize this, we define three key parameters. The first captures teachers' direct contributions to student learning (φ_1), the second measures how much exposure to effective teachers raises the effectiveness of future teachers (ρ), and the third summarizes the resulting spillover to the students taught by those future teachers (φ_2).

The *baseline VA* parameter, φ_1, captures the direct effect of teacher effectiveness on student achievement, assumed linear:

$$\varphi_1 = \frac{\partial Y_{ijct}}{\partial VA_j},$$

where Y_{ijct} denotes the achievement of student i, taught by teacher j in classroom c and year t, and VA_j is teacher j's VA, standardized to mean zero and unit variance. Teacher VA is estimated using established approaches: in our main specification, the empirical Bayes two-step procedure of Kane and Staiger (2008). Identification relies on the assumption that, conditional on prior achievement and observables, assignment to teachers is as good as random. Our results are highly robust once we include school and classroom level means of control variables, consistent with a formal result that group means of observables absorb across-group variation in unobservables (Altonji and Mansfield, 2018).

The second parameter, ρ, captures the *intergenerational transmission* of VA which measures how much a teacher's own effectiveness is influenced by exposure to a high-VA teacher

while a student:

$$\rho = \frac{\partial VA_j}{\partial VA_{g(j)}},$$

where $VA_{g(j)}$ denotes the VA of teacher j's teacher, or "grandteacher." We estimate ρ by linking three generations of teachers and students, first estimating VA for the grandteacher and teacher generations using standard methods, and then correlating these VA estimates across generations.

Identification of ρ relies on the same selection-on-observables assumption used for baseline teacher VA, together with the assumption that grandteacher VA ($VA_{g(j)}$) does not influence students' selection into the teaching profession. We corroborate this latter assumption by showing that any selection effects are small and implementing an inverse-probability weighting strategy, which leaves our estimates unchanged.

As we discuss in Section 6, the total social spillover across generations can be recovered from just two quantities: the intergenerational transmission coefficient ρ, and the number of students per teacher who go on to become teachers. However, VA estimates and the resulting estimate of intergenerational transmission are likely attenuated by measurement error. To test this, we directly estimate VA models linking grandteachers to their grandstudents' outcomes, allowing us to estimate the *indirect VA* after one generation:

$$\varphi_2 = \frac{\partial Y_{ijct}}{\partial VA_{g(j)}}.$$

Assuming an autoregressive process without memory—which is natural, as grandteachers and grandstudents have no direct exposure—it follows that $\varphi_2 = \rho\varphi_1$. With perfect measurement, this identity would extend to the estimated parameters $\hat{\varphi}_1, \hat{\rho}, \hat{\varphi}_2$. In practice, departures from this identity are likely to result from measurement error in VA. We explore this discrepancy in Section 7 to propose a bias-corrected magnitude of intergenerational spillovers.

3. Setting and data

The empirical setting of this study is Chile, covering all three branches of its secondary system: public schools (40% of enrolment), private subsidized (53%), and private paid schools (7%). In our analysis, we focus on teachers' VA in the 10th grade. Although secondary education is compulsory until 12th grade, 10th is the last grade where students are exposed to a common curriculum. A more detailed description of Chilean secondary education, including aspects related to teacher preparation, evaluation, and professional development, is provided in Appendix A.

3.1. SIMCE assessments

To measure teacher VA, we use *Sistema Nacional de Medición de la Calidad de la Educación* (SIMCE), Chile's national standardized assessment in math and reading. SIMCE is applied nationwide in grades 2, 4, 6, 8, and 10. However, because not all grades are tested every year, only some students have scores in multiple grades. To ensure lagged achievement for estimating VA, we focus on four cohorts with both 8th- and 10th-grade scores: those in

10th grade in 2006, 2013, 2015, and 2017. Characteristics of these cohorts are reported in Appendix Table B1.

As described below, we rely on 2006 estimates to measure grandteachers' effectiveness between 2004 and 2008. We focus on math because reading VA is considerably less stable over time (Table F1 in Appendix F), which likely leads to severe mismeasurement when assigning estimates across years. In contrast, math VA is more stable and is estimated reliably enough to pass standard validation checks (Appendix F). [1]

3.2. Additional student data

We complement the SIMCE data with several administrative sources on students. First, publicly available Ministry of Education records provide each student's school and assigned teacher over time, as well as annual grade point average (GPA) and attendance. Second, the Department of Educational Evaluation, Measurement, and Records supplies information on 12th-grade students' test scores from the national university entrance exams. These scores allow us to estimate longer-term effects of teacher VA. Third, Ministry of Education administrative records provide information on higher education enrollment and completion, covering the institution, degree program, and year of graduation, allowing us to trace students into the teaching profession.

3.3. Teacher evaluations

We also use data from the National Teacher Evaluation System (NTES). Between 2005 and 2015, this evaluation system was mandatory for all teachers in public schools and included a teaching portfolio (video observations and written materials), a peer interview, and supervisor and self-evaluations. Such performance ratings contain detailed information that validly predicts teacher effectiveness (Bacher-Hicks et al., 2019; Kraft et al., 2020). Appendix C provides further details of the NTES assessments. Using portfolio scores for both generations of teachers, we explore the extent to which teachers' classroom performance reflects the practices they observed from grandteachers during their schooling.

3.4. Sample construction

Figure 1 summarizes the data structure. We construct the *grandteacher* sample by estimating VA for all 10th-grade math teachers in 2006. We then identify students who attended 10th grade two years earlier or later (2004, 2005, 2007, and 2008) and were taught by one of these grandteachers, defining them as the *index* cohort. Excluding the 2006 cohort avoids mechanical correlation between grandteacher VA and the outcomes of the index cohort. Some index students later become math teachers themselves, allowing us to construct the *grandstudent* sample, which includes 10th graders taught by these former students in 2013, 2015, or 2017.

This timeline implies that we follow students who were active as teachers nine to ten years after we observe them in the 10th grade. Completing teacher training in Chile typically takes

[1] The SIMCE math test consists of multiple-choice questions assessing mathematical knowledge and problem-solving skills.

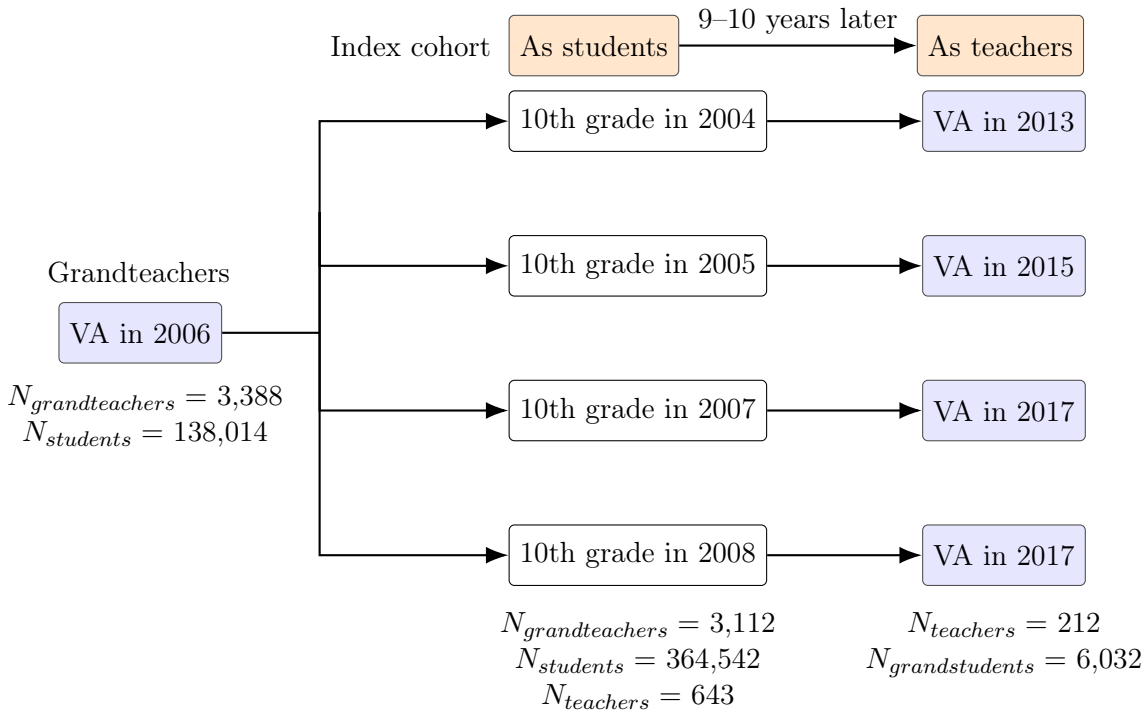


Figure 1: Data structure and timeline.

Note: We estimate grandteacher VA in 2006 for 3,388 math teachers and 138,014 students in 10th grade. Of these grandteachers, 92% (3,112 teachers) also appear as teachers of the index cohort, which includes 364,542 students. Among index cohort students, 643 later become math teachers (approximately 0.2%). Of these, 33% are observed as teachers of the grandstudent cohorts (2013, 2015, and 2017), comprising a final grandstudent sample of 212 teachers and 6,032 students.

at least seven years after 10th grade, so teachers in these cohorts are likely to have no more than three years of experience when their VA is estimated. We focus on this early-career stage to avoid attrition from the teaching profession, which can lead to bias if poorer performing teachers are more likely to leave the teaching profession (Goldhaber et al., 2020; Harris and Sass, 2011).

About 60% of 10th graders in 2004, 2005, 2007, and 2008 had a teacher with a VA estimate from the 2006 SIMCE, and these students resemble the full student population on observed characteristics (Appendix Table B2). Among them, roughly 7% later completed an education-related degree. Future high school math teachers represent about 2.5% of those with an education degree, and 0.2% of index students ($N = 643$). These future teachers tend to come from average-SES, high-performing schools and were themselves high achievers, particularly in the subject they later teach (Appendix Table B3). Our main analysis focuses on those who taught 10th grade in 2013, 2015, or 2017, when grandstudents took the SIMCE, comprising 212 teachers and 6,032 grandstudents. These teachers are comparable to the broader population of index teachers in terms of socioeconomic background, educational trajectories, and school characteristics (Appendix Table B3).

4. Value-added model

We estimate teacher VA on all 10th-graders in 2006, 2013, 2015, and 2017 using a standard VA model (VAM). Let Y_{ijct} denote the standardized 10th-grade math test score for student i, taught by teacher j in classroom c and year t. The model controls for prior achievement in both reading and math Y_{it-1}, individual characteristics X_{it}, the lagged mean performance of classmates $\bar{Y}_{ct-1}$, classroom-level characteristics $\bar{X}_{ct}$, and school-level characteristics Z_{st}. The estimating equation is:

$$Y_{ijct} = \beta_0 f\left(Y_{it-1}\right) + \beta_1 X_{it} + \gamma_0 f\left(\bar{Y}_{ct-1}\right) + \gamma_1 \bar{X}_{ct} + \gamma_2 Z_{st} + v_{ijct}, \tag{1}$$

where the composite error term v_{ijct} captures unobserved teacher, classroom, and student-level components:

$$v_{ijct} = \mu_{jt} + \theta_{ct} + \varepsilon_{ijct}.$$

Specifically, μ_{jt} is a teacher-specific random effect representing teacher VA with variance $\sigma^2_{\mu t}$, θ_{ct} is a classroom-specific random effect capturing exogenous class shocks with variance $\sigma^2_{\theta t}$, and ε^s_{ijct} is an idiosyncratic student-level error with variance $\sigma^2_{\varepsilon t}$. Hence, $\text{Var}(v_{ijct}) = \sigma^2_{\mu t} + \sigma^2_{\theta t} + \sigma^2_{\varepsilon t}$. In this specification, μ_{jt} is year-specific, which allows us to estimate grandteachers' effectiveness in 2006, around the time they taught the students who later became teachers. In Appendix E, we show that our results are similar when we treat teacher VA as time-invariant fixed effects.

4.1. Control variables

We include 8th-grade math and reading scores Y_{it-1} through a third-degree polynomial $f(\cdot)$. Student-level covariates X_{it} include gender, parental years of schooling, log household income, and GPA and attendance in the 8th and 9th grades.[2] Moreover, we control for students' school transitions (from 8th to 9th and 9th to 10th grade), accounting for the instability of their prior trajectories, and whether they had the same teacher in 9th grade (45% of students). Appendix Table B1 describes these variables.

The model further includes classroom-level means of the lagged outcomes and student characteristics ($\bar{Y}_{ct-1}$ and $\bar{X}_{ct}$), along with school characteristics (Z_{st}). These last variables include the type of school administration (public, private subsidized, private paid), region, enrollment (second-degree polynomial), and the across-time average shares of low (primary) and high (university) educated parents in the school. Specification curve tests in Appendix Figure B1 show that results are stable across specifications that use *school- and classroom means* of contextual variables as controls, consistent with Altonji and Mansfield (2018).

4.2. Bayes two-step approach

In our preferred specification, we estimate teacher VA using the empirical Bayes two-step procedure in Kane and Staiger (2008). In the first step, we estimate equation (1) via OLS

[2]SIMCE data include parents' responses to a questionnaire collecting information on socioeconomic status. Parental education represents the total number of years of schooling completed by the parents. Household income is reported via a multiple-choice question with 15 income brackets. We impute the midpoint of each bracket.

and obtain residuals. In the second step, we average these residuals at the teacher-year level. Because most high school teachers (63%) teach multiple classrooms, we first average residuals within classrooms and then compute precision-weighted averages across classrooms (Chetty et al., 2014a). The final empirical Bayes estimates of each teacher's VA are obtained by multiplying the weighted average of teacher-year residuals by an estimate of its reliability. Appendix D provides further details and the resulting VA distribution. Appendix E reports one-step fixed-effects estimates, including models with school fixed effects, and shows that our main results are robust to these alternatives.

4.3. Validation checks

We report several validation checks in Appendix F. First, the correlation of teacher VA with its two-year lag is 0.44, suggesting a year-to-year reliability of $\sqrt{0.44} = 0.66$, near the upper bound of the 0.47–0.64 range reported by Bacher-Hicks and Koedel (2023). We address residual measurement error in Section 7, using a novel method to infer the degree of measurement error from the multigenerational structure of the data. Second, math teacher VA explains most (but not all) of the variation in schools' average math performance following the exit and entry of teachers with differing VA (cf. Chetty et al., 2014b). Lastly, teacher VA rises during the first decade of a teacher's career before flattening out, which is consistent with prior work (Ost, 2014; Papay and Kraft, 2015; Rivkin et al., 2005).

5. Main results

5.1. Baseline VA estimates

We begin by quantifying the variation in grandteacher VA and its implications for student achievement. We first estimate baseline VA in grade 10 of the 2006 cohort. We then extend this to later outcomes for the index cohort.

Assuming class shocks (θ_{ct}) and student-level noise (ε_{ijct}) in equation (1) are independent across classrooms, the baseline VA parameter φ_1 can be recovered from the covariance of average residuals across classrooms taught by the same grandteacher in 2006 (Chetty et al., 2014a; Jackson, 2018). Column (1) of Table 1 shows that the SD of grandteacher VA is 0.20, implying that a 1 SD difference in grandteacher quality raises student test scores by 0.20 SD, which corresponds to moving an average student from the 50th to roughly the 58th percentile. This value falls within the range (0.09–0.32) reported for US middle and high school teachers (Aaronson et al., 2007; Chetty et al., 2014a; Goldhaber et al., 2013; Jackson, 2018; Mansfield, 2015). We find similar results for 10th-grade teachers between 2013 and 2017, whose SD of VA is 0.21 (Appendix Table D1).

Our next question is whether teacher effects extend beyond 10th-grade test scores. We turn to the index sample, composed of 10th-graders up to two years before and after 2006, and specifically the 76% of these students who sat qualifying exams for higher education at

the end of 12th grade.[3] We estimate the model:

$$Y_{ijct} = \omega_0 + \omega_1 \widehat{VA}_{g(j)} + \omega_2 C_i + u_{ijct}, \tag{2}$$

where $\widehat{VA}_{g(j)}$ is the estimated VA of the grandteacher, standardized to mean zero and unit variance, and Y_{ijct} represents two outcomes: standardized math scores on the 12th-grade qualifying exam and eventual university graduation. For the latter, we report the average marginal effects estimated via a logit model. Control variables (C_i) correspond to those in equation (1), except index students did not take 8th- and 10th-grade SIMCE exams, so for prior achievement we rely on GPA and attendance, and additionally include school-level test scores and family background from the 2006 cohort.[4]

Column (2) of Table 1 shows that a 1 SD increase in 10th-grade grandteacher VA raises 12th-grade math scores by 0.04 SD. Column (3) shows that such students are 0.7 percentage points more likely to graduate from university, equivalent to 2% of the sample mean (36%). Within the selected sample of students graduating from higher education, we also find positive effects on graduating from a STEM degree (about 5% of the sample mean) and on expected post-graduation earnings, which increase by 1.2% (Appendix Table B4).[5] Appendix Table B5 confirms that including only students who took the university admission exam does not bias estimates, showing negligible effects on on-time high school graduation and exam participation for the full index sample.

5.2. Recruitment effects

Do effective teachers encourage students to become teachers? We test this by estimating equation (2) for two additional outcomes: completion of an education-related degree (described as such in the ISCED-13 classification) and becoming a high school math teacher. Column (4) of Table 1 shows no effect on earning an education degree, indicating that the earlier positive effect on university completion reflects degrees outside education. Additionally, although we find a statistically significant effect on becoming a math teacher, the magnitude is small: at the 95% confidence level, we rule out effects larger than 0.0001 percentage points (5% of the sample mean, 0.2%). Lastly, Appendix Figure B2 illustrates minor differences in the VA to which future 10th-grade teachers were exposed while studying the 10th grade compared to the overall index sample.[6]

[3] Because we condition on participation in the university admission process, we weight observations by the inverse of the estimated probability of participation, computed as a function of the model's control variables.

[4] The vector C_i includes index cohort, GPA and attendance in 8th and 9th grade, indicators for school changes, whether the student had the same teacher in 9th grade, and the classroom averages of these variables. At the school level, we control for school type, size, and a set of contextual variables imputed from the 2006 cohort: average 8th-grade math and reading SIMCE scores (third-degree polynomial), average household income and parental education, and time-averaged shares of low- and high-educated parents.

[5] These estimates are similar to those in Chetty et al. (2014b) for elementary and middle-school teachers in the US, where a 1 SD increase in teacher VA raises college attendance by 0.82 percentage points (relative to a mean of 37%) and earnings at age 28 by 1.3%.

[6] Our main findings remain unchanged when reweighting the teacher sample by their inverse estimated probability of entering our sample based on grandteachers' VA and their own within-school percentile rank (Appendix Table B6).

Even without significantly increasing the proportion of their students who become teachers, higher-VA grandteachers could still shape *who* enters the profession by attracting top achievers from their class into teaching. To test this, we examine whether students who became math teachers and were taught by higher-VA grandteachers ranked higher on the 12th-grade math exam than their classmates, i.e., have a higher within-school percentile rank (from 1 to 100). Column (6) of Table 1 shows no such association: grandteacher VA is unrelated to the relative math performance of their students who later became math teachers.[7] Overall, we find no evidence that high-VA grandteachers increase their students' probability of becoming teachers, which aligns with evidence showing that increased professional opportunities often encourage students to pursue higher-status employment opportunities outside teaching (Bacolod, 2007).

5.3. Intergenerational transmission of VA

We now turn to our main question of how VA is transmitted to future teachers, captured by the parameter ρ. First, to quantify the intergenerational transmission of VA, we regress the VA of teachers on the VA of grandteachers, both measures standardized to mean zero and unit variance across all teachers within the respective evaluation year. We then go on to estimate three-generation VA directly, by linking grandteachers to their grandstudents.

Column (1) of Table 2 reveals that having a 10th-grade teacher 1 SD above the mean increases future teachers' effectiveness by 0.14 SD.[8] This effect is comparable in magnitude to the improvement between a novice math teacher and one with 4–5 years of experience (Appendix F3). As illustrated in Appendix Figure B3, this effect is close to linear: teachers exposed to higher-VA grandteachers are likelier to have a higher VA, while exposure to lower–VA grandteachers is correspondingly associated with lower VA. Given that the SD of teacher VA is 0.20 (Appendix Table D1), we would expect the extended effect of grandteacher VA on grandstudent achievement to amount to 0.03 SD (0.20×0.14).

To directly test influence across three generations, we estimate a model similar to the VAM in equation (1), but linking grandstudent achievement directly to grandteacher VA:

$$Y_{ijct} = \beta_0 f\left(Y_{it-1}\right) + \beta_1 X_{it} + \gamma_0 f\left(\bar{Y}_{ct-1}\right) + \gamma_1 \bar{X}_{ct} + \gamma_2 Z_c + \mu_{jt} + \epsilon_{ijct}, \tag{3}$$

where

$$\mu_{jt} = \varphi_2 \widehat{VA}_{g(j)} + \eta_{jt}.$$

Here, Y_{ijct} represents grandstudents math scores in the 10th grade standardized within cohort. Equation (3) corresponds to a multilevel model, with a random effect at the teacher level (μ_{jt}). The coefficient of interest φ_2 captures the extended effect of grandteachers. Results are presented in column (2) of Table 2. Having a grandteacher with a VA 1 SD above the mean increases grandstudents' scores by 0.04 SD, slightly higher than the 0.03 SD implied by extrapolating VA across two generations. As we show in Section 7 below, this gap is consistent with attenuation from classical measurement error.

[7] We further confirm no association between grandteachers' VA and teachers' gender, socioeconomic background, 8th-grade GPA and attendance, and their school's average performance in the 8th grade (Appendix Table B7).

[8] The results are robust to controlling for teachers' GPA and attendance in 8th and 9th grades, characteristics of the schools they attended, and characteristics of the schools where they work (Appendix Table B8).

Table 1: Baseline VA estimates and absence of recruitment effects.

	2006 cohort	Index cohort				Math teachers
	10th-grade math scores (1)	12th-grade math scores (2)	Graduated university (3)	Graduated education (4)	High school math teacher (5)	School rank (6)
Grandteacher VA	0.20 [0.20; 0.20] (3568.49)	0.04 [0.03; 0.04] (9.48)	0.007 [0.003; 0.010] (3.99)	0.0004 [-0.0009; 0.0018] (0.62)	0.00005 [0.00001; 0.00010] (2.17)	1.19 [-0.15; 2.37] (1.74)
Mean of dep. var.	0	0	0.36	0.07	0.002	83.80
N students	138,014	364,542	364,542	364,542	364,542	643
N grandteachers	3,388	3,112	3,112	3,112	3,112	507

Note: 95% confidence intervals in square brackets; t-statistics in parentheses. Standard errors are clustered at the school level. Column (1) reports the estimated standard deviation of grandteachers' value-added, calculated as the square root of the covariance of residuals across classrooms taught by the same teacher. The standard error and confidence intervals in column (1) are estimated through 1,000 bootstrap repetitions. Columns (2)–(5) use the sample of 10th-grade students in 2004, 2005, 2007, and 2008 for whom information on their teachers' value-added and university admission scores is available, weighting observations by the inverse of the estimated probability of participating in the university admission process based on the model's control variables. The outcome in column (4) equals 1 if the student graduated from an education degree. The outcome in column (5) equals 1 if the student was working as a high school math teacher up to eight years after graduating from high school. In column (6), school ranking represents the percentile rank of the teacher within the cohort of students graduating from the same high school. All estimates in columns (2)–(6) control for student cohort, GPA and attendance in the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. At the school level, they control for school type, size, average 8th-grade math and reading SIMCE scores (third-degree polynomial), average household income and parental education, and time-averaged shares of low- and high-educated parents. Columns (3)–(5) report marginal effects evaluated at the mean of all other variables.

Table 2: Intergenerational transmission of VA and grandstudent effects.

		Grandstudents' scores		
	Teachers' VA	Math scores	Predicted math scores	Reading scores
	(1)	(2)	(3)	(4)
Grandteacher VA	0.14	0.04	-0.02	-0.00
	[0.01; 0.27]	[0.01; 0.08]	[-0.10; 0.06]	[-0.04; 0.04]
	(2.05)	(2.40)	(-0.56)	(-0.05)
Teacher's student cohort	✓	✓	✓	✓
VAM controls		✓		✓
N Students		6,032	6,032	6,032
N Teachers	212	212	212	212

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. Standard errors are clustered at the level of the school where teachers work. VAM controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and the school's type of administration, region, and socioeconomic status (SES) composition are also included. Predicted math scores are estimated based on the VAM controls.

Table 2 also reports two placebo tests: one using predicted math scores based on exogenous grandstudent characteristics, and another using grandstudents' reading scores. Predicted scores are generated from all controls in equation (1). Since these factors influence achievement but cannot be affected by 10th-grade teachers, any association with grandteacher VA would indicate bias. Column (3) shows no significant relationship, suggesting that effects are not driven by pre-existing student differences. Similarly, column (4) shows no association between grandteacher VA and students' 10th-grade reading scores, consistent with the expectation that math teachers have little direct impact on reading achievement (Aaronson et al., 2007; Chetty et al., 2014a).

Further evidence supports that our findings are not driven by teacher sorting across schools. First, our results are robust to excluding the 27 teachers who teach in the same school where they attended as students (Appendix Table B9). They also hold when estimating a VAM with school fixed effects, thereby addressing the potential sorting of teachers who graduated from higher-quality schools into higher-quality schools (Appendix E).

5.4. Continuous exposure effects

To further assess the plausibility of our findings, we exploit variation in the number of years future teachers are exposed to the same grandteacher during high school. Gradual dose–response relationships, in which effects increase with the duration of exposure, are a well-documented feature of social spillovers (e.g., Chetty and Hendren, 2018; Chetty et al., 2016; Lee, 2018). Because families are unlikely to select schools based on how many years a

Table 3: Intergenerational transmission of VA by years of exposure.

	Teachers' VA (1)	Grandstudents' scores (2)
Grandteacher VA	-0.04	-0.01
	[-0.24; 0.16]	[-0.06; 0.05]
	(-0.41)	(-0.19)
Additional year of exposure	-0.03	-0.02
	[-0.14; 0.08]	[-0.05; 0.01]
	(-0.58)	(-1.44)
Grandteacher VA × Additional year of exp.	0.14	0.04
	[0.01; 0.27]	[0.01; 0.07]
	(2.08)	(2.32)
Teacher's student cohort	✓	✓
VAM controls		✓
N Students		6,032
N Teachers	212	212

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. VAM controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and the school's type of administration, region, and socioeconomic status (SES) composition are also included. Additional year of exposure ranges from 0 (they were together only in the 10th grade) to 3 (together throughout high school). Test scores are standardised across all students within their respective cohorts (2013, 2015, and 2017).

student will be assigned to the same math teacher, and because actual exposure is shaped by factors largely outside their control, such as teacher turnover and grade reallocation, variation in overlap provides a sharper test of a causal grandteacher effect.

We re-estimate the intergenerational transmission parameter ρ, allowing it to vary with years of exposure, using the following specification:

$$VA_j = \alpha + \rho_1 \widehat{VA}_{g(j)} + \rho_2 \big(Years_j \times \widehat{VA}_{g(j)} \big) + \psi\, Years_j + \varsigma_j, \tag{4}$$

where $Years_j$ denotes the number of additional years teacher j was taught by the same grandteacher, ranging from 0 (together only in the 10th grade) to 3 (throughout high school), and $\widehat{VA}_{g(j)}$ is the empirical Bayes estimate of the grandteacher's VA. In this formulation, ρ_1 captures the transmission effect associated with one year of exposure, while ρ_2 measures how intergenerational transmission scales with additional years of contact. The total transmission effect after m years is therefore $\rho(m) = \rho_1 + (m-1)\rho_2$.[9] We also estimate an analogous regression for the indirect VA, φ_2.

[9]Appendix Table B10 reports the number and characteristics of teachers by years of exposure. The mean overlap in the sample is 2.4 years, SD = 1.

Table 3 shows little evidence of transmission after a single year of exposure, for either teacher VA (ρ) or grandstudent outcomes (φ_2). Effects then rise approximately linearly with additional exposure: each extra year increases teacher VA by about 0.14 SD and grandstudent achievement by roughly 0.04 SD. With four years of exposure, the cumulative effect reaches 0.36 SD for teacher VA and 0.10 SD for grandstudent achievement, as illustrated in Appendix Figure B4.[10]

These estimates suggest that sustained exposure to effective teaching during schooling might be as influential for later instructional quality as formal mentoring later in the career. For instance, the cumulative four-year effect is comparable in magnitude to the 0.12 SD gain estimated by Goldhaber et al. (2020) for early-career teachers with high-VA mentors and the 0.12 SD improvement documented by Papay et al. (2020) from pairing teachers with higher-skilled colleagues.

6. Scaling returns across generations

Is the spillover large enough to matter for policy? To understand the broader social value, we provide a simple back-of-the-envelope calculation quantifying the implied multiplier effect of teacher quality. Define I as the impact in standardized learning gains multiplied by the number of affected students. The direct impact of improving a teacher's VA by 1 SD can then be expressed as:

$$I_1 = \varphi_1 N, \tag{5}$$

where subscript $k = 1$ indexes the student generation, φ_1 corresponds to the effect of increasing a teacher's VA by 1 SD, and N is the number of students exposed to one teacher. The extended impact on the grandstudent generation can be written:

$$I_2 = (\rho\varphi_1)(N^2 P), \tag{6}$$

where ρ is the effect of VA on VA and P is the probability that an exposed student becomes a same-subject high school teacher. This reasoning can be extended to an arbitrary number of generations:

$$I_k = (\rho^{k-1}\varphi_1)(N^k P^{k-1}) = \varphi_1 N(\rho\, NP)^{k-1}. \tag{7}$$

If ($|\rho\, NP| < 1$), the infinite sum converges and the total impact across all generations is a geometric series:

$$I_{\text{total}} = \sum_{k=1}^{\infty} I_k = \varphi_1 N \sum_{k=1}^{\infty} (\rho\, NP)^{k-1} = \frac{\varphi_1 N}{1 - \rho\, NP}. \tag{8}$$

The multiplier relative to the direct effect on generation 1 is then:

$$\textit{Multiplier} = \frac{I_{\text{total}}}{I_1} = \frac{1}{1 - \rho\, NP} = \frac{1}{1 - \rho T}, \tag{9}$$

[10] Results are robust to controlling for characteristics of teachers' schools of origin that may affect overlap, including administration type, academic versus vocational tracks, and school size (Appendix Figure B5).

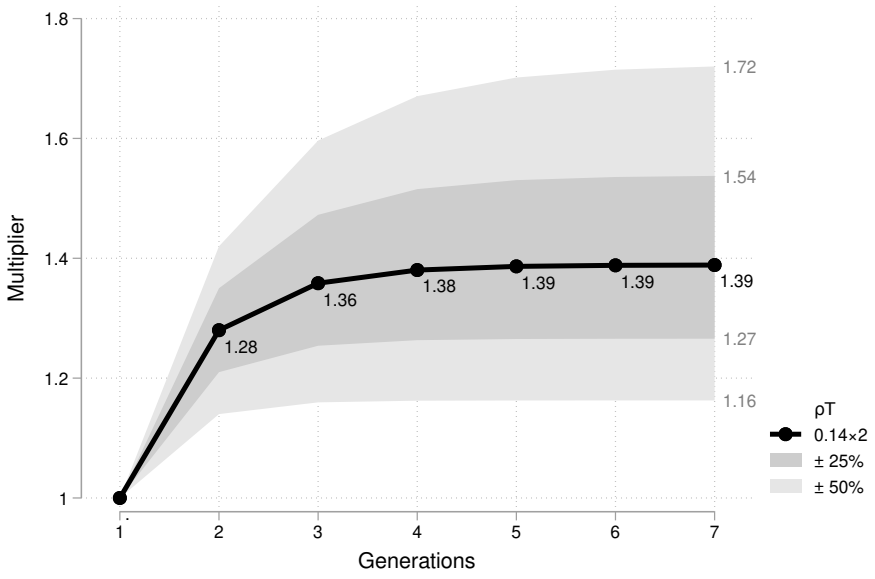


Figure 2: Social multiplier in teacher VA by generation.

Note: The figure plots the cumulative multiplier on teacher VA over successive generations of teachers. The middle line shows the estimate based on $\rho = 0.14$ and $T = 2$. The shaded areas vary the product ρT by $\pm 25\%$ and $\pm 50\%$, corresponding to alternative assumptions about intergenerational transmission or the number of future teachers taught by each grandteacher.

where $T = NP$ is the *number* of students exposed to one teacher who go on to become a same-subject high school teacher.

In other words, the total social multiplier can be recovered from two key quantities: the intergenerational transmission coefficient, ρ, and the number of future teachers a grandteacher instructs over a career, T. From our estimates, $\rho = 0.14$. To calculate T, note that each teacher instructs on average 2 classrooms of 25 students per year (Appendix Table A1). With an annual attrition rate of 5% in Chile, the expected career length is $1/0.05 = 20$ years (Centro de Investigación Avanzada en Educación, 2025). Thus, the average teacher instructs roughly $2 \times 25 \times 20 = 1000$ students during their career. Given that 0.2% of students eventually become same-subject high school teachers (Table 1, column 4), this implies $T = 1000 \times 0.002 = 2$. Plugging in ρ and T, we estimate a social multiplier $1/(1 - \rho T) = 1/(1 - 0.14 \times 2) = 1.39$, implying VA rises by 39% taking the generational spillover into account.

We visualize the implied social multiplier in Figure 2. The line reports our main estimate. Although the long-run multiplier reaches 1.39, it already equals 1.28 after one generation (the grandstudent cohort) and 1.36 after two generations (the great-grandstudent cohort). Hence, most of the cumulative gains materialize within a short horizon. The shaded areas illustrate the sensitivity of the multiplier to changes in the key term ρT by 25% and 50%, respectively. A decrease of 50% corresponds to a ρ of 0.07 or assuming that each teacher only instructs one future teacher. Gains are meaningful even in this conservative scenario, stabilizing around 1.16 in 2–3 generations.

7. Measurement error correction

Baseline estimates of VA and its intergenerational transmission are likely attenuated by measurement error. As shown in Appendix F, the correlation of math teacher VA with its two-year lag is 0.44 and forecast estimates imply a reliability of about 0.76. However, these estimates may confound error with genuine temporal variation. To provide stronger evidence on the potential role of measurement error, we leverage the gap between direct and indirect two-generation VA estimates.

Define $\hat{\varphi}_1$ as the estimated effect of teacher VA on students' test scores, $\hat{\rho}$ as the estimated effect of a grandteacher's VA on teacher VA, and $\hat{\varphi}_2$ as the estimated effect of grandteacher VA on grandstudent test scores. In a perfect-measurement scenario, the true effect on grandstudents would equal the product of the effects along the causal chain:

$$\hat{\varphi}_2 = \hat{\rho}\hat{\varphi}_1. \tag{10}$$

However, in practice, VA measures are noisy, and classical measurement error introduces attenuation bias. Let $\lambda \in [0, 1]$ denote the reliability of the VA measure, defined as the signal-to-total variance ratio (Fuller, 1987). Then, for an OLS regression of observed VA on outcomes, the estimated coefficients are attenuated toward zero by a factor of λ:

$$\hat{\varphi}_1 = \lambda\varphi_1, \quad \hat{\rho} = \lambda\rho, \quad \hat{\varphi}_2 = \lambda\varphi_2. \tag{11}$$

Now consider the estimated product along the chain:

$$\hat{\rho}\hat{\varphi}_1 = (\lambda\rho)(\lambda\varphi_1) = \lambda^2(\rho\varphi_1). \tag{12}$$

In contrast, the direct regression of grandstudent outcomes on observed grandteacher VA yields

$$\hat{\varphi}_2 = \lambda\varphi_2 = \lambda(\rho\varphi_1), \tag{13}$$

assuming the true effect $\varphi_2 = \rho\varphi_1$. Notice that

$$\hat{\varphi}_2/(\hat{\rho}\hat{\varphi}_1) = \frac{\lambda(\rho\varphi_1)}{\lambda^2(\rho\varphi_1)} = \frac{1}{\lambda} > 1, \tag{14}$$

Hence, the observed effect on grandstudents exceeds the product of the two stepwise effects because $\hat{\rho}\hat{\varphi}_1$ compounds the measurement error twice, while $\hat{\varphi}_2$ only suffers attenuation once. Using our empirical estimates ($\hat{\varphi}_1 = 0.20$, $\hat{\rho} = 0.14$, $\hat{\varphi}_2 = 0.04$), we can back out an implied reliability of

$$\lambda = \frac{\hat{\rho}\hat{\varphi}_1}{\hat{\varphi}_2} = \frac{0.20 \times 0.14}{0.04} = 0.70, \tag{15}$$

suggesting that roughly 30% of the variation in VA is noise. Adjusting for this implies a direct VA of $\hat{\varphi}_1/\lambda = 0.20/0.70 = 0.29$, a VA-on-VA of $\hat{\rho}/\lambda = 0.14/0.70 = 0.20$, and an indirect VA of $\hat{\varphi}_2/\lambda = 0.04/0.70 = 0.06$. Plugging these estimates into the long-term social multiplier would increase it to 1.66 (Appendix Figure B6).

8. Evidence for role modeling

How is teacher VA transmitted from grandteachers to future teachers? One channel could be the academic gains and improved pre-service qualifications associated with having more effective teachers, enhancing future teachers' subject-matter knowledge and higher education opportunities. However, aligning with Goldhaber et al. (2020), we find no evidence that pre-service qualifications account for the intergenerational transmission of VA. Our estimates are unchanged when controlling for teachers' 12th-grade math scores and the fixed effect of the university they graduated from (Appendix Table B11).

Prior research highlights role modeling as a key mechanism through which effective teachers improve the skills of peers and mentees (Papay et al., 2020). Yet, unlike colleagues, students observe only what happens inside the classroom: how teachers manage classes, deliver lessons, and engage students, and not the planning or pedagogical reasoning behind them (Lortie, 1975). If role modeling drives our results, grandteachers' influence should be strongest in the observable dimensions of instruction.

We use Chile's National Teacher Evaluation System (NTES) to test this. Between 2005 and 2015, both grandteachers and teachers were assessed under the same framework (see Appendix C). Teachers' portfolios combined a videotaped lesson rated on classroom management, lesson structure, and pedagogical interactions, and written materials assessing lesson planning, analysis, assessments, and reflection, scored on a four-point scale. We locate 32 math teachers as students in the 10th grade in 2002–2005 who are public high school teachers in 2012–2015, and their grandteachers' NTES scores from 2005–2008. Although small, this sample provides insight into which instructional practices are more likely to be transmitted across generations.

Appendix Figure C1 shows stronger correlations in video-based (classroom) dimensions than in written components: 0.33 versus 0.06. Appendix Table C3 disaggregates these effects: classroom management shows the highest association (0.42), followed by lesson structure (0.30), while written dimensions are near zero. Prior evidence has highlighted classroom management as an especially important teaching skill in the promotion of students' math achievements (Kane et al., 2011; Tyler et al., 2010). Overall, these findings suggest that exposure to effective grandteachers provides a lasting model for future teachers, through the instructional dimensions they directly observe as students.

9. Conclusion

Teachers spend thousands of hours as students observing instruction long before formal training begins. Because teacher effectiveness varies widely, some future teachers are repeatedly exposed to high-quality instruction while others are not. This study is the first to quantify how such differences in exposure shape the effectiveness of future teachers. Our findings show that exposure to a teacher 1 SD above the mean in VA raises a student's own VA by 0.14 SD if they later enter teaching, increasing their students' test scores by 0.04 SD. These spillovers imply that the long-run returns to teacher quality are about 39% larger than conventional estimates suggest. Our results strengthen the case for policies that identify, retain, and reward effective teachers, while highlighting the risk of persistent low-quality instruction:

students repeatedly taught by ineffective teachers are more likely to become less effective teachers themselves. Ensuring that all future teachers experience high-quality instruction during their schooling can therefore generate cumulative returns, as today's effective teachers become tomorrow's grandteachers.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25:95–135.

Altonji, Joseph G and Richard K Mansfield. 2018. "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects." *American Economic Review* 108:2902–2946.

Bacher-Hicks, Andrew, Mark J Chin, Thomas J Kane, and Douglas O Staiger. 2019. "An Experimental Evaluation of Three Teacher Quality Measures: Value-added, Classroom Observations, and Student Surveys." *Economics of Education Review* 73:101919.

Bacher-Hicks, Andrew and Cory Koedel. 2023. "Estimation and Interpretation of Teacher Value Added in Research Applications." In *Handbook of the Economics of Education*, edited by Eric A Hanushek, Stephen Machin, and Ludger Woessmann, volume 6, pp. 93–134. Elsevier.

Bacolod, Marigee P. 2007. "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality." *The Review of Economics and Statistics* 89:737–751.

Bietenbeck, Jan. 2025. "Do Motivated Classmates Matter for Educational Success?" *The Economic Journal* 135:36–58.

Björklund, Anders and Kjell G Salvanes. 2011. "Education and family background: Mechanisms and policies." In *Handbook of the Economics of Education*, volume 3, pp. 201–247. Elsevier.

Borowiecki, Karol Jan. 2022. "Good reverberations? Teacher influence in music composition since 1450." *Journal of Political Economy* 130:991–1090.

Braun, Sebastian Till and Jan Stuhler. 2018. "The transmission of inequality across multiple generations: testing recent theories with evidence from Germany." *The Economic Journal* 128:576–611.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review* 108:3377–3415.

Carrell, Scott E and Mark L Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics* 2:211–228.

Centro de Investigación Avanzada en Educación. 2025. "Abandono docente en el sistema educativo chileno 2015-2024." Informe del Observatorio de Formación Docente 12, Universidad de Chile, CIAE, Santiago, Chile.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-added Estimates." *American Economic Review* 104:2593–2632.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-added and Student Outcomes in Adulthood." *American Economic Review* 104:2633–2679.

Chetty, Raj and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133:1107–1162.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence From the Moving to Opportunity Experiment." *American Economic Review* 106:855–902.

Contreras, Dante, Paulina Sepúlveda, and Sebastián Bustos. 2010. "When Schools Are the Ones that Choose: The Effects of Screening in Chile." *Social Science Quarterly* 91:1349–1368.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg. 2017. "Peer Effects in the Workplace." *American Economic Review* 107:425–456.

Danielson, Charlotte. 1996. *Enhancing Professional Practice: A Framework for Teaching*. Association for Supervision and Curriculum Development, 1250 N. ERIC Number: ED403245.

De Grip, Andries and Jan Sauermann. 2012. "The Effects of Training on Own and Co-Worker Productivity: Evidence from a Field Experiment." *The Economic Journal* 122:376–399.

Feng, Shuaizhang, Jun Hyung Kim, and Zhe Yang. 2025. "Effects of Childhood Peers on Personality Traits." *Journal of Labor Economics* forthcoming.

Fuller, Wayne A. 1987. *Measurement Error Models*. New York: John Wiley & Sons.

García-Echalar, Andrés, Sebastián Poblete, and Tomás Rau. 2024. "Teacher Value-added and the Test Score Gender Gap." *Labour Economics* 89:102588.

Goldhaber, Dan, John Krieg, and Roddy Theobald. 2020. "Effective Like Me? Does Having a More Productive Mentor Improve the Productivity of Mentees?" *Labour Economics* 63:101792.

Goldhaber, Dan D, Pete Goldschmidt, and Fannie Tseng. 2013. "Teacher Value-added at the High-school Level: Different Models, Different Answers?" *Educational Evaluation and Policy Analysis* 35:220–236.

Golsteyn, Bart HH, Arjan Non, and Ulf Zölitz. 2021. "The Impact of Peer Personality on Academic Achievement." *Journal of Political Economy* 129:1052–1099.

Harris, Douglas N and Tim R Sass. 2011. "Teacher Training, Teacher Quality and Student Achievement." *Journal of Public Economics* 95:798–812.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. "The causal effect of parents' schooling on children's schooling: A comparison of estimation methods." *Journal of economic literature* 49:615–651.

Jackson, C Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-test Score Outcomes." *Journal of Political Economy* 126:2072–2107.

Jackson, C Kirabo and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1:85–108.

Kane, Thomas J and Douglas O Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." NBER Working Paper 14607, National Bureau of Economic Research.

Kane, Thomas J, Eric S Taylor, John H Tyler, and Amy L Wooten. 2011. "Identifying Effective Classroom Practices Using Student Achievement Data." *Journal of Human Resources* 46:587–613.

Koschnick, Julius. 2025. "Teacher-directed scientific change: The case of the English Scientific Revolution." Technical report, EHES Working Paper.

Kraft, Matthew A, David Blazar, and Dylan Hogan. 2018. "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-analysis of the Causal Evidence." *Review of Educational Research* 88:547–588.

Kraft, Matthew A, William H Marinell, and Darrick Shen-Wei Yee. 2016. "School Organizational Contexts, Teacher Turnover, and Student Achievement: Evidence From Panel Data." *American Educational Research Journal* 53:1411–1449.

Kraft, Matthew A, John P Papay, and Olivia L Chi. 2020. "Teacher Skill Development: Evidence From Performance Ratings by Principals." *Journal of Policy Analysis and Management* 39:315–347.

Lee, Se Woong. 2018. "Pulling Back the Curtain: Revealing the Cumulative Importance of High-Performing, Highly Qualified Teachers on Students' Educational Outcome." *Educational Evaluation and Policy Analysis* 40:359–381.

Lortie, Dan C. 1975. *Schoolteacher: A Sociological Study*. Chicago: University of Chicago Press.

Mansfield, Richard K. 2015. "Teacher Quality and Student Inequality." *Journal of Labor Economics* 33:751–788.

Mas, Alexandre and Enrico Moretti. 2009. "Peers at Work." *American Economic Review* 99:112–145.

Mizala, Alejandra and Hugo Ñopo. 2016. "Measuring the Relative Pay of School Teachers in Latin America 1997–2007." *International Journal of Educational Development* 47:20–32.

Ost, Ben. 2014. "How Do Teachers Improve? The Relative Importance of Specific and General Human Capital." *American Economic Journal: Applied Economics* 6:127–151.

Papay, John P and Matthew A Kraft. 2015. "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-term Career Improvement." *Journal of Public Economics* 130:105–119.

Papay, John P, Eric S Taylor, John H Tyler, and Mary E Laski. 2020. "Learning Job Skills From Colleagues at Work: Evidence From a Field Experiment Using Teacher Performance Data." *American Economic Journal: Economic Policy* 12:359–388.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73:417–458.

Shure, Nikki. 2021. "Non-cognitive Peer Effects in Secondary Education." *Labour Economics* 73:102074.

Treviño, Ernesto, Juan Pablo Valenzuela, and Cristóbal Villalobos. 2016. "Within-school Segregation In The Chilean School System: What Factors Explain It? How Efficient Is This Practice For Fostering Student Achievement And Equity?" *Learning and Individual Differences* 51:367–375.

Tyler, John H, Eric S Taylor, Thomas J Kane, and Amy L Wooten. 2010. "Value-added Modeling: A Review Using Student Performance Data to Identify Effective Classroom Practices." *American Economic Review* 100:256–260.

Valenzuela, Juan Pablo, Cristian Bellei, and Danae de los Ríos. 2014. "Socioeconomic School Segregation in a Market-oriented Educational System: The Case of Chile." *Journal of Education Policy* 29:217–241. Publisher: Routledge.

Teaching Teachers

Supplemental Material

A. Education in Chile

A1. Secondary education

To illustrate key features of the Chilean secondary education system, Panel A in Table A1 provides yearly statistics on the number of classrooms, students, and teachers in the 10th grade. The average school has two math teachers for the 10th grade, with students spread across four different classrooms, each containing an average of 33 students. The distribution of students across classrooms is generally not random. Previous studies in Chile suggest that nearly half of high schools segregate students across classrooms based on test scores and GPA (Treviño et al., 2016). We take this into account in our VA measures by controlling for classroom averages of prior SIMCE scores, GPA, and attendance.

Table A1: High schools, classrooms, and teachers in Chile.

	Mean	SD	p10	p50	p90
Panel A: the 10th grade by school					
N classrooms	4	2	1	3	7
N students per classroom	33	8	22	33	42
N math teachers	2	1	1	2	4
Panel B: math teachers by school					
N grades taught	3	1	1	3	4
N classrooms taught	5	2	2	5	8
N 10th-grade classrooms taught	2	1	1	2	4

Note: Own elaboration based on data from the Ministry of Education. Values correspond to yearly statistics for the period 2014–2017.

A2. High school teachers

Regarding high school teachers, most teach only one subject; for instance, math teachers exclusively teach math. Panel B in Table A1 summarizes the number of grades and classrooms (in total and in the 10th grade) taught by math teachers in Chilean high schools. Many teachers instruct more than two grades each year, teaching an average of five classrooms annually (two classrooms in the 10th grade).

Teachers in public schools in Chile are governed by the Teacher Statute, a national legal framework regulating working conditions, professional responsibilities, and dismissal procedures. Until 2016, there were no national academic requirements for entering the teaching profession beyond holding a teaching degree from a recognized higher education institution. Universities and non-universities were free to offer 4-year teacher education programs, though most in-service teachers graduated from universities (around 95% by 2015). Private and public schools recruited teachers independently, with decentralized recruitment processes across municipalities (administrators of all public schools until 2016). While public schools could only hire teachers without a teaching qualification under exceptional authorization by the Ministry of Education, private subsidized schools had greater discretion in employing professionals without formal teacher training, particularly in subjects with teacher shortages. The profession remained of relatively low status, with high school teachers earning 15–20% less than similarly educated professionals in 2007 (Mizala and Ñopo, 2016).

A3. Recent reforms

Significant reforms were introduced between 2015 and 2017 to enhance student learning and equity. The 2017 Public Education Law initiated the transfer of public school administration from 346 municipalities to 70 newly created Local Education Services (SLEP), intended to strengthen public education and reverse decreasing enrolment in this sector. In parallel, the 2016 Inclusion Law prohibited subsidized private schools from selecting students, charging co-payments, or operating for profit, which had previously exacerbated socioeconomic segregation of the system (Contreras et al., 2010; Valenzuela et al., 2014).

Reforms targeting the teaching profession were introduced through the 2016 Teacher Professional Development Law, which sought to address long-standing concerns about low salaries, inadequate preparation, and poor working conditions. It introduced minimum entry scores for teacher education programs, limited them to accredited universities, and created a structured national teaching career for all teachers in state-funded schools, linking pay and promotion to performance in a new evaluation system. It also introduced an induction and mentoring program for novice teachers and increased non-teaching time by reducing classroom hours from 75% to 65% of the workweek. This study focuses on teachers who entered the profession before the 2016 reform, under the previous preparation, recruitment, and evaluation systems.

A4. Admission to higher education

Students apply to higher education through a centralized admission system based on standardized exam scores (math, reading, history, and sciences) and high school GPA. About 75% of students participate in this system each year. The Chilean higher education system includes both universities (50–60% of higher education enrollment) and non-university institutions offering shorter, vocationally oriented programs.

B. Additional Tables and Figures

Table B1: Descriptive statistics by SIMCE cohort.

	2006 (1)	2013 (2)	2015 (3)	2017 (4)
10th-grade SIMCE scores				
Math	262 (65)	277 (64)	272 (63)	277 (62)
Reading	261 (52)	262 (54)	255 (52)	259 (51)
Socioeconomic background				
Household income (US$)	718 (814)	1150 (1173)	938 (891)	972 (869)
Parental education	12 (3)	12 (3)	13 (3)	13 (3)
Type of school administration				
Public	41%	32%	30%	29%
Private subsidized	50%	59%	63%	63%
Private paid	8%	8%	8%	8%
Prior performance				
8th-grade math	264 (49)	269 (48)	271 (47)	273 (48)
8th-grade reading	263 (50)	265 (48)	264 (48)	253 (48)
Attendance 8th grade	95 (6)	94 (5)	94 (5)	94 (6)
Attendance 9th grade	95 (6)	94 (5)	94 (6)	94 (6)
8th-grade GPA	5.81 (0.54)	5.72 (0.53)	5.69 (0.53)	5.74 (0.53)
9th-grade GPA	5.52 (0.54)	5.51 (0.55)	5.54 (0.55)	5.59 (0.56)
School transfer				
8th–9th school change	69%	59%	57%	51%
9th–10th school change	10%	8%	8%	8%
Female	53%	52%	52%	52%
N students	149,804	107,602	114,398	110,358
N math teachers	3,389	3,434	3,648	3,598
N schools	2,141	2,355	2,459	2,424

Note: Sample of students with full information in each SIMCE cohort. Standard deviations in parentheses. Attendance corresponds to the percentage of school days students attended their school during the school year. GPA ranges from 1.0 (minimum grade) to 7.0 (maximum); the minimum GPA required to approve the year is 4.0.

Table B2: Comparison of students in the 10th grade in 2004, 2005, 2007, and 2008 with available information on their teacher's VA from the 2006 SIMCE.

	All students (1)	With teacher's VA (2)
10th-grade cohort		
2004	24%	24%
2005	27%	28%
2007	24%	24%
2008	25%	24%
Type of school administration		
Public	42%	44%
Private subsidized	50%	49%
Private paid	8%	8%
School with vocational track	35%	36%
High school enrollment (mean)	212 (146)	211 (144)
School's share of low-educated parents	0.17 (0.16)	0.18 (0.16)
School's share of high-educated parents	0.15 (0.22)	0.15 (0.22)
School's average 8th-grade scores		
Math	0.19 (0.55)	0.19 (0.55)
Reading	0.19 (0.54)	0.19 (0.54)
School's average 10th-grade scores		
Math	0.10 (0.62)	0.10 (0.62)
Reading	0.10 (0.62)	0.10 (0.62)
N students	753,369	448,998

Note: Standard deviations in parentheses. The shares of low- and high-educated parents correspond to the across-time average share of parents with only primary education and a university degree, respectively. Schools' average scores in the 8th and 10th grades correspond to the SIMCE assessments in 2004 and 2006, respectively.

Table B3: Description of the index sample.

		Future math teachers	
	Index students (1)	High school teachers (2)	10th-grade teachers (3)
10th-grade cohort			
2004	23%	27%	28%
2005	26%	24%	28%
2007	26%	25%	26%
2008	25%	25%	17%
Female	54%	57%	58%
High school where they studied			
Type of administration			
Public	39%	42%	41%
Private subsidized	51%	51%	52%
Private paid	10%	7%	8%
School's share of low-educated parents	0.14 (0.15)	0.13 (0.14)	0.14 (0.15)
School's share of high-educated parents	0.18 (0.24)	0.19 (0.21)	0.17 (0.20)
8th-grade average score in math	0.30 (0.55)	0.48 (0.49)	0.43 (0.48)
8th-grade average score in reading	0.30 (0.55)	0.48 (0.49)	0.43 (0.48)
University admission exam scores			
Math	0.05 (0.96)	0.91 (0.43)	0.91 (0.42)
Reading	0.03 (0.95)	0.62 (0.47)	0.59 (0.52)
Higher education			
Graduated	58%	96%	100%
Type of higher education institution			
University	36%	94%	100%
Non-university	22%	2%	0%
Graduated from education-related degree	7%	83%	92%
High school where they work			
Type of administration			
Public		31%	28%
Private subsidized		65%	67%
Private paid		4%	5%
School's share of low-educated parents		0.15 (0.15)	0.16 (0.16)
School's share of high-educated parents		0.14 (0.17)	0.14 (0.18)
8th-grade average score in math		0.07 (0.45)	0.07 (0.42)
8th-grade average score in reading		0.06 (0.46)	0.07 (0.42)
N	364,542	643	212

Note: Index sample of students in the 10th grade in 2004, 2005, 2007, or 2008 for whom the VA of their 10th-grade math teachers can be measured in 2006 and who participated in the university admission process (76% of the index sample). The shares of low- and high-educated parents correspond to the across-time average share of parents with only primary education and a university degree, respectively. The 8th-grade average scores of the school where they studied come from the 2006 SIMCE assessment. The 8th-grade average scores of the school where they work come from the 2013, 2015, and 2017 SIMCE assessments.

Table B4: Effects on the probability of graduating from STEM and expected earnings.

	STEM degree (1)	Expected earnings (log) (2)
Grandteacher VA	0.006	0.012
	[0.003; 0.009]	[0.008; 0.016]
	(3.88)	(5.92)
Mean of dep. var.	0.12	13.7
N students	229,766	140,226
N grandteachers	3,102	3,091

Note: 95% confidence intervals in square brackets; t-statistics in parentheses. Standard errors are clustered at the school level. The sample includes 10th-grade students observed in 2004, 2005, 2007, and 2008 who graduated from higher education between 2012 and 2017.The dependent variable in column (2) is the log of the fourth-year mean earnings after graduation from the corresponding degree–institution pair, measured using 2014–2016 tax records. Earning data come from the System of Higher Education Information (www.mifuturo.cl) and are available for 1,000 different degree-institution pairs, covering 61% of higher education graduates. Observations in column (2) are weighted by the inverse of the estimated probability of having earnings information, based on the model's control variables. All estimates control for student cohort, GPA and attendance in 8th and 9th grade, indicators for school changes between 8th–9th and 9th–10th grade, and whether the student had the same teacher in 9th grade. At the classroom level, regressions additionally control for classroom averages of these variables. At the school level, they control for school type, size, average 8th-grade math and reading SIMCE scores (third-degree polynomial), average household income and parental education, and time-averaged shares of low- and high-educated parents.

Table B5: Effects on on-time graduation and participation in higher education admission.

	On-time graduation (1)	Participating in HE admission (2)
Grandteacher VA	0.005	0.002
	[0.003; 0.007]	[-0.002; 0.006]
	(5.68)	(1.13)
Mean of dep. var.	0.83	0.76
N students	478,454	478,454
N grandteachers	3,112	3,112

Note: 95% confidence intervals in square brackets; t-statistics in parentheses. Standard errors are clustered at the school level. The outcome in column (1) equals 1 if the student graduated on-time from high school. The outcome in column (2) equals 1 if the student participated in the university admission process within two years from graduating from high school. All estimates control for student cohort, GPA and attendance in 8th and 9th grade, indicators for school changes between 8th–9th and 9th–10th grade, and whether the student had the same teacher in 9th grade. At the classroom level, regressions additionally control for classroom averages of these variables. At the school level, they control for school type, size, average 8th-grade math and reading SIMCE scores (third-degree polynomial), average household income and parental education, and time-averaged shares of low- and high-educated parents. Estimates from a logit model, reporting marginal effects evaluated at the mean of all other variables.

Table B6: IPW-adjusted grandteachers' effects

	Teachers' VA		Grandstudents' scores	
	(1)	(2)	(3)	(4)
Grandteacher VA	0.14	0.24	0.04	0.06
	[0.01; 0.26]	[0.01; 0.46]	[0.01; 0.07]	[0.01; 0.11]
	(2.17)	(2.05)	(2.37)	(2.49)
Teacher's student cohort	✓	✓	✓	✓
VAM controls			✓	✓
Reweighting based on:				
Grandteacher VA	✓		✓	
Within-school rank		✓		✓
N Students			6,032	6,032
N Teachers	212	212	212	212

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. Standard errors are clustered at the level of the school where teachers work. Columns (1) and (3) adjust by the inverse of the probability of being in the sample based on grandteachers' VA. Columns (2) and (4) adjust by the inverse of the probability of being in the sample based on teachers' within-school rank (1 to 100) as students, which is determined by their average GPA from the 9th to the 12th grades. Weights are estimated using the full index sample. VAM controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and the school's type of administration, region, and socioeconomic status (SES) composition are also included.

Table B7: Effects on teachers' gender, socioeconomic background, and performance before the 10th grade.

	Female (1)	Parents with higher education (2)	8th-grade GPA (3)	8th-grade attendance (4)	School average 8th-grade score (5)
Grandteacher VA	-0.023	-0.009	0.031	0.062	-0.061
	[-0.096; 0.050]	[-0.070; 0.053]	[-0.028; 0.090]	[-0.498; 0.622]	[-0.130; 0.009]
	(-0.63)	(-0.28)	(1.05)	(0.22)	(-1.72)
Mean of dep. var.	0.58	0.22	6.34	96.44	0.43
N teachers	212	212	212	212	212
N grandteachers	195	195	195	195	195

Note: 95% confidence intervals in square brackets; t-statistics in parentheses. Standard errors are clustered at the school level. Sample of 10th-grade teachers who were in the 10th grade in 2004, 2005, 2007, and 2008 for whom information on their teachers' value-added is available. All estimates control for student cohort. The outcome in column (1) equals 1 if the teacher is female. The outcome in column (2) equals 1 if any of the teacher's parents completed higher education. This last variable is derived from the student questionnaire they had to complete during the 12th-grade university admission exam. The outcome in column (5) corresponds to the average 8th-grade SIMCE math scores of the school they attended. Columns (1) and (2) report marginal effects evaluated at the mean of all other variables.

Table B8: Impact of math grandteachers' value-added on math teachers' value-added.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grandteacher VA	0.14 [0.01; 0.27] (2.05)	0.13 [-0.01; 0.27] (1.88)	0.17 [0.04; 0.31] (2.52)	0.15 [0.02; 0.28] (2.29)	0.15 [0.01; 0.29] (2.08)	0.16 [0.03; 0.29] (2.39)	0.14 [0.01; 0.27] (2.14)	0.14 [0.02; 0.26] (2.22)
Student cohort	✓	✓	✓	✓	✓	✓	✓	✓
GPA and attendance in the 8th–9th grades		✓						
School where they studied:								
SES of the school			✓					
Average 8th-grade scores				✓				
Type of school, region, and size					✓			
School where they work:								
SES of the school						✓		
Average 8th-grade scores							✓	
Type of school, region, and size								✓
N grandteachers	212	212	212	212	212	212	212	212

Note: 95% confidence intervals in square brackets; t-statistics in parentheses.Standard errors are clustered at the school level. Variations in school SES variables correspond to a one percentage point change in the across-time average of these variables considering all 10th-grade students in 2006, 2013, 2015, and 2017.

Table B9: Intergenerational transmission of VA without teachers teaching where they studied.

	Teachers' VA (1)	Grandstudents' scores (2)
Grandteacher VA	0.16	0.04
	[0.02; 0.30]	[-0.00; 0.08]
	(2.22)	(1.88)
Teacher's student cohort	✓	✓
VAM controls		✓
N Students		5,168
N Teachers	185	185

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. Standard errors are clustered at the level of the school where teachers work. We excluded 27 teachers from the main sample who teach in the same school where they studied. VAM controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and the school's type of administration, region, and socioeconomic status (SES) composition are also included.

Table B10: Characteristics of the school of origin by years of exposure to the same grandteacher.

		Years of overlap			
	Total	1	2	3	4
Type of administration					
Public	41%	44%	48%	18%	44%
Private subsidized	52%	47%	44%	69%	54%
Private paid	8%	9%	8%	13%	2%
School with vocational track (%)	16%	12%	19%	18%	13%
High school enrolment (mean)	195	236	211	151	162
	(145)	(167)	(148)	(122)	(117)
School's share of low-educated parents	0.14	0.13	0.16	0.12	0.16
	(0.15)	(0.13)	(0.17)	(0.13)	(0.17)
School's share of high-educated parents	0.17	0.19	0.16	0.20	0.14
	(0.20)	(0.22)	(0.20)	(0.24)	(0.16)
8th-grade average score in math	0.43	0.45	0.42	0.45	0.43
	(0.48)	(0.48)	(0.51)	(0.45)	(0.49)
8th-grade average score in reading	0.43	0.45	0.42	0.44	0.43
	(0.48)	(0.48)	(0.51)	(0.45)	(0.49)
N teachers	212	57	64	39	52

Note: Own elaboration based on the sample of math teachers with information on their own math teachers' VA. Standard deviations in parentheses. High school enrollment corresponds to the number of students enrolled in grades 9th to 12th. The shares of low- and high-educated parents correspond to the across-time average share of parents with only primary education and a university degree, respectively.

Table B11: Potential mechanisms behind the intergenerational transmission of VA.

	Teachers' VA			Grandstudents' scores		
	(1)	(2)	(3)	(4)	(5)	(6)
Grandteacher VA	0.14 [0.01; 0.27] (2.05)	0.13 [0.00; 0.26] (2.00)	0.15 [0.01; 0.30] (2.09)	0.04 [0.01; 0.08] (2.40)	0.04 [0.01; 0.08] (2.35)	0.04 [0.01; 0.08] (2.24)
Teacher's student cohort	✓	✓	✓	✓	✓	✓
12th-grade math scores		✓			✓	
University FE			✓			✓
VAM controls				✓	✓	✓
N Students				6,032	6,032	6,032
N Teachers	212	212	212	212	212	212

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. Standard errors are clustered at the level of the school where teachers work. The 12th-grade math scores are standardized to have a mean of 0 and a variance of 1 across all test-takers for each year. University FE refer to the fixed effects associated with the universities from which the teachers graduated (there are 38 different universities in the sample). VAM controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and the school's type of administration, region, and socioeconomic status (SES) composition are also included. Predicted math scores are estimated based on the VAM controls.

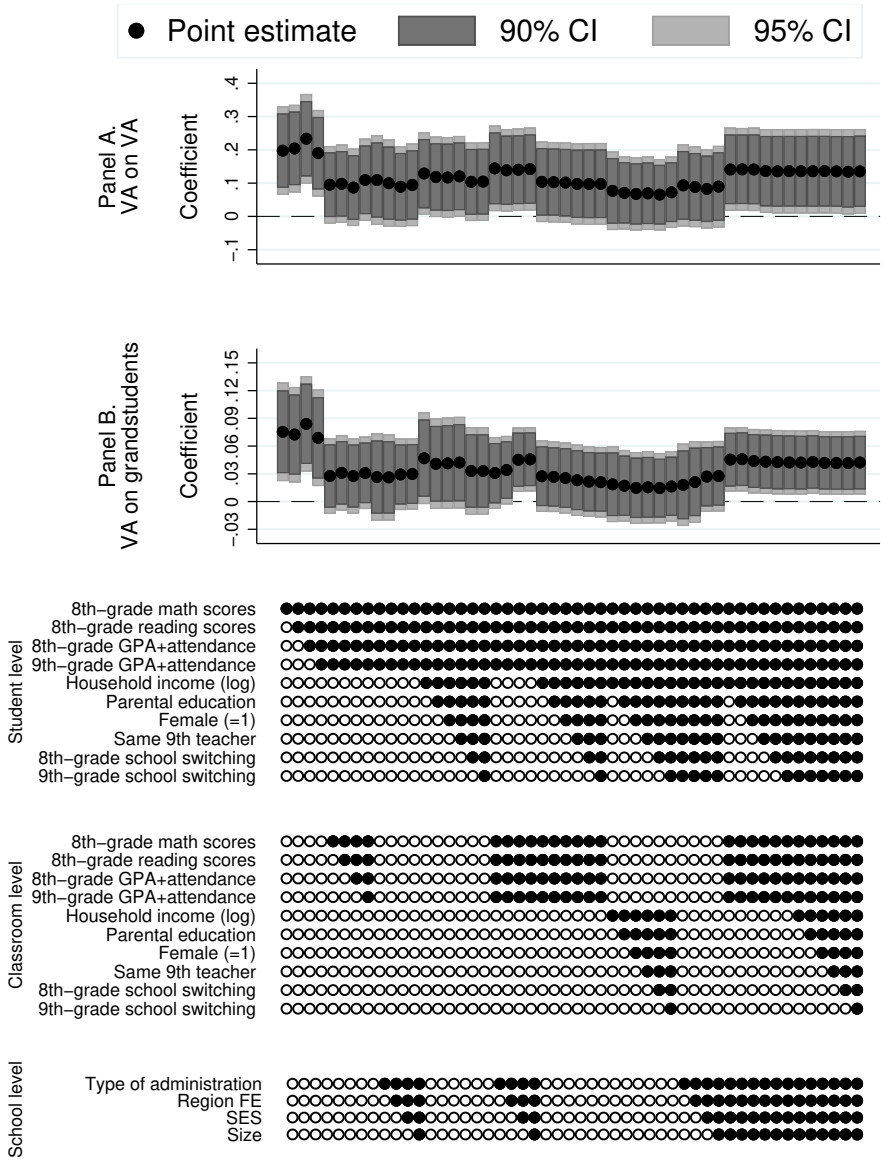


Figure B1: Specification curve analysis.

Note: Standard errors are clustered at the school level. The specification curve analyses shows how the estimated relationship between grandteacher and teacher VA, as well as gransteacher VA and grandstudent scores, change when varying the set of controls included in the VAM in equations (1) and (3). Black dots indicates that the variable enters the model as a control. 8th-grade math and reading scores are included as a third-degree polynomial. Classroom-level controls correspond to the classroom averages of these variables. School SES corresponds to the across-time average of the shares of high- and low-educated parents considering all 10th-grade students in 2006, 2013, 2015, and 2017. All estimates control for the teachers' student cohort.

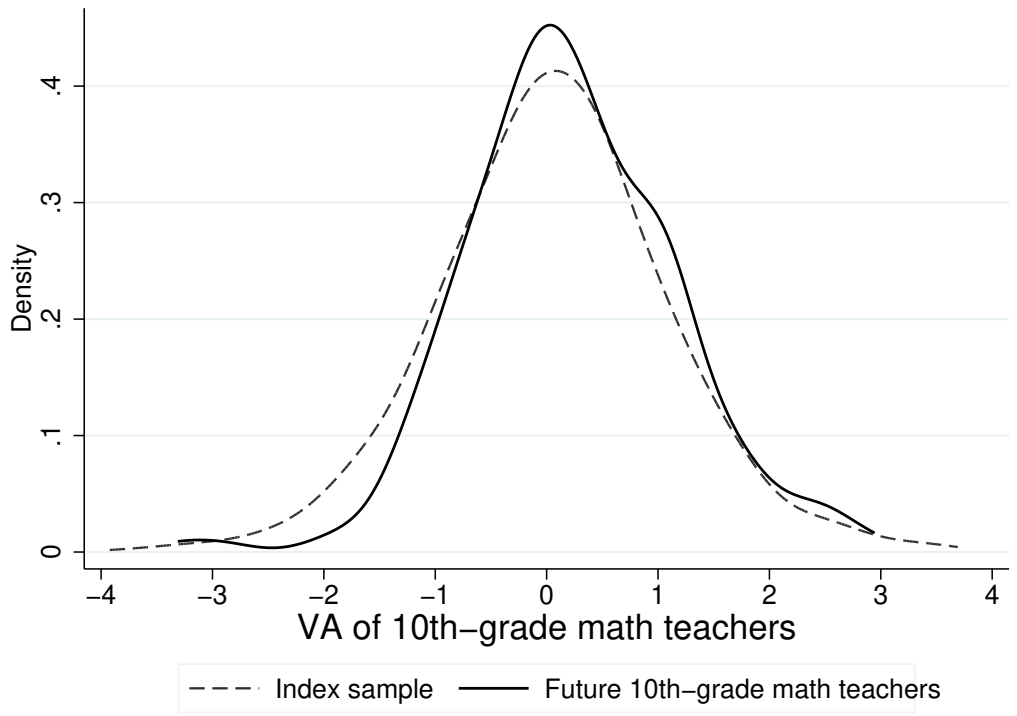
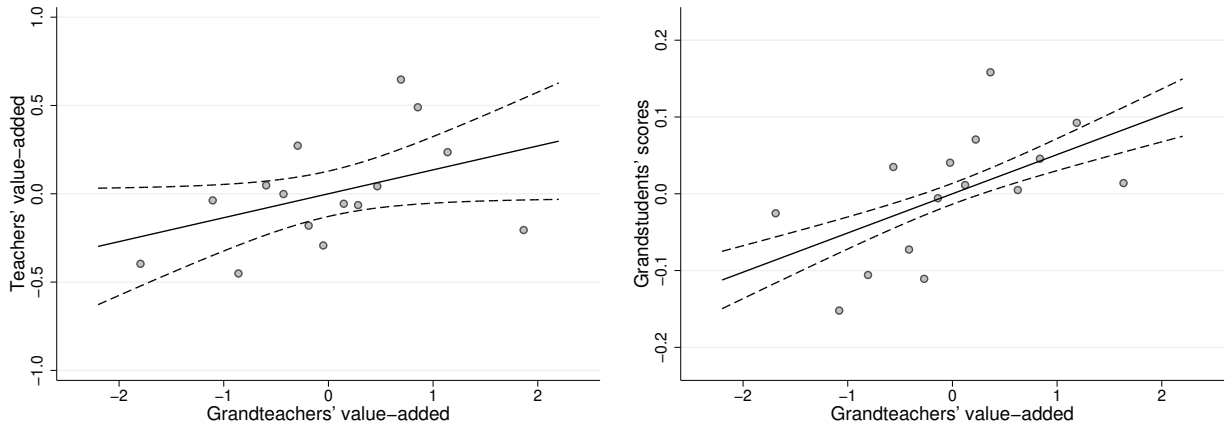


Figure B2: Distribution of VA to which students were exposed in the 10th grade.

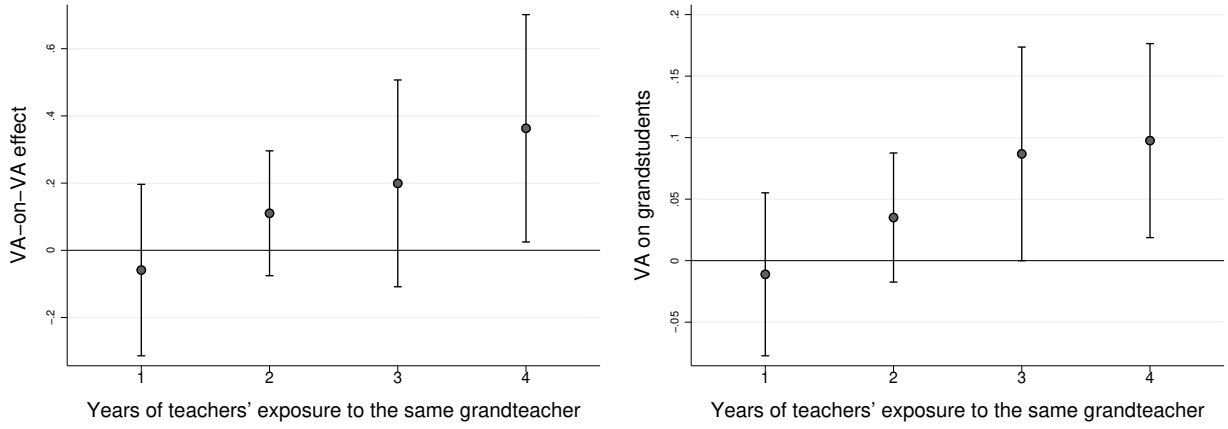
Note: The index sample includes 364,542 students, of whom 212 became 10th-grade math teachers 9–10 years later. The figure plots the kernel density estimation, implemented via the *kdensity* package in Stata. We select an optimal bandwitdth of 0.28 from the sample of future 10th-grade teachers.



(a) VA-on-VA (b) VA-on-grandstudents

Figure B3: Scatterplot of grandteachers' effects on teachers' VA and grandstudents' scores.

Note: Estimates based on 6,032 students taught by the 212 teachers in the sample. Observations are grouped into 15 equal-sized bins based on teachers' VA (grandstudents' scores), showing the mean of each bin. The fitted lines the full sample, not the binned averages. Panel (a) shows the residuals of grandteacher VA and teacher VA after controlling for the teacher's student cohort. Panel (b) shows the residuals of grandteacher VA and grandstudent test scores after controlling for all the variables included in equation (3) (grandstudents' prior test scores, socioeconomic background, classroom composition, and school characteristics).



(a) VA-on-VA (b) VA-on-grandstudents

Figure B4: Continuous exposure effects.

Note: Standard errors are clustered at the school level. The figure shows the impact of 1 SD increase in grandteacher VA by years of overlap between grandteacher and teacher. We interact grandteacher VA with years of overlaps treated as a categorical variable (from 1 to 4). Panel (a) shows effects of grandteacher VA on teacher VA and (b) of grandteacher VA on grandstudent test scores (b). Estimates based on 6,032 students taught by the 212 teachers in the sample. In panel (a), estimates control for the teacher's student cohort. Teachers' and grandteachers' VA are standardised across all teachers within evaluation year (2006, 2013, 2015, or 2017). In panel (b) estimates control for all the variables included in equation (3): grandstudents' prior test scores, socioeconomic background, classroom composition, and school characteristics. Test scores are standardised across all students within their respective cohorts (2013, 2015, and 2017).



(a) VA-on-VA

(b) VA-on-grandstudents

Figure B5: Continuous exposure effects with additional control variables.

Note: Standard errors are clustered at the school level. The figure shows the impact of 1 SD increase in grandteacher VA by years of overlap between grandteacher and teacher. We interact grandteacher VA with years of overlaps treated as a categorical variable (from 1 to 4). Panel (a) shows effects of grandteacher VA on teacher VA and (b) of grandteacher VA on grandstudent test scores (b). Estimates based on 6,032 students taught by the 212 teachers in the sample. In panel (a), estimates control for the teacher's student cohort and characteristics of the school they graduated from (type of school administration, whether it offers academic and/or vocational upper secondary education, and the number of students in the school). Teachers' and grandteachers' VA are standardised across all teachers within evaluation year (2006, 2013, 2015, or 2017). In panel (b) estimates additionally control for all the variables included in equation (3): grandstudents' prior test scores, socioeconomic background, classroom composition, and school characteristics. Test scores are standardised across all students within their respective cohorts (2013, 2015, and 2017).

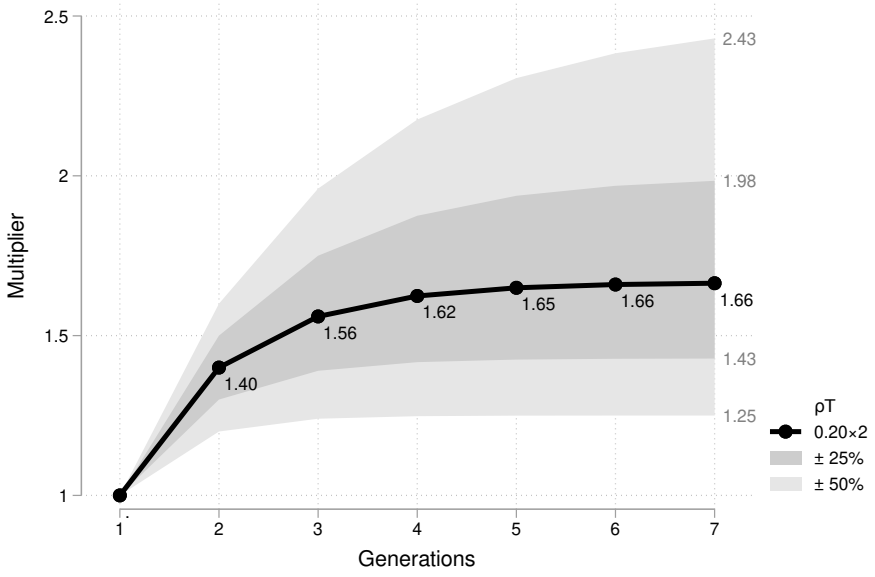


Figure B6: Social multiplier in teacher VA by generation, corrected for measurement error.

Note: The figure plots the cumulative multiplier on teacher VA over successive generations of teachers. The middle line shows the bias-adjusted estimate based on $\rho = 0.20$ and $T = 2$, in contrast to our main estimate of $\rho = 0.14$. The shaded areas vary the product ρT by $\pm 25\%$ and $\pm 50\%$, corresponding to alternative assumptions about intergenerational transmission or the number of future teachers taught by each grandteacher. Unlike Figure 2 in the main text, this version applies the measurement error correction developed in Section 7.

C. National Teacher Evaluation System

In 2004, the Ministry of Education introduced the Guidelines for Good Teaching, based on the framework of Danielson (1996), which set professional standards for teachers in Chile. Building on this framework, the National Teacher Evaluation System (NTES) operated from 2005 to 2015, assessing all public school teachers. The NTES combined four instruments: (1) a self-evaluation questionnaire, (2) a structured peer interview, (3) questionnaires completed by two school leaders, and (4) a portfolio. The portfolio was the central component and included (a) a full teaching unit of eight lessons, an assessment instrument used in the unit, and written reflections, and (b) a video recording of a lesson, filmed by an external team, with accompanying information on its goals. Each element was scored by trained scorers on a four-point rubric ranging from unsatisfactory (1) to outstanding (4). Final results were a weighted average: 60% portfolio, 20% peer interview, 10% supervisor assessments, and 10% self-evaluation.

The National Teacher Evaluation System (NTES) portfolio included (a) a full teaching unit of eight lessons, an assessment instrument used in the unit, and written reflections, and (b) a video recording of a lesson, filmed by an external team, with accompanying information on its goals. From these modules, the portfolio rated teachers in seven dimensions, each evaluated through multiple indicators and scored on a four-point scale ranging from 1 (unsatisfactory) to 4 (outstanding).

The classroom video module assesses three dimensions, which can be summarized as follows:

1. Classroom Management: teacher's ability to maintain student focus, manage group dynamics, and promote equitable participation.

2. Lesson Structure: coherence of class activities, the effective use of instructional time, and the teacher's ability to introduce and consolidate learning objectives.

3. Pedagogical Interactions: clarity of explanations, the quality of questions posed, and the teacher's ability to build on student contributions to deepen understanding.

The written module assesses four additional dimensions:

1. Organization of the Unit: the formulation and coherence of learning objectives, the pedagogical justification for activities, and the sequencing of lessons.

2. Analysis of Classroom Activities: teacher's ability to identify the strengths and weaknesses of the unit and design strategies to address students' learning difficulties.

3. Assessments: the adequacy of assessment items, their alignment with learning goals, and the justification for the evaluation instruments.

4. Reflection: how teachers use student assessment results to provide feedback, plan improvements to their teaching, and engage in broader reflection on their pedagogical practices.

Table C1 presents the correlations between teachers' VA scores (the one- and two-step VA measures) and their results on each component, as well as the overall evaluation score.[11] For math teachers, there is a low positive correlation (r between 0.15 and 0.21) between VA and overall evaluation results. Correlations with VA scores are greater for the video-based components of the portfolio (r between 0.12 and 0.22) and supervisor ratings (r between 0.19 and 0.26).

Table C1: Correlation coefficients between teacher VA and NTES scores

	Two-step VA (1)	One-step VA (2)
Total	0.15 (0.04)	0.21 (0.04)
Self-evaluation	-0.03 (0.04)	-0.02 (0.04)
Peer interview	0.01 (0.04)	0.06 (0.04)
Supervisor evaluation	0.19 (0.04)	0.26 (0.04)
Portfolio		
Video-taped lesson		
Classroom management	0.13 (0.04)	0.15 (0.04)
Lesson structure	0.12 (0.04)	0.13 (0.04)
Pedagogical interactions	0.22 (0.04)	0.22 (0.04)
Written part		
Lesson planning	0.03 (0.04)	0.04 (0.04)
Analysis of activities	0.03 (0.03)	0.03 (0.03)
Assessments	0.08 (0.04)	0.12 (0.04)
Reflections	0.11 (0.04)	0.13 (0.04)
N	609	

Note: Standard errors in parentheses. Full sample of math teachers in public schools with VA scores and NTES assessments in 2006, 2013, and 2015. All scores are normalized to have a mean of zero and unit variance across all teachers assessed by the NTES each year.

We construct a sample of 32 math teachers with available information on their portfolio scores and their grandteachers' portfolio scores. These teachers attended 10th grade in public schools between 2002 and 2005 and were later evaluated as teachers in the public sector between 2012 and 2015. Their grandteachers were assessed under the NTES between 2005 and 2008. Table C2 compares the distribution of scores for each of these dimensions across all public-school teachers and those in the 32-teacher sample. When examining the average scores and standard deviations of grandteachers in comparison to the broader teacher population between 2005 and 2008, the results indicate that their performance is quite sim-

[11] The models used for estimating the one- and two-step VA are described in sections E of the Appendix and 4 in the main text, respectively.

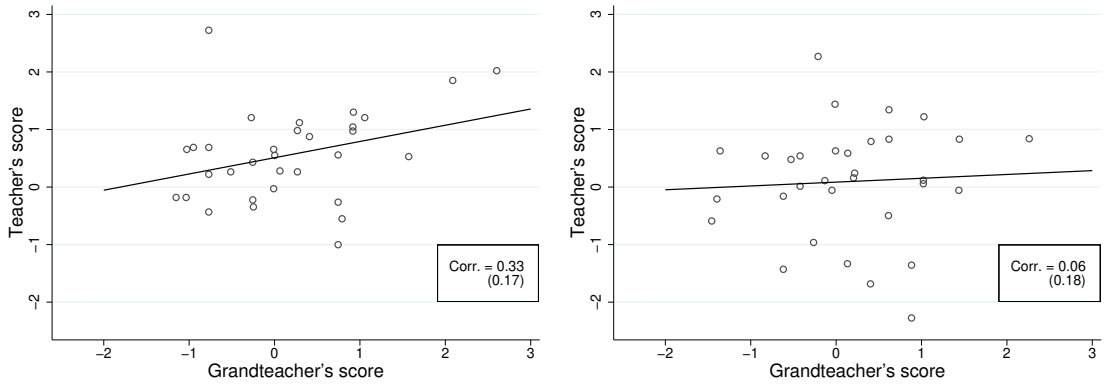
ilar. This suggests that new teachers had equal exposure to both low-rated and high-rated grandteachers.

Table C2: Average and standard deviation of NTES dimension scores: broader population and grandteacher–teacher sample

Dimension	2005–2008 period		2012–2015 period	
	Total (1)	Sample (Grandteachers) (2)	Total (3)	Sample (Teachers) (4)
Video-taped				
Classroom Management	2.6 (0.5)	2.7 (0.5)	2.8 (0.2)	2.8 (0.2)
Lesson structure	2.1 (0.5)	2.3 (0.5)	2.7 (0.4)	2.8 (0.3)
Pedagogical interactions	2.0 (0.5)	2.1 (0.5)	2.3 (0.5)	2.6 (0.4)
Written part				
Lesson planning	2.3 (0.5)	2.3 (0.5)	2.6 (0.6)	2.6 (0.6)
Analysis of activities	2.5 (0.6)	2.4 (0.6)	2.1 (0.6)	2.2 (0.6)
Assessment instruments	2.2 (0.8)	2.2 (0.9)	2.4 (0.7)	2.5 (0.6)
Reflection	1.9 (0.5)	2.0 (0.5)	2.1 (0.7)	2.2 (0.7)
N	1,403	29	1,638	32

Note: Standard deviation in parentheses. "Total" corresponds to all math teachers evaluated during the 2005–2008 and 2012–2015 periods. For teachers with more than one evaluation within each period, we consider their first evaluation.

Table C3 reports the correlations between grandteachers' and teachers' scores across the seven different dimensions. These scores have been standardized to mean zero and unit variance for each year of evaluation. Correlations are higher for those aspects of instruction that teachers directly experienced in the classroom as students, particularly in classroom management and lesson structure, which yielded correlation coefficients of 0.42 and 0.3, respectively. For the written dimensions (lesson planning, analysis, assessments, and reflection), the correlations are generally weaker, ranging from -0.18 to 0.15.



(a) Video-based scores (b) Written-based scores

Figure C1: Association between grandteacher's and teacher's NTES scores.

Note: Scores are standardized to mean zero and unit variance across all assessed teachers for each year of evaluation. The solid line shows the best linear fit. Correlation coefficients and their standard errors are reported as insets in each panel. The video-based scores average three dimensions: classroom management, lesson structure, and pedagogical interactions. The written-based scores average four dimensions: lesson planning, analysis of activities, assessments, and reflection.

Table C3: Correlation between grandteacher's and teacher's NTES scores by dimension

Dimension	Correlation
Video-taped lesson	
Classroom management	0.42 (0.17)
Lesson structure	0.30 (0.17)
Pedagogical interactions	0.15 (0.18)
Written part	
Lesson planning	0.11 (0.18)
Analysis of activities	0.06 (0.18)
Assessments	-0.17 (0.18)
Reflections	0.15 (0.18)
N	32

Note: Standard errors in parentheses. The table shows the correlation coefficient for each dimension between the grandteacher's score and the teacher's score. These scores have been standardized to have a mean of zero and unit variance across all assessed teachers for each year of evaluation.

D. Value-added estimation

We estimate teacher effects (μ_{jt}) using the empirical Bayes two-step approach described by Kane and Staiger (2008). In the first step, we estimate equation (1) via OLS and compute residuals ($\hat{v}_{ijct}$). In the second step, we average these residuals at the teacher-year level: $\overline{\hat{v}}_{jt}$. Because most high-school teachers teach more than one classroom (63% of teachers), we first average these residuals at the classroom level ($\overline{\hat{v}}_{jct}$) and then collapse the data at the teacher-year level considering precision-weighted averages, assigning more weights to classrooms with more precise estimation, i.e., more students (Chetty et al., 2014a). Defining n_{jct} as the number of students in classroom c in year t, we estimate the residual at the teacher-year as follows:

$$\overline{\hat{v}}_{jt} = \sum_{c} w_{jct} \overline{\hat{v}}_{jct}, \tag{D1}$$

where

$$w_{jct} = \frac{h_{jct}}{\sum_c h_{jct}}, \quad \text{and} \quad h_{jct} = \frac{1}{\hat{\sigma}^2_{\theta} + \frac{\hat{\sigma}^2_{\varepsilon}}{n_{jct}}}$$

We estimate $\hat{\sigma}^2_{\varepsilon}$ as Var $\left(\hat{v}^s_{ijct} - \overline{\hat{v}}_{jct}\right)$ and $\hat{\sigma}^2_{\theta}$ as Var $\left(\hat{v}^s_{ijct}\right) - \hat{\sigma}^2_{\varepsilon t} - \hat{\sigma}^2_{\mu t}$. The component $\hat{\sigma}^2_{\mu t}$, corresponding to the estimated variance of teacher VA, is computed using the covariance of average residuals across classrooms taught by the same teacher in a given year. We estimate this component by randomly sorting classrooms within each teacher-year cell and estimating the covariance between pairs of classroom-level residuals $\overline{v}_{jct}$, weighting each pair by the sum of students taught (Chetty et al., 2014a; Jackson, 2018). Table D1 shows the estimated standard deviation of teacher VA for each year, along with the correlation between classrooms taught by the same teacher. Math teachers show an SD of 0.20–0.21 throughout this period. The positive and moderately high correlation between classrooms (0.60) in columns (1) suggests consistency in teacher effects across classrooms.

Table D1: Correlation between classrooms and the standard deviation of VA.

Year	Correlation (1)	SD of VA (2)
2006	0.57	0.20
2013	0.61	0.21
2015	0.57	0.20
2017	0.64	0.21

Note: Columns (1) shows the correlation between the residuals of pairs of classrooms taught by the same teacher. The variance of teachers' VA is recovered from the covariance of average residuals across classrooms. The SD of teachers' VA in columns (2) is the square root of this covariance.

The final empirical Bayes estimates of each teacher's VA result from multiplying the weighted average of the teacher-year residuals $\overline{v}_{jt}$ by an estimate of its reliability:

$$\widehat{VA}_{jt} = \lambda_{jt}\bar{\widehat{v}}_{jt}, \qquad \text{where} \qquad \lambda_{jt} = \frac{\widehat{\sigma}^2_{\mu}}{\text{Var}\left(\bar{\widehat{v}}_{jt}\right)} = \frac{\widehat{\sigma}^2_{\mu}}{\widehat{\sigma}^2_{\mu} + \frac{1}{\sum_c h_{jct}}} \tag{D2}$$

This reliability ratio $\lambda_{jt} \in [0, 1]$ shrinks noisy estimates toward zero, with more shrinkage for teachers with fewer students.

E. One-step VA model

We estimate an alternative VAM that specifies teacher VA as standard fixed effects, also referred to as one-step VAM:

$$Y_{ijct} = \beta_0 f\left(Y_{it-1}\right) + \beta_1 X_{it} + \gamma_0 f\left(\bar{Y}_{ct-1}\right) + \gamma_1 \bar{X}_{ct} + \gamma_2 Z_{st} + \tau_j + v_{ijct} \tag{E1}$$

Here, τ_j is a teacher fixed effect estimated jointly with the other parameters in the model. Unlike our main two-step specification, here we assume τ_j to be time-invariant, which also allows us to estimate parameters on school-year covariates (Z_{st}). The time-invariant assumption means that teachers with multiple SIMCE assessments are assigned a single VA estimate summarizing their effects over years. To form the empirical Bayes estimates, we consider the reliability ratio to be equal to the ratio between the estimated true variance (the observed variance of $\hat{\tau}$ minus the mean of the square of the standard error estimates of $\hat{\tau}$) and the observed variance of $\hat{\tau}$ (Mansfield, 2015).

Columns (1) and (2) in Table E1 show the estimated VA-on-VA effect and the indirect effect of grandteachers' VA on grandstudents' scores. The indirect effect is obtained by estimating equation (3) using grandteachers' VA from the one-step model in equation (E1). Results mirror our main findings, with a VA-on-VA effect of 0.14 SD and an indirect effect of 0.04 SD.

As an additional check, we estimate a one-step VAM that includes school fixed effects:

$$Y_{ijct} = \beta_0 f\left(Y_{it-1}\right) + \beta_1 X_{it} + \gamma_0 f\left(\bar{Y}_{ct-1}\right) + \gamma_1 \bar{X}_{ct} + \gamma_2 Z_{st} + \tau_j + \delta_s + v_{ijct} \tag{E2}$$

where δ_s denotes school fixed effects and Z_{st} includes only time-varying school covariates (enrollment). Identification of τ_j and δ_s relies on teachers transferring across schools over time. As in equation E1, we must assume teachers effects are time-invariant to include school fixed effects. Again, this assumption hinders us from estimating the VA of grandteachers specific to 2006, around the time they taught the students who later became teachers. However, adding school fixed effects in the VAM allows us to address potential sorting of teachers with higher-VA grandteachers into higher-quality schools. Columns (3) and (4) in Table E1 show the results with these VA estimates, which remain highly consistent with our main results.

Table E1: Intergenerational transmission of VA and grandstudent effects (one-step VAM).

	VAM without school FE		VAM with school FE	
	Teachers' VA	Grandstudents' scores	Teachers' VA	Grandstudents' scores
	(1)	(2)	(3)	(4)
Grandteacher VA	0.14	0.04	0.17	0.04
	[0.02; 0.27]	[0.00; 0.07]	[0.02; 0.32]	[0.01; 0.08]
	(2.33)	(2.01)	(2.20)	(2.32)
Teacher's student cohort	✓	✓	✓	✓
VAM controls		✓		✓
N Students		6032		6032
N Teachers	212	212	212	212

Note: 95% confidence intervals in squared brackets and t statistics in parentheses. Standard errors are clustered at the level of the school where teachers work. VAM additionally controls include students' gender, socioeconomic background (log household income and parental education), a third-degree polynomial of 8th-grade scores, GPA and attendance in both the 8th and 9th grades, school changes between 8th–9th and 9th–10th grades, and whether the student had the same teacher in the 9th grade. Classroom averages of these variables and school enrollment are also included. The VAM in columns (1) and (2) also includes school's type of administration, region, and socioeconomic status (SES) composition.

Table F1: Temporal stability of teacher VA

Sample	Two-year correlation (r) (1)	Year-to-year correlation (sqrt(r)) (2)	N (3)
Panel A. Math			
2013–2015	0.42	0.65	1,827
2015–2017	0.47	0.69	1,841
Pooled sample	0.44	0.66	3,668
Panel B. Reading			
2013–2015	0.27	0.52	1,772
2015–2017	0.25	0.50	1,760
Pooled sample	0.26	0.51	3,532

Note: The 2013–2015 and 2015–2017 samples include teachers observed in t and $t+2$ within each period. The pooled sample combines all teachers within these samples. Year-to-year correlations are computed as the square root of the corresponding two-year correlation. Column (3) reports the number of teachers used to estimate each correlation.

F. Validation checks

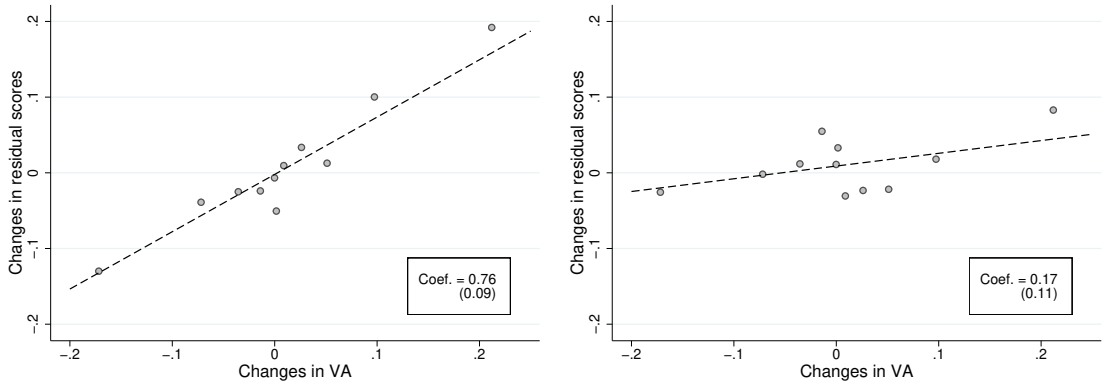
F1. Temporal stability of VA

Because our SIMCE cohorts are spaced at least two years apart (2013, 2015, and 2017), we assess the stability of teacher effects by correlating teachers' VA in year t with their VA two years earlier ($t-2$). Column (1) of Table F1 reports the two-year correlation coefficients for teachers observed in 2013–2015 and those observed in 2015–2017, as well as for the pooled sample combining both subsamples. The two-year correlation is 0.44, suggesting a year-to-year correlation of $\sqrt{0.44} = 0.66$, which lies near the upper bound of the 0.47–0.64 range reported by Bacher-Hicks and Koedel (2023). The results are nearly identical when examined separately for each period (0.42 for 2013–2015 and 0.47 for 2015–2017). In contrast, the persistence of reading VA is substantially lower, with a two-year correlation of 0.27, corresponding to a year-to-year correlation of 0.51.

F2. Estimates of forecast bias

Following Chetty et al. (2014a), we assess VA estimates' forecast bias using teacher switching as quasi-experiments. The core idea is that when a school experiences the departure and/or arrival of teachers, the resulting changes in the average VA of the school's teaching staff should predict variations in the average student scores, assuming teacher turnover is unrelated to other changes in school characteristics.

Let $\Delta\bar{\hat{v}}_{st}$ represent the change in the average student residual score of students in school s between t and $t-2$, and $\Delta\,\widehat{\overline{VA}}_{s\{t-2\}}$ denote the change in the school's student-weighted mean VA of teachers. This last variable is calculated considering only VA estimates from $t-2$, so variations are driven by teacher turnover rather than changes in VA (around 65%–70% of schools experienced teacher turnover or the arrival of a new teacher between 2013–2015 and



(a) Math scores (b) Reading scores

Figure F1: Forecast bias of math teachers' VA.

Note: 1,877 schools with information on their average teacher VA and residual scores (math and reading) in two years (2013–2015 or 2015–2017). Standard errors in parenthesis. The fitted lines, regression coefficients, and standard errors reported in this figure are estimated using the full school-level data, not the binned averages.

2015–2017). Assuming that teacher turnover is unrelated to shifts in other determinants of student scores, $\Delta\overline{\widehat{VA}}_{s\{t-2\}}$ should closely predict $\Delta\overline{\hat{v}}_{st}$. We estimate the following model:

$$\Delta\overline{\hat{v}}_{st} = a + b\Delta\overline{\widehat{VA}}_{s\{t-2\}} + \zeta_{st} \tag{F1}$$

The extent to which b differs from 1 reflects the level of forecast bias of the estimates. The binned scatterplot in Panel A of Figure F1 illustrates the results. These plots group the data into 10 equal-sized bins based on changes in school-level VA and plot the mean student residual score change in each bin. The figure shows that changes in the average VA of the school's teaching staff (based on estimates from $t-2$) closely predict variations in the average student scores between t and $t-2$. The two-year $\hat{b}$ equals 0.76, which suggests a forecast bias of 24%. This is larger than the 3% reported by Chetty et al. (2014a), but consistent with recent research on Chile (García-Echalar et al., 2024), and of a similar magnitude as our measurement error estimate in Section 7 in the main text.

To evaluate the validity of this test, we conduct a placebo test using test scores in the other subject, i.e., assessing how changes in the average math VA predict changes in student reading scores. An equally positive and significant relationship would suggest that the previous results are likely driven by changes in other school factors. Panel B in Figure F1 shows these results. The estimated $\hat{b}$ in the placebo test is significantly lower (0.17) and statistically indistinguishable from zero ($p = 0.14$).

F3. Teacher VA and experience

Figure F2 illustrates the relationship between teachers' VA and years of teaching experience using binned scatterplots. These plots group the data into 20 equal-sized bins based on years of experience, showing the mean teacher VA in each bin. The fitted lines are estimated

using a locally weighted scatterplot smoothing (LOWESS) procedure applied to the full teacher-level data rather than the binned means. This association shows a concave profile: VA increases significantly during the first five years of teaching, leveling off after 8-10 years, which aligns with international evidence (Ost, 2014; Papay and Kraft, 2015; Rivkin et al., 2005).



Figure F2: Association between years of teaching experience and VA.

Note: Own elaboration based on 6,151 math teachers with VA estimation in 2006, 2013, 2015, or 2017. Teacher VA is is standardized to have mean zero and unit variance based on all teachers with available measures. Information on years of experience comes from records of the Ministry of Education. The analysis considers only teachers with 20 or fewer years of teaching experience. Observations are grouped into 20 equal-sized bins based on years of experience, showing the mean teacher VA in each bin. The fitted lines reported in this figure are estimated through a locally weighted regression (*lowess* package in STATA) using the full teacher-level data, not the binned averages.

To complement this analysis, we estimate the following model:

$$VA_j = \pi_0 + \pi_1 Experience_j + \xi_j \tag{F2}$$

Here, $Experience_j$ represents a five-category variable: 0–1, 2–3, 4–5, 5–10, and more than 10 years of teaching experience. Table F2 shows the results, considering teachers with one or fewer years of experience as the reference group. Experience is more significant for math teachers, suggesting teachers reach peak VA between 5 and 10 years. It is important to note that these findings are purely descriptive and may be influenced by performance-driven attrition of teachers. For instance, lower-VA teachers may be more likely to leave the profession, while higher-VA teachers might have better opportunities to take on different roles within schools.

Table F2: Coefficients of years of teaching experience on teacher VA.

	(1)
Years of experience (base = 0 or 1 year)	
2–3 years	0.13
	(3.40)
	[1,602]
4–5 years	0.15
	(3.84)
	[1,346]
5–10 years	0.22
	(6.38)
	[2,220]
$>$ 10 years	0.17
	(4.97)
	[2,703]
N	9,305

Note: t-statistics in parentheses. The number of observations in each category are in square brackets. Full sample of 6,151 math teachers with 20 or fewer years of teaching experience and with VA measures in 2006, 2013, 2015, or 2017. Around 58% of teachers in this sample appear in the data more than once (VA measures for more than one year). Teacher VA is is standardized to have mean zero and unit variance based on all teachers with available measures. All estimations control for year fixed effects. Information on years of experience comes from records of the Ministry of Education.