

NBER WORKING PAPER SERIES

HUMAN CAPITAL DEPRECIATION AND RETURNS TO EXPERIENCE

Michael Dinerstein  
Rigissa Megalokonomou  
Constantine Yannelis

Working Paper 27925  
<http://www.nber.org/papers/w27925>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue  
Cambridge, MA 02138

October 2020, Revised September 2022

Prior versions of the paper were titled “Human Capital Depreciation.” The authors wish to thank Stéphane Bonhomme, Niels Gormsen, Caroline Hoxby, Victor Lavy, Magne Mogstad, Steve Rivkin, Jonah Rockoff, Seth Zimmerman, and seminar participants at NBER Labor Studies and Economics of Education Summer Institute, Columbia University, Princeton University, the University of Chicago, Athens University of Economics and Business, CESifo Area Conference on the Economics of Education, Public Choice Conference at Queensland University Technology, the Chicago Booth Household Finance Conference, the Rome Junior Conference on Applied Economics, the University of Adelaide, the University of Piraeus, the University of Toronto, University of Queensland, and Warwick University for helpful discussion and comments. We are grateful to Spyros Kypraios and Katerina Nikalexi for superb research assistance and to Panos Voulgaris, Veronica Blau, and Ilias Arvanitakis for helping to compile the data. Constantine Yannelis is grateful to the Becker Friedman Institute and Fama Miller Center for generous financial support. This work was approved by the Research Committee by the University of Queensland (clearance number: 2020000158 / IRB19-1614). The data repository is openicpsr-165721 (see Dinerstein et al. (2022)). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Michael Dinerstein, Rigissa Megalokonomou, and Constantine Yannelis. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Human Capital Depreciation and Returns to Experience  
Michael Dinerstein, Rigissa Megalokonomou, and Constantine Yannelis  
NBER Working Paper No. 27925  
October 2020, Revised September 2022  
JEL No. H52,I26,J24

**ABSTRACT**

Human capital can depreciate if skills are unused. But estimating human capital depreciation is challenging, as worker skills are difficult to measure and less productive workers are more likely to spend time in non-employment. We overcome these challenges with new administrative data on teachers' assignments and their students' outcomes, and quasi-random variation from the teacher assignment process in Greece. We find significant losses to output, as a one-year increase in time without formal employment lowers students' test scores by 0.05 standard deviations. Using a simple production model, we estimate a skill depreciation rate of 4.3% and experience returns of 6.8%.

Michael Dinerstein  
Department of Economics  
University of Chicago  
1126 East 59th Street  
Chicago, IL 60637  
and NBER  
mdinerstein@uchicago.edu

Constantine Yannelis  
Booth School of Business  
University of Chicago  
5807 S. Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
constantine.yannelis@chicagobooth.edu

Rigissa Megalokonomou  
Department of Economics  
University of Queensland  
University Drive, St Lucia  
Brisbane, QLD 4072  
Australia  
r.megalokonomou@uq.edu.au

Human capital – the knowledge and skills of workers – is a key factor driving economic growth in the aggregate and labor market outcomes at the individual level. Human capital can be developed at home within the family, through formal education, and through labor market experience (Becker, 1962, 1964). But if acquired skills are not actively used, they may depreciate over time. Thus, high rates of skill depreciation may amplify the costs to unemployment or labor force detachment by lowering a worker’s future productivity.

Indeed, several studies find evidence of structural non-employment duration dependence, where the probability of callbacks, probability of re-employment, and wages upon re-employment decline in the length of the non-employment spell (e.g., Jacobson, LaLonde and Sullivan (1993), Kroft, Lange and Notowidigdo (2013), Autor, Maestas, Mullen and Strand (2015)). Whether this duration dependence derives from skill depreciation or other explanations, such as stigma, changes to reservation wages (Schmieder, von Wachter and Bender, 2016), changes to match quality (Neal, 1995), or cohort effects (Oreopoulos, Von Wachter and Heisz, 2012) is important for optimal policy design. For instance, if unemployment costs are driven by human capital depreciation, then policies that maintain part-time or temporary employment or provide structured activities that require using job-related skills might be particularly effective. Such policies could include part-time working subsidies (e.g., the German Kurzarbeit scheme), public works programs, or even banning non-compete clauses. Furthermore, high degrees of skill depreciation imply particularly large costs from increases in aggregate non-employment through a depletion in aggregate levels of human capital. Yet, despite their policy importance, well-identified empirical estimates of skill depreciation rates remain elusive.

This paper studies how human capital changes with time spent without formal employment. To do so, we compile a new dataset on teachers in Greece and their assignments and exploit an institutional feature that quasi-randomly assigns time spent in formal employment. We find that an additional year without formal employment, with the associated forgone experience, leads to a 0.05 student standard deviation ( $\sigma$ ) decline in a teacher’s students’ test scores. Using aggregated data covering the entire country, we estimate that a district’s student test scores fall by  $0.07\sigma$  per-class if the teachers’ mean time without formal employment increases by a year. We further decompose the effects into human capital depreciation and returns to experience using a simple production model and find a skill depreciation rate of 4.3% and experience

returns of 6.8% per year, for inexperienced teachers.

There are two challenges in measuring the changes in human capital due to non-employment. First, in many contexts worker productivity or ability is difficult to measure, especially over time (Bertrand and Schoar, 2003; Zivin and Neidell, 2012). Second, even if productivity can be measured, unproductive workers are less likely to receive job offers, and hence are likely to spend more time not working. This may generate a negative correlation between time spent without employment and productivity, even if time without employment does not directly affect productivity.

We overcome these challenges by studying teachers in Greece, using administrative data from the entire country. We address the measurement issue by focusing on employees for whom we have a direct measure of output. Following a large literature (e.g., Rockoff (2004); Chetty, Friedman and Rockoff (2014)), we infer a teacher's productivity based on her students' test scores, while we avoid the strong identification assumptions common in the value-added literature. We use students' results on the Panhellenic Examinations, national exams that all Greek high school students take in grade 12, and which are the primary determinants of university admission. We also include university enrollment outcomes.

We address the second empirical challenge by exploiting the unique system of teacher assignments in Greece. Individuals who graduate in good standing with an education degree are guaranteed a public sector teaching position. In nearly all years, however, there are not enough positions immediately available, and thus graduates enter long waitlists that determine assignments. By law, all university graduates are assigned waitlist rankings in order of their date of degree conferral. Because small differences in degree conferral dates are driven by heterogeneous course schedules, timing of oral defenses, and bureaucratic delays, we argue that within a degree conferral month-year, remaining variation is exogenous. This quasi-randomness in waitlist position can translate to considerable variation in how long similar teachers wait for assignments. Because some teachers may work in the informal sector while waiting, the waitlist variation generates exogenous differences in time without *formal* employment.

In addition to the data on national test scores, we compile novel data from the Greek Ministry of Education on the universe of Greek deputy teacher waitlist rankings and assignments between 2004 and 2011. The assignments designate the school district the teacher is assigned

to for the following year. We supplement this with hand-collected data from 22 high schools that includes student-teacher assignments for each course and test scores by subject for all high school grades.

Our main specification relates a teacher's students' test scores to the accumulated number of years the teacher spent without formal employment. To address the concern that years without formal employment may be correlated with a teacher's "potential" productivity,<sup>1</sup> we control for the month-year in which the teacher earned her degree and further instrument for years without formal employment with initial waitlist rank. Our estimates indicate significant loss in productivity from not working formally: a  $0.052\sigma$  decline in test scores per year. Losses are similar in the first and second semester exams, and continue to post-secondary outcomes. Students with a teacher who had an additional year without formal experience are 1.8 percentage points (2.7%) less likely to be admitted to a university.

To extend the analysis beyond these 22 schools, we employ a second specification where we aggregate to the school district level and estimate the effect of the labor force's average time spent without formal employment on test scores. These estimates are robust to within-district sorting of students to teachers. Our estimates are very similar to the estimates from the teacher-level specification: a combined depreciation and forgone experience effect of  $0.068\sigma$  per class, or 1.5 percentage points in university admissions probability.

Our identification strategy relies on the initial waitlist position, conditional on degree month-year, being orthogonal to (1) teacher potential productivity and (2) unobserved ability of a teacher's assigned students. We assess the validity of our identification strategy in several ways. First, we examine attrition, which we view as the most obvious threat. We find no evidence of attrition related to initial waitlist position, conditional on degree month-year. Second, we find no evidence that teachers' initial waitlist position, conditional on degree month-year, correlates with the teacher's university achievement. Third, we find no statistical relationship between the mean (conditional) waitlist rank of the teachers assigned to a district and district characteristics like unemployment rates or class size. Finally, as a placebo check we repeat our analysis using teachers who are assigned via the same system but teach in subjects that do not

---

<sup>1</sup>We refer to "potential" productivity as the teacher's productivity had she been continuously employed as a public school teacher. We distinguish this from realized productivity, which may depend on how long a teacher has been without formal employment.

appear on the national tests. When we estimate the effect of time without formal employment for these teachers on students' national test scores (in other subjects), we find no statistical relationship.

The main estimates capture the combined effect of skill depreciation, independent of employment, plus forgone skill appreciation that would have accrued with experience. We decompose the effects of these channels by comparing teachers with the same levels of prior experience but who have exogenously waited different amounts of time for their assignments. Using our district-level model, we estimate that, controlling for experience, an extra year without formal employment lowers test scores by  $0.044\sigma$  per class. We specify a simple production model of student test scores and teacher human capital that relates our causal estimates to experience returns and depreciation rates. Our estimates imply a skill depreciation rate of 4.3% and experience returns of 6.8% per year. Our individual-level estimates are noisier and driven primarily by more experienced teachers. For these teachers, we estimate a skill depreciation rate of 17.2% and experience returns of 4.4% per year.<sup>2</sup>

This paper delineates a specific mechanism that generates structural unemployment duration dependence. An extensive literature has documented the duration dependence of non-employment for callback rates (Kroft, Lange and Notowidigdo (2013), Farber, Silverman and von Wachter (2017)), wages upon re-employment (Jacobson, LaLonde and Sullivan (1993), Card, Chetty and Weber (2007), Centeno and Novo (2012), Schmieder, von Wachter and Bender (2016)), and re-employment (Autor, Maestas, Mullen and Strand, 2015). By focusing on a profession where output is observable, our paper distinguishes changes in productivity as a cause of duration dependence from other explanations, primarily stigma. Further, because our variation in non-employment is concentrated among early-career workers, our results imply that skill depreciation may be an important factor behind the negative effects of graduating during a recession (Kahn, 2010; Oreopoulos et al., 2012). Several papers have moved beyond wages to estimate effects on skills or output. Edin and Gustavsson (2008) estimate the effect of unemployment duration on skill measures, using a panel data fixed effects approach. Wiswall (2013) and Benhenda (2017) look at teachers' career interruptions and absences, respectively, and their effects on output. Our contribution is to combine output measures with

---

<sup>2</sup>We describe the human capital depreciation as skill depreciation rather than obsolescence because the teachers in our context teach under the same curriculum and technologies throughout the sample period.

quasi-random variation in non-employment that is robust to time-varying shocks.

The paper also contributes to a literature on human capital and productivity by estimating a key parameter used in many structural models in macroeconomics, labor, and finance, many of which study duration dependence (e.g., Alvarez, Borovičková and Shimer (2016)). Comparing estimates of depreciation across models is complicated as the form of depreciation is often specific to the model. Keane and Wolpin (2001) estimate a large amount of depreciation (31%) for white collar workers changing industries, which is similar to that in Imai and Keane (2004). Blundell, Dias, Meghir and Shaw (2016) estimate human capital depreciation rates of 6-8%. Macroeconomic models incorporating human capital depreciation tend to use very different parameterizations, including Ljungqvist and Sargent (1998) (20% chance of losing skills), Kehoe, Midrigan and Pastorino (2019) (1.4% during a non-employment spell), Jarosch (2015) (15% depreciation rate), and Manuelli and Seshadri (2014) (21% depreciation rate). We offer a depreciation rate estimate that leverages quasi-random variation in time not working.

Our work also contributes to a literature in public finance and labor economics, by estimating a parameter relevant to evaluating the effects of many policies. A large literature on optimal unemployment insurance (Baily, 1978; Chetty, 2006; Shimer and Werning, 2006) derives optimal benefit levels by trading off consumption smoothing benefits with changes to labor supply. If human capital depreciates when individuals are not working, this effect can affect optimal benefit levels. Non-compete clauses may also have harmful effects in the presence of human capital depreciation (Marx, Strumsky and Fleming, 2009). Countries such as Germany incentivize employers to reduce hours rather than lay off workers. Giroud and Mueller (2017) discuss how these policies led to lower unemployment during the Great Recession, with substitution in hours worked. These policies may have positive effects on human capital accumulation if they prevent deterioration. The importance of policies preventing the deterioration of human capital due to workers being out of the labor force is particularly relevant at the time of writing. The COVID-19 pandemic and recession has led to historically high unemployment and furlough rates in many countries. Our estimates suggest that the resulting decline in human capital may have aggregate effects.

Finally, in the most direct sense this paper joins a literature on the returns to experience

(e.g., Angrist (1990); Altonji and Williams (1992)), especially for teachers. A number of papers have estimated high experience returns early in a teacher’s career that then flatten (Rockoff, 2004; Rice, 2013), while Wiswall (2013) and Papay and Kraft (2015) estimate positive returns even at high levels of experience. Herrmann and Rockoff (2012) estimate the effects of teacher absence on students, mainly driven by the productivity of the substitute teacher. Jackson, Rockoff and Staiger (2014) provide a review of the broader literature on teacher productivity.<sup>3</sup> This paper estimates the effect of not working on productivity and decomposes the effect into within-worker experience returns and depreciation rates. We rely on quasi-random variation that avoids the common assumption that teachers select into experience samples based only on time-invariant factors.

The remainder of this paper is organized as follows. Section I discusses the institutional details that form the basis for our empirical strategy. Section II describes the data used in our analysis. Section III introduces our empirical strategy. Section IV presents the student-level estimates and then Section V extends the analysis to the larger sample and conducts a district-level analysis. Section VI decomposes the effects into skill depreciation and forgone returns to experience. Section VII concludes and discusses avenues for future research.

## **I Institutional Details**

The education system in Greece appoints teachers to either permanent or temporary positions. Our analysis focuses on the temporary teachers, formally called substitute or deputy teachers, who are employed hourly or full-time on a contract basis for up to ten months. Schools may request a deputy teacher when there is a shortage of teaching staff. This occurs through retirements among permanent teaching staff, teachers taking long-term unpaid leaves (maternity, pregnancy, post-graduate studies, serious illness, or temporary moves abroad), or unexpected demand shocks. For example, if a school’s enrollment increases more than expected, the school may request a deputy teacher to cover the additional classes. Most temporary teacher assign-

---

<sup>3</sup>Other recent work has focused on measurement (Hanushek and Rivkin, 2010), information and evaluation (Rockoff, Staiger, Kane and Taylor, 2012; Taylor and Tyler, 2012), instructor characteristics (Hoffmann and Oreopoulos, 2009), match quality (Jackson, 2013), peer effects (Jackson and Bruegmann, 2009; Opper, 2019), and performance pay (Lavy, 2002, 2009).



ments occur in September or October, but some of these events require mid-year assignments. Over the last two decades, the fraction of teachers in temporary assignments has grown considerably such that now 15-20% of the teacher workforce are on temporary contracts (OECD, 2018). Online Appendix A provides more background about the rise in temporary assignments and the entire assignment process.

University graduates with a degree in education prior to 2011 were entitled to a teaching position in a Greek public school. Graduates, however, have to wait until a position opens in their academic subject. The day the prospective teacher receives her university degree, she is added to the end of a subject-specific waitlist consisting only of teachers yet to receive any assignment. We thus refer to it as the “inexperienced” waitlist. If two teachers in the same subject have identical degree conferral dates, they are ordered by university grade-point-average. Unlike other higher education systems where a university’s graduates receive their degrees on the same day, in Greece degree conferral occurs once the pivotal course’s grade is entered. The pivotal course is often a student-teaching assignment or involves a written thesis or oral defense. Heterogeneous course schedules, grading congestion, and bureaucratic delays lead students to finish their degrees on different days, and this generates considerable variation in when students graduate, as seen in the top row of Figure 1. The left panel is a histogram of the month of the year in which teachers earn their degrees while the right panel is a histogram of the day of the month. Some months and days produce more degrees than others, but there is still extensive variation across and within month.

Once a teacher rises to the top of the list, she is offered a position the next time a school requests a temporary teacher. Teachers toward the back of the waitlist thus wait longer for their first assignment, which we show for the 2003 mathematics list in Appendix Figure A.1, Panel A. The figure plots the mean year of first assignment by waitlist position, with teachers toward the front of the list typically receiving assignments in 2005 while individuals with high waitlist numbers obtaining positions several years later. Waitlist position does not perfectly predict timing of assignment because teachers with special circumstances – having at least three children, having a special needs child, having military experience, or reading Braille – may receive faster assignments.

During the first assignment, which lasts up to ten months, the teacher accrues credits based

on the assignment's length and location. Upon completing the assignment, the teacher does not have the option to remain at the school but instead returns to the waitlist system. This time, the teacher enters a second subject-specific list (the "experienced" list) for teachers with prior assignments. This list is ordered by a teacher's accumulated number of credits (having more credits is better). The teacher then waits to rise to the front of the "experienced" list to receive her second teaching assignment. We show the histogram of wait times and the joint distribution of first and second assignment years in Appendix Figure A.2 for the 1995 degree cohort. While wait times are long for the first assignment, teachers tend to receive a second assignment quickly. The teacher accrues more credits during her second assignment, after which she returns to the "experienced" list, likely with a better position due to the additional credits. After each subsequent assignment, the teacher returns to the same "experienced" list. Eventually the teacher may hope to receive a permanent position; but due to a hiring freeze during our sample period, no teachers actually received permanent assignments.

While they wait for an assignment, teachers may find alternate ways to generate income. Importantly, these activities may not include taking a full-time job, which would remove a teacher's public school teaching eligibility. Greece's active informal employment sector means that some teachers may still take full-time jobs that are hidden from the government. In Appendix Table A.1 we present data from an online survey, which we describe further in Section II, of 200 current and former teachers. Just over half (54%) of the surveyed teachers used their time waiting to offer private lessons, 39% worked in non-education positions (e.g., as a restaurant waiter), and 33% got more formal education. A smaller number of teachers used their wait to start a family or on other activities. We will thus interpret our estimates as capturing the potential loss of skills relevant to a worker's desired profession (teaching) when the worker is unable to work formally in her desired profession. This, notably, does not rule out the acquisition of skills relevant for other professions nor does it preclude all activities that might use the same skills that matter for formal full-time employment. If the prospective teacher takes full-time formal employment or notifies the government that she no longer wants to be considered for temporary public school teaching positions, then she will no longer appear on the lists. We will explore attrition in more detail in Section III as certain forms of selective attrition would violate our identification assumptions.

This institutional setting has several features that might produce high skill depreciation. First, because nearly all full-time teaching positions are in the formal sector, there are few available ways to spend one's waiting time that approximate the desired job. Second, teachers face limited financial or professional incentives to perform well once assigned. Student test scores or other outcomes do not factor into any form of teacher evaluation. Neither financial compensation nor future assignments – both where and when – depends on performance. Thus, the incentives to insure against or take investments to counteract depreciation are limited, and the available tools are directly restricted.

The nature of work experience also stands out in this setting. On-the-job learning may involve general and firm-specific skills. Here, because each temporary position lasts less than a year and reassignment to the same school is exceedingly rare, the effect of experience on output may be limited as school-specific skills generate no future return. Furthermore, the short-term nature of the national assignment process means that each year of experience likely involves a costly relocation to a different part of the country. Finally, unlike many settings with time spent not working, here there is no need to search for a job, at least in the desired profession. While there is some uncertainty about when a position will be available, the teacher's actions do not affect it.

## **II Data**

We compiled novel national data on teacher waitlists, teacher assignments, and student test scores. We supplement these data sets with comprehensive data from 22 high schools that includes student-teacher linkages for each course taken and a survey of current and former teachers. We provide more details about the data set construction in Online Appendix B.

### **II.A Teacher Waitlists**

The Ministry of Education compiles teacher waitlists centrally and maintains an online archive in which it posts some waitlists from prior years. We tracked down the archives and constructed the waitlists for deputy and hourly high school teachers from 2003 – 2011 (Greece Ministry of Education and Religious Affairs, 2003-2012). Each year includes separate lists for each

teaching subject and for fresh graduates versus experienced teachers. A teacher may appear on both the deputy and hourly lists, corresponding to her experience level, in the same year. The waitlists include each teacher's position on the list plus any characteristic or outcome that determines the waitlist order. Importantly, this includes when a teacher's university degree was conferred and teaching experience accrued in each year. We restrict our sample to teachers who earned their university degrees prior to 2006, as our identification strategy, described in Section III, will rely on cohorts where at least some members reach a second assignment in our sample period.

Panel A of Table 1 displays summary statistics for the waitlist data, for teachers in subjects that appear on the national examinations. Under 40% of the waitlist teachers have yet to accumulate experience in the public Greek system, and prior experience in private schools or other EU countries is rare. Teachers of Greek and History comprise the majority of the list, with the rest are split between math and statistics, physics and biology, economics, and computer science. We will later compare teachers in the same subject with degree conferral dates in the same month-year. These cohorts are fairly large, with an average size of 62 teachers.

In the bottom row of Figure 1 we show the full distribution of total (left) or consecutive (right) years without formal employment for our sample of teachers that appear on the waitlists between 2003 and 2011. Most teachers wait at least two years, with waits of 3-4 years being common. While subsequent assignments beyond the first tend to occur more quickly, the distribution of consecutive years without formal employment is similar to the distribution of total years, as most teachers on the waitlists are awaiting their first assignment.

## **II.B Micro School Data**

To measure teacher productivity, we link teachers to their students' outcomes using micro data from 22 public high schools (Greece High School Archives, 2001-2011). We obtained this data by visiting these schools in person, requesting all of their records, and digitizing them. These schools are distributed throughout the country and cover a diverse set of areas (see Appendix Figure A.3 for the schools' locations).

The data set's key features are student and teacher course schedules for all high school grades that allow us to link a student to a specific teacher for each course and subject-specific

exam scores. Teachers cover an average of 2.17 subjects and teach 2.94 classes per subject (Panel B of Table 1). While the sample is limited to 22 schools, the more precise match between teacher and her students' outcomes will form the basis for our student-level empirical analysis. The initial data for the schools spans 2001 to 2011, with variation across schools in which years are available.

This micro dataset's additional outcomes allow us to test for treatment effects on subjects and grades not part of the national exam. We also observe grade point averages and semester-specific test scores. Finally, we merged the sample to university admissions outcomes, which allows us to extend our analysis to teacher effects on consequential, non-test outcomes. In our sample, 64% of students are admitted to a university and 62% to a university with an academic focus, with considerable dispersion in the selectivity of university attended.

## **II.C Teacher Assignment Data**

Because our micro school data only includes 22 schools, for a more representative analysis we use national data on assignments and outcomes. The Ministry of Education and the school districts collect information on teachers' temporary full-time assignments to high schools around the country. They publicly announce these lists to inform teachers of their assignments and for transparency. These assignments come from both teachers' waitlists – the fresh graduates waitlist and the experienced teachers' waitlists – and are usually announced in September or October based on schools' needs.

We obtained the assignment lists from the online archives of the Ministry of Education and school district authorities (Greece Ministry of Education and Religious Affairs, 2004-2011a). These lists contain an assigned teacher's name, teaching categorization based on the subjects that they teach, and the assigned school's district. We obtained these lists from 2004 – 2011, with an average of 2,491 high school teachers assigned per year. Unless a district has only a single high school, we do not know the exact school within the district that the teacher was assigned to. But districts are fairly small; the 608 districts in our sample have an average of 2.3 high schools.

We present summary statistics on the assigned teachers' subjects in Panel C of Table 1. The distribution of full-time assignments is similar to the distribution of teachers on the waitlists,

with math and statistics and physics and biology having somewhat larger relative assignment shares while economics and computer sciences have lower relative assignment shares. Different regions of the country rely more on the assignment of temporary teachers. In Appendix Figure A.4 we show how the number of assignments varies geographically and note that the islands are particularly dependent on temporary teachers.

## **II.D Test Score Data**

We also obtained student-level data (Greece Ministry of Education and Religious Affairs, 2003-2011b) from the Ministry of Education with test scores on the Panhellenic Examinations, national exams that all Greek high school students take in twelfth grade. Our data span 2004 – 2011 and include each student’s total score and school attended. These exams are the most important determinant for university admission; given these high stakes, the exams are graded by external markers. The exams cover the core subjects (Mathematics, Greek Language, History, Biology, Physics), and thus we will restrict our analysis to teachers in these subjects unless otherwise noted. Students also take exams in other subjects depending on whether they have chosen the Classics, Science, or Exact Science track. We convert the test scores to student standard deviation units by normalizing scores to have mean 0 and standard deviation 1 in each year.

## **II.E Teacher Survey**

Finally, we supplement the micro and national data with an online survey, of 200 current and former teachers, we conducted from December, 2019 – January, 2020 (Dinerstein et al., 2019-2020). The responses yield information on teachers’ awareness and expectations of the assignment system, non-teaching activities while waiting for an assignment or once assigned, and investments made to maintain skills. We provide recruitment details and the survey text, in Greek and English, in Online Appendix C.

### III Empirical Model and Strategy

We seek to estimate the effect of a year out of the formal labor force on teacher output, as measured by student outcomes. To fix ideas, consider the hypothetical example, illustrated below, of two teachers who earned their degrees in July, 2004. While otherwise identical, teacher 1 had her degree conferred earlier in the month than teacher 2, so teacher 1 started with a better waitlist position. This led to an immediate assignment in fall 2004 for teacher 1, while teacher 2 remained unassigned. Thus, at the end of the 2004-2005 school year, teacher 1 had accumulated 0 years without formal employment and 1 year of formal experience while teacher 2 had accumulated 1 year without formal employment and 0 years of formal experience. For the following school year, both teachers received assignments, such that at the end of the 2005-2006 school year, teacher 1 had accumulated 0 years without formal employment and 2 years of formal experience while teacher 2 had accumulated 1 year without formal employment and 1 year of formal experience.

	2004-05	2005-06
Teacher 1	Graduates in July and Assigned (Years Not Emp, Experience) = (0,1)	Assigned (0,2)
Teacher 2	Graduates in July but Not Assigned (Years Not Emp, Experience) = (1,0)	Assigned (1,1)

Our target parameter is the causal effect on some output measure of a year without formal employment instead of a year with formal employment. Identifying this effect involves comparing outcomes for teachers in the same school year, but with exogenous variation in the fraction of years since degree conferral without formal employment. Specifically, we can compare teachers 1 and 2 in the same academic year – 2005-2006 – where the difference is that teacher 1 has one fewer year without formal employment and one more year of formal experience.<sup>4</sup> The difference in outcomes is a relevant input for certain policymaking problems. In this section, we further characterize this estimand with a simple production model. The

---

<sup>4</sup>As our assignments data cover full-time assignments, we will compare full-time teachers with different levels of prior experience or years waiting. We can measure prior experience as either full-time-only experience or inclusive of part-time experience. We discuss the robustness of our estimates to the experience measure at the end of Section IV.C.

model serves several purposes: (a) it motivates the functional forms we use in our estimating equations; (b) it clarifies the relationship between the estimand and human capital technological parameters; and (c) it highlights the potential threats to identification. We then lay out our empirical strategy to obtain consistent estimates.

### III.A Production Model

Let  $i$  index students,  $j$  index teachers, and  $t$  index academic years. We start with a general model of student achievement and teacher human capital:

$$(1) \quad y_{it} = f(H_{jt}, \varepsilon_{it})$$

$$(2) \quad H_{jt} - H_{j,t-1} = g(H_{j,t-1}, e_{j,t-1})$$

where  $y_{it}$  is student  $i$ 's output (e.g., test score) in year  $t$ ,  $H_{jt}$  is teacher  $j$ 's stock of teaching-relevant human capital in year  $t$ ,  $e_{j,t-1}$  is whether  $j$  accrued formal teaching experience in  $t-1$ , and  $\varepsilon_{it}$  captures student heterogeneity and student-year shocks. We label  $f()$  as the student educational production function and  $g()$  as the teacher human capital production function.

We first parameterize the teacher human capital production function. We follow Blinder and Weiss (1976), Rosen (1976), and a generalized version of Ben-Porath (1967) in assuming that  $g()$  is homogeneous of degree one in the prior year's human capital:<sup>5</sup>

$$(3) \quad H_{jt} - H_{j,t-1} = (\gamma e_{j,t-1} - \delta_\tau) H_{j,t-1}$$

where  $\gamma$  is the return to experience and  $\delta_\tau$  is the human capital depreciation rate.<sup>6</sup> We let the depreciation rate differ between periods prior to any formal employment ( $\tau = 0$ ) and pe-

---

<sup>5</sup>These papers focus on optimal human capital investments to maximize discounted lifetime earnings and their production functions characterize how human capital varies with the fraction of time spent working. Here we focus only on the technology and treat each year's work experience as binary.

<sup>6</sup>The literature sometimes refers to the depreciation rate as the change in human capital only when an individual is not working, or the change associated with a disruption like switching industries. We could redefine  $\delta_\tau$  as being multiplied by non-employment and our model would be the same except the multiplicative change in human capital while working would be  $\gamma$  instead of  $\gamma - \delta_\tau$ .



riods after a teacher's first formal employment spell ( $\tau = 1$ ) because depreciation rates may vary depending on whether human capital includes prior on-the-job experience (Kahn, 2010; Oreopoulos et al., 2012). Given this production function, we can derive human capital as a function of teacher  $j$ 's human capital at the time of graduating from university ( $H_{j0}$ ), the number of years since graduating from university ( $N_{jt}$ ), the number of years since graduating from university before first formal employment ( $V_{jt}$ ), and the number of prior years of experience  $E_{jt} = \sum_{s=0}^{t-1} e_{j,s}$ :

$$(4) \quad H_{jt} = (1 - \delta_0)^{V_{jt}} (1 + \gamma - \delta_1)^{E_{jt}} (1 - \delta_1)^{N_{jt} - E_{jt} - V_{jt}} H_{j0}.$$

The literature provides less guidance on parameterizing student achievement as a function of teacher human capital, which is typically an unobserved input. But we follow the teacher value-added literature in assuming that a teacher's impact on student outcomes is additively separable from other determinants. Specifically, we adopt a log-linear specification, which is equivalent to Cobb-Douglas if the outcome is expressed in logs:

$$(5) \quad y_{it} = \ln(H_{jt}) + \varepsilon_{it}.$$

As human capital is not measured directly, we will make inference about a teacher's human capital based on her students' output. Thus, we insert Equation 4 into Equation 5 to express student output as a function of the determinants of human capital:

$$(6) \quad y_{it} = \ln(1 - \delta_0)V_{jt} + \ln(1 + \gamma - \delta_1)E_{jt} + \ln(1 - \delta_1)(N_{jt} - E_{jt} - V_{jt}) + \ln(H_{j0}) + \varepsilon_{it}.$$

If we rearrange terms to express student output in terms of number of years since graduating, number of years not working before first employment, and total number of years not working ( $U_{jt} \equiv N_{jt} - E_{jt}$ ), we have:

$$(7) \quad y_{it} = (\ln(1 - \delta_1) - \ln(1 + \gamma - \delta_1))U_{jt} + (\ln(1 - \delta_0) - \ln(1 - \delta_1))V_{jt} + \ln(1 + \gamma - \delta_1)N_{jt} + \ln(H_{j0}) + \varepsilon_{it}.$$

This model, with student outcomes as a linear function of  $U_{jt}$ ,  $V_{jt}$ , and  $N_{jt}$ , will form the basis

for our estimating equations, and in Section III.B we discuss how we identify the coefficient on  $U_{jt}$ . As the model clarifies, this target parameter combines experience returns ( $\gamma$ ) and depreciation ( $\delta_\tau$ ). We implicitly measure human capital in terms of student output. The combined effect is the relevant parameter for the effects of policies that shift employment status, as any additional year spent not working automatically includes forgone experience. But in Section VI we will decompose the total effect into depreciation and forgone experience channels.

Before proceeding, we make several observations about the parameter interpretations and functional forms. The return to experience,  $\gamma$ , captures any effect of formal experience on teaching-relevant human capital. This may include increased curricular knowledge, human capital investments complementary to being formally employed, improved time management skills, and changes in motivation or effort from working. The effect of formal experience is relative to the actions taken while not formally working, as characterized in Section I.

Similarly, the depreciation rate,  $\delta_\tau$ , captures any changes in teaching-relevant human capital independent from whether the teacher is formally employed. This may include forgetting material learned in formal education, (constant) age effects, and changes in motivation from being part of a complicated assignment system. The composite coefficient on  $U_{jt}$  is relevant for characterizing the current assignment process and for considering policies that shift employment status but not the experience returns or depreciation rates. But because some of these components might change depending on policy, might involve costly actions, or might have spillover effects outside of education, we will comment on the extent to which the estimates seem to be driven by different factors. Finally,  $\varepsilon_{it}$  may capture student heterogeneity and student-year shocks as well as any transient shocks to teacher productivity independent of experience.

The linear relationship we will estimate between student outcomes and teacher experience is the subject of debate in the economics of education literature. Some consensus has formed that experience returns are high for the first years of a teacher's career and then become flatter (Rockoff, 2004; Rice, 2013). We are hesitant to rely on this literature to guide our functional form choices for two reasons. First, most of the literature employs within-teacher estimators, ruling out selection into experience levels based on time-varying factors, and estimates a combined experience and age effect. Beyond the difference in estimands (whether the age effect is

included), our analysis will use exogenous variation in experience that is robust to selection in which teachers reach different experience levels based on productivity shocks.<sup>7</sup> Second, without exogenous disruptions to the accrual of experience, papers in the literature need to impose restrictions to solve the age-period-cohort problem, usually by imposing a flat profile for teachers above a certain experience level. Wiswall (2013) and Papay and Kraft (2015) show that reasonable variations of these restrictions, including Wiswall (2013)’s preferred specification, can generate a linear profile.

Despite these critiques, we will impose linearity instead of tracing out a non-parametric experience profile. We thus address the linearity assumption in two ways. First, the papers that find a non-linear relationship between experience and student outcomes typically find that the relationship is close to linear at low levels of experience. As we describe below, our variation will be largely driven by teachers early in their careers – just 1% and 19% of our district and individual estimation samples, respectively, have greater than 4 years of experience.<sup>8</sup> Our variation thus lies in the experience interval where the literature has found linear returns. Second, we will provide robustness checks to our main results that allow student output to vary with the log of experience and that vary the curvature in the student outcome.

### III.B Empirical Model and Identification

We now turn to our empirical implementation and how we will identify the model’s parameters. We start by imposing that  $\delta_0 = \delta_1 = \delta$  such that skills depreciate from non-formal employment at a similar rate before and after on-the-job experience is accrued. We make this restriction because our empirical strategies will limit our ability to pool non-employment spells at different career points, but we will return to it in Section VI. We would like to estimate the effect of a year without formal employment ( $U_{jt}$ ) on an output measure ( $y_{it}$ ) where  $j = j(i)$ . But if the teacher hiring process favors more productive teachers, then regressing  $y_{it}$  on  $U_{jt}$  will

---

<sup>7</sup>When we use these standard assumptions, we estimate early career experience returns of 0.03 student standard deviations per year and later career returns of 0.00 student standard deviations per year. For different identifying assumptions, Wiswall (2013) relies on the exogeneity of measured career interruptions like decisions to take time off. Herrmann and Rockoff (2012) use teacher absences before versus after student testing dates, though the parameter of interest is the change in the productivity of the classroom’s teacher rather than how an individual teacher’s productivity changes as her experience varies exogenously.

<sup>8</sup>While accrued experience is low, teachers have much more potential experience. The district and individual estimation samples have means of 7 and 15 years of potential experience, respectively.

not yield a consistent estimate of the causal effect. Greece’s centralized assignment process provides us with useful variation in  $U_{jt}$  that we argue is unrelated to a teacher’s potential productivity. Define teacher  $j$ ’s risk set  $m = m(j)$  in year  $t$ , to be the set of teachers who had their degrees conferred in the same year-month as  $j$ , teach the same subject as  $j$ , and are eligible for assignment in year  $t$  (i.e., they have not attrited from the lists). If we control for a teacher’s risk set, then we isolate variation in  $U_{jt}$  among teachers who completed their education at very similar points in time. But this remaining variation may still be related to a teacher’s potential productivity. For instance, teachers may receive an assignment via priority from having a special needs child.

Thus, we instrument for within-risk set variation in  $U_{jt}$  with a teacher’s normalized waitlist position from the fresh graduates list in the first year the teacher appears in our sample.<sup>9</sup> Because a teacher may appear on both the deputy and hourly lists, we use the minimum normalized waitlist position across the two lists and label it as  $z_j$ . We only use the waitlist position from the fresh graduates list because the position on the experienced list may be related to a teacher’s potential productivity. The fresh graduates list position, however, still strongly predicts the speed of assignments on the experienced list because an earlier first assignment starts the credit accrual process faster and moves the teacher up the experienced list. We provide more details on instrument construction in Online Appendix D.

Our empirical model is:

$$(8) \quad y_{it} = \beta U_{jt} + \mu_{mt} + \eta_{it}$$

$$(9) \quad U_{jt} = \lambda z_j + \theta_{mt} + \nu_{jt}$$

where  $\mu_{mt}$  and  $\theta_{mt}$  are vectors of risk set-year fixed effects. Relating Equation 8 to the production model (Equation 7),

---

<sup>9</sup>We normalize the waitlist position by the length of the list so it runs from 0 to 1. The first year the teacher appears in our sample is the maximum of 2003 and the teacher’s degree conferral year.

$$(10) \quad \beta = (\ln(1 - \delta) - \ln(1 + \gamma - \delta))$$

$$(11) \quad \mu_{mt} = \ln(1 + \gamma - \delta)N_{jt} + \ln(\bar{H}_{j0})$$

$$(12) \quad \eta_{it} = (\ln(H_{j0}) - \ln(\bar{H}_{j0})) + \varepsilon_{it}$$

where the average  $\ln(\bar{H}_{j0})$  is taken over teachers in risk set  $m$  in year  $t$ .

As written, our empirical model implies a constant causal effect of a teacher's years without formal employment on student output. The fairly rigid structure of the assignment system makes it likely that in the presence of heterogeneous treatment effects, compliers are close to representative, a statement we return to empirically in Section IV.A. But in case compliers are selected, we would be estimating a local average treatment effect and specify our identifying assumptions accordingly. First, the instrument is (conditionally) randomly assigned. Second, we require an exclusion restriction that the instrument only affects student output through its effect on the teacher's years without formal employment. Third, we need a relevance condition that normalized waitlist position affects assignment:  $cov(z_j, U_{jt}|m, t) \neq 0$ . Fourth, we assume monotonicity: a higher initial waitlist position weakly increases a teacher's years not formally employed.

The rigidity of the public, centrally-administered assignment process makes us confident in the final two assumptions. As assignments must follow waitlist order, the waitlist position is likely to affect how quickly a teacher receives an assignment. We test this empirically in Section IV.A and find a strong relationship. Similarly, the difficulty in changing one's waitlist position speaks to the monotonicity condition. The main threat would be that teachers with poor initial positions make costly investments that improve their subsequent waitlist positions. While these might be unlikely, if the special circumstances (e.g., having a special needs child) we described in Section I were responsive to initial waitlist position, we could have monotonicity violations. As a robustness check, we will therefore estimate our IV model without these teachers and find

very similar results.

For the assumptions of random assignment of the instrument and the exclusion restriction, our model clarifies the likely sources of identification threats. First, the waitlist position could be (conditionally) correlated with a teacher’s potential productivity ( $cov(z_j, \ln(H_{j0})|m, t) \neq 0$ ). Second, the waitlist position could be (conditionally) correlated with student types or shocks to teacher productivity ( $cov(z_j, \varepsilon_{it}|m, t) \neq 0$ ). In this section, we argue that our instrument is independent of a teacher’s potential productivity upon entering the waitlist system and remains independent as years elapse and some teachers attrit. We also evaluate the threat to the exclusion restriction that waitlist position might affect teachers’ time allocation for teaching, once assigned to a position. In Sections IV and V, we evaluate threats to identification that are specific to each empirical strategy.

As described in Section I, waitlist position is determined by the date of degree conferral, with ties broken by grade-point average. Teachers graduating in different semesters may differ in many ways. Thus, we isolate only fine timing differences by controlling for the month-year of degree conferral and argue that remaining conferral date variation within the month is exogenous. Our identification strategy fails if within-month variation in graduation timing correlates with a teacher’s potential productivity, perhaps because more productive prospective teachers pressure faculty members to enter grades quickly. Even if there were a pattern in which more productive teachers graduate sooner, there is nothing special in the education system about graduating at the beginning of a calendar month. Thus, “expedited” graduates could earn their degrees faster than other students and still be at the end of a calendar month.

But to provide more evidence that within-month timing of graduation appears unrelated to teacher type, we regress the teacher’s university grade point average (GPA) (out of 10) on the teacher’s waitlist position percentile (position normalized by the length of the list). In the first column of Table 2, we show a strong relationship between waitlist position, as determined by degree conferral date, and the teacher’s university GPA. This could reflect grade inflation as later cohorts have higher (worse) waitlist positions and higher grades. In the second column, we add fixed effects for each graduation month-year combination, separately for each academic subject because the waitlists are subject-specific. Once we rely only on waitlist position variation from within-month differences in graduation timing (and estimate

Equation 9, but replacing the left-hand-side with teacher's GPA and keeping each teacher's first waitlist observation), the relationship between waitlist position and teacher grades goes away, with a high degree of statistical precision.

Even if initial waitlist position were (conditionally) uncorrelated with teacher type when teachers graduate, the assignment process could induce a correlation over time via attrition. As prospective teachers wait for an assignment, some may find full-time employment that would make them ineligible for public school teaching. If the prospective teachers who attrit differ in productivity from the teachers who remain on the waitlists, then initial waitlist position may be correlated with teacher potential productivity among the remaining teachers, even if we account for the direct impact of different time without formal employment. Whether attriting teachers would be positively or negatively selected on teaching potential productivity is unclear and depends on how teacher potential productivity correlates with productivity for other jobs.

Attrition is quite common, which is perhaps unsurprising given the long wait times until assignment. For example, 98% of the 2003 graduating cohort remain on the waitlists through 2004 but by 2010, 52% have dropped off. We explore the relationship between our instrument – fresh graduates' waitlist position conditional on degree month-year – and attrition in the middle columns of Table 2, where attrition corresponds to dropping off the waitlists before the end of the sample period and without accruing any formal experience. Our data sample includes all teachers belonging to risk sets that lead to any assignments, as these are the risk sets that will identify our causal estimates. We estimate Equation 9, but replace the left-hand-side with an indicator for attrition, and we fail to reject the null hypothesis that there is no relationship between conditional waitlist position and attrition.<sup>10</sup> Thus, while within risk-set variation still predicts how quickly fresh graduates receive assignments, the within risk-set variation is much smaller than the across risk-set variation, and attrition rates only appear responsive to these larger differences. We provide further visual evidence in Appendix Figure A.5, where we summarize attrition rates by the teacher's degree conferral day. When we control for risk set, this within-month degree conferral timing generates our identifying variation. The figure shows no clear pattern between attrition rates and time of the month.

---

<sup>10</sup>Column 3 of Table 2 shows a negative relationship between waitlist position and attrition, when we do not control for risk set. Recent graduating cohorts are less likely to have attrited by the end of our sample and appear at the back of lists that grow each year.

Attrition appears to be unrelated to our instrument and also teachers' university academic achievement. In the last column of Table 2 we see that teachers with higher university grade point averages are no more likely to attrit when we control for risk set. Thus, while our assumption that there is no selective attrition is fundamentally untestable, the balanced attrition rates across our instrument and lack of relationship between attrition and teacher academic achievement are encouraging. One potential reason for a lack of selective attrition is that teachers are neither evaluated nor compensated on the basis of student performance. Thus, the attractiveness of remaining a teacher may be relatively unrelated to productivity.

To interpret the losses in output as measures of productivity changes, we need to rule out two possible violations of the exclusion restriction. First, teachers could alter their effort in response to more time without formal employment. Effort certainly matters for how much student output a teacher helps create; but because the educational system has no evaluation or rewards based on a teacher's performance, we speculate that our estimates are unlikely to be explained by changes in effort due to incentives. Changes in motivation unrelated to incentives could also constitute part of a teacher's human capital. Perhaps a more direct effort channel comes from how teachers spend their time waiting. If teachers develop non-teaching human capital during part-time work – e.g., they open a successful tutoring business – then their teaching output may be lower once they are assigned, but this could reflect split effort between their formal work and continuing the activities they started while waiting. From the perspective of the formal educational system, the loss in teaching output remains high. But in linking the change to human capital, we might overstate the losses to the teacher's complete set of skills.

While we cannot rule out such a possibility, we note that many deputy assignments involve relocating to a new part of the country. Thus, the teachers' activities, to the extent they depend on geography, are likely to change. Second, we collected survey data on teachers' activities while working as a deputy and offer a summary in Appendix Table A.2. Over a quarter of teachers engage in part-time work *while* serving as a deputy teacher, and the most common type of work is offering private lessons. In Greece, this is not specific to deputy teachers. Permanent teachers regularly offer private lessons on the side. A very similar percentage of teachers reports engaging in part-time work, and specifically private lessons, that they con-



tinued from their waiting period. That the propensity to offer private lessons is similar for teachers continuing and not continuing activities from their waiting period makes us skeptical that the waiting time itself is causing teachers to change how they spend their teaching time. We further find no correlation between the years a teacher waited until her first assignment and whether she worked part-time once assigned (Appendix Table A.3). We thus argue our estimates of large output loss likely reflect changes in human capital.

## IV Individual-Level Estimates

We estimate two versions of our empirical model, each adapted to a different level of aggregation in our data. First, we present a student-level model that describes how students' outcomes vary with the years without formal employment of their assigned teacher. This model exploits the detailed data on student outcomes and teacher classroom assignments from the 22 schools in the micro data. Our micro data set includes student test results for each subject. We thus extend our empirical model to a student-subject-year unit of analysis. Let  $i$  denote a student,  $s$  denote a subject, and  $j = j(i, s, t)$  denote student  $i$ 's teacher for subject  $s$  in year  $t$ . Let  $m = m(j)$  denote teacher  $j$ 's risk set. We specify our model of a student-subject exam outcome  $y_{ist}$  as:

$$(13) \quad y_{ist} = \beta U_{jt} + \mu_{mt} + \eta_{ist}$$

$$(14) \quad U_{jt} = \lambda z_j + \theta_{mt} + \nu_{jt}.$$

While our micro data only has outcomes from 22 schools, we have waitlist and experience measures for the whole country. To leverage both sets of data, we implement a within transformation on our empirical model (Frisch and Waugh, 1933; Lovell, 1963; Giles, 1984). Let  $\tilde{x}_{ist}$  denote  $x_{ist}$  demeaned at the risk set level ( $\tilde{x}_{ist} = x_{ist} - \bar{x}_{mt}$ ). If we apply the within-risk set transformation, our model becomes:

$$(15) \quad \tilde{y}_{ist} = \beta \tilde{U}_{jt} + \tilde{\eta}_{ist}$$

$$(16) \quad \tilde{U}_{jt} = \lambda \tilde{z}_{jt} + \tilde{v}_{jt}.$$

For the model observables that do not require student-level data ( $\tilde{U}_{jt}, \tilde{z}_{jt}$ ), we can perform this within-risk set transformation using the full national sample of teachers. But because we only observe outcome  $y_{ist}$  for a subset of 22 schools, we cannot estimate this model on the full sample. Instead, define  $\xi_{ist} = \tilde{\eta}_{ist} + \bar{y}_{mt}$  and rewrite our model as:

$$(17) \quad y_{ist} = \beta \tilde{U}_{jt} + \xi_{ist}$$

$$(18) \quad \tilde{U}_{jt} = \lambda \tilde{z}_{jt} + \tilde{v}_{jt}.$$

We estimate this model as our student-level specification, clustering standard errors by teacher. We construct  $\tilde{U}_{jt}$  and  $\tilde{z}_{jt}$  using the risk set means from the full sample of teachers and then estimate the instrumental variables specification using the students and teachers in our micro data set. Provided our micro sample is representative, the identification assumptions are unchanged: if  $\mathbb{E}[\tilde{z}_{jt} \tilde{\eta}_{ist}] = 0$  then  $\mathbb{E}[\tilde{z}_{jt} \xi_{ist}] = 0$  by construction.<sup>11</sup> This procedure allows us to use the national sample to control for differences across risk sets while using the micro sample for linked outcomes.

## IV.A First Stage

Before estimating the model, we present distributions of the demeaned waitlist positions ( $\tilde{z}_{jt}$ ) and accumulated years without formal employment ( $\tilde{U}_{jt}$ ) for the full sample in Figure 2. We see that, even when controlling for the degree month-year, there is still a lot of residual variation left in our instrument and the years without formal employment. This occurs in part because earlier first assignments generate faster subsequent assignments, given how the experienced teachers waitlist is structured. Thus, we are able to implement fine controls for degree timing and still have enough variation left over to generate precise causal estimates.

Consistent with the details of the assignment process, we observe a strong relationship

---

<sup>11</sup>Note that  $\mathbb{E}[\tilde{z}_{jt} \xi_{ist}] = \mathbb{E}[\tilde{z}_{jt} \tilde{\eta}_{ist}] + \mathbb{E}[\tilde{z}_{jt} \bar{y}_{mt}] = 0 + \mathbb{E}[\tilde{z}_{jt} \bar{y}_{mt}] = \mathbb{E}_{mt}[\mathbb{E}[\tilde{z}_{jt} \bar{y}_{mt} | m, t]] = \mathbb{E}_{mt}[\bar{y}_{mt} \mathbb{E}[\tilde{z}_{jt}]] = \mathbb{E}_{mt}[\bar{y}_{mt} \cdot 0] = 0$ .

between demeaned waitlist positions and demeaned years without formal employment. We show a binscatter plot of the two variables in Figure 3. Worse waitlist position from the fresh graduates list, even after controlling for risk set, strongly predicts a higher number of years without formal employment. Relative to a teacher at the front of the waitlist, a teacher at the waitlist's median is expected to wait almost an additional year. We see a similar relationship when we estimate Equation 18 in the full sample. In the first column of Table 3, Panel A, we see that moving from the front to the back of the inexperienced waitlist leads to 1.5 additional years of waiting prior to an assignment. This relationship is very strong, with an F-statistic above 100.

This strong relationship comes in part from a structured assignment system. This also implies that compliers – those for whom waitlist position affects time without formal employment – are unlikely to be very selected. While the data do not allow us to identify individual compliers, we can analyze their characteristics relative to the general population of prospective teachers in Greece. We follow Angrist and Pischke (2008) and estimate the first-stage equation separately for different sub-populations, based on pre-determined characteristics. We stratify the sample based on degree subject, time of degree, grade marks, gender (inferred from names), and degree conferral time. For each subpopulation, the ratio of the first stage coefficient in the sub-population to the overall first-stage coefficient indicates the relative likelihood that compliers come from a given sub-population. The first stage regression for each sample is shown in Table 3. Panel A shows results by subject and whether a teacher graduated before or after 2000. After the pooled first stage in the first column, the next four columns split the sample by subject. We see strong first stages in Greek, math, and science. The magnitude is largest in science and smallest in Greek. We do not find a significant first stage for other subjects, but these subjects account for just 6% of the full sample. We do not see significant heterogeneity in the first stage based on whether teachers obtained their degrees before or after 2000, with both coefficients being very close to the full sample first stage.

Panel B shows results for grade marks, gender, and degree conferral dates. We split the sample between those who receive marks above and below 7, which roughly corresponds to the median of the sample. We see very little difference for teachers with high or low marks. Men and women also see similar first-stage coefficients, with that of men being slightly higher.

Teachers whose degrees are conferred in the fall see a slightly lower first-stage coefficient than those who see conferral in the summer or other time periods. Generally, the results in Panels A and B of Table 3 indicate that there is little meaningful heterogeneity in the first stage. In Panel C, we split teachers based on characteristics that may vary over time and directly affect waitlist positions. We label a teacher as having special circumstances if she has at least three children, has a special needs child, has military experience, or reads Braille. We then run separate first stage regressions for teachers based on whether they *currently* have special circumstances or *ever* have special circumstances and present the results in Panel C. As expected, the first stage is much weaker for teachers with special circumstances as these are the teachers who have priority in assignment beyond when they earned their degree.

## IV.B Individual Sample and Tests of Student Sorting

While the first stage regressions in Table 3 use the full sample of teachers in the country, we only observe student-linked outcomes for a smaller subset of teachers. We thus maximize our individual sample size in two ways. First, we include deputy teachers even if they graduated too long before our sample to appear on a fresh graduates list in our sample period. While we do not observe these teachers' waitlist positions on the "inexperienced" lists, we use the waitlist's assignment rule – by degree conferral date – to determine original waitlist orderings, for all teachers. To translate these to waitlist percentiles, we divide the teacher's position by the number of teachers in our data with earlier degree conferral dates. We provide more details about how we scale the instrument, with analysis showing our results are robust to different forms of scaling, in Online Appendix D. Second, we include permanent teachers in the analysis and treat each as having her own risk set. In case permanent teachers teach at different types of schools, and thus violate  $\mathbb{E}[\tilde{z}_{jt} \tilde{\eta}_{ist}] = 0$ , we add a dummy variable for whether the teacher is a deputy teacher. These permanent teachers simply add precision to our controls, introduced below. In total, this leaves us with 42 deputy teachers and 348 permanent teachers. We show robustness to excluding permanent teachers in Online Appendix D.

Before estimating our IV model, we assess an identification threat specific to the individual analysis – that teachers with less experience are assigned different students. If a principal realizes the new deputy teacher is unprepared, she may want to place the teacher in a classroom

with a non-random set of students. This is unlikely to be an issue because Greek law requires high school teachers to be assigned randomly to students, who themselves are assigned to classrooms based on last name (Lavy and Megalokonomou, 2019, 2020). But we can use our micro school data to test this empirically. We estimate Equation 18 but instead of years not formally employed as the outcome, we use measures related to student sorting or teacher workload and present the results in Table 4. We find that a teacher’s time spent not working is unrelated to her students’ lagged grade-point average (column 1), which suggests limited non-random assignment of students to teachers. We further test whether a teacher’s time not formally employed varies with her workload (number of students per grade) or likelihood of being assigned to the nationally tested grade (grade 12). We show the results in columns 2 and 3 and again find no statistical relationship.

#### **IV.C IV Estimates**

We now turn to estimating the effect of time out of formal employment on students’ average exam score in a subject-year. In the first column of Table 5, we show OLS estimates from a model that controls for each teacher’s risk set but does not instrument for the number of years without formal employment. We find a negative estimate of the effect of years without formal employment on student total exam scores, but we cannot statistically reject no effect. We next introduce our instrument and present reduced form and first stage estimates in columns 2 and 3, respectively.<sup>12</sup> We find a strong relationship between (demeaned) waitlist percentile and (demeaned) years without formal employment. Column 4 displays our IV estimate on average test score during the year (“Score ( $\sigma$ )”). We estimate that an additional year without formal employment lowers student test scores by 0.052 student standard deviations, or 1.5 percentiles, with the ability to reject no effect.<sup>13</sup>

Before continuing with the IV estimates on other outcomes, we compare the OLS and IV estimates for average test score. In many settings, we might expect the OLS estimates to be

---

<sup>12</sup>We show the reduced form variation in a binscatter in Appendix Figure A.6.

<sup>13</sup>The 2SLS standard error on years without formal employment is lower than the OLS standard error. This occurs when we cluster the errors, where there is no general result that 2SLS standard errors are necessarily higher. When we do not cluster our errors, the OLS standard errors are slightly lower than the 2SLS standard errors.

biased downward, if potential productivity correlates negatively with years without formal employment and negatively with output. The usual bias may not apply here as the Greek teachers assignment process shuts down the school’s ability to select teachers based on productivity. But this does not necessarily explain why the OLS estimate is less negative than the IV estimate. We thus turn back to the compliers’ characteristics and note that the teachers whose (initial) waitlist position corresponds least to years waiting – and therefore are mostly likely to drive differences between OLS and IV – were those with special circumstances. The assignment system gives priority to these teachers because they face additional challenges. If these challenges reduce the amount of time teachers have for lesson planning or other aspects of teaching, then they may produce lower output. This could explain the sign in the difference between the OLS and IV estimates as teachers with the least time waiting are those with challenges that reduce output. We investigate this in Appendix Table A.4 and find that when we drop teachers with special circumstances from the sample, the OLS and IV estimates converge.

Returning to the IV estimates in Table 5, our estimated effect of  $0.05\sigma$  is large. The standard deviation of teacher value-added, estimated using a model with drift as in Chetty et al. (2014), is  $0.23\sigma$  in our data.<sup>14</sup> Thus, the effect of a year without formal employment is equivalent to 23% of the cross-sectional standard deviation. The effect appears to be strong for teachers in STEM fields, though for non-STEM fields our estimates are very imprecise (Appendix Table A.5). Such a large effect of a year without formal employment may reflect skill loss that rebounds quickly. Teachers who spent a few years not working may forget some curricular material; but if they are able to re-learn the material quickly, then short-run estimates may overstate the impact. Further, optimal unemployment benefits depend on how skill depreciation rates vary during an unemployment spell (Shimer and Werning, 2006). We therefore break out our estimates by first and second semester tests and end-of-year exam, in columns 6 through 8 of Table 5. We see a very stable set of estimates that get slightly larger in magnitude for later tests. Thus, the change in teacher output is persistent for at least a year.

Students with teachers who have been out of formal employment for longer are at a large

---

<sup>14</sup>We estimate value-added pooled across subjects, controlling for a cubic function of lagged grade-point-average, gender, and subject indicators, and allowing for value-added to drift flexibly for up to 4 years. In related work, Lavy and Megalokonomou (2020) find that a one standard deviation in teacher value added increases test scores by 0.21 to 0.22 standard deviations. This is a bit larger than effect sizes in Chetty et al. (2014).

disadvantage on tests. We show similar effects on students' grade-point averages in Appendix Table A.6. We now explore whether this disadvantage carries over to post-secondary outcomes in Table 6. The Greek university admissions process takes the Panhellenic exam scores and reweights the sections to generate an admissions score. Students then submit ordered preference lists of institution-program combinations and are admitted in order of admissions score. Column 1 presents the results on the natural log of the admission score. Students with less experienced high school teachers have lower admissions scores. Perhaps to compensate for these lower scores, the students include more university-programs on their preference lists (column 2), but they remain at an admissions disadvantage. Students with teachers who had an additional year without formal employment are 1.8 percentage points less likely to be admitted to a university, with a baseline admissions rate of 67% (column 3). Of this 1.8 percentage point effect, 1.3 percentage points comes from reduced admissions to academic university programs (column 4) while the rest consists of students not attending technical universities. On the margin of *where* a student is admitted, we see that students with inexperienced teachers are admitted to less selective institution-programs. Column 5 shows that having a teacher with one year fewer of formal experience causes the student to be admitted to an institution-program that is 2.2 percentiles lower-ranked in terms of admitted students' admissions scores. These large effects on post-secondary outcomes confirm that formal experience matters beyond the time the student is in the teacher's class. We provide more interpretation of the estimates after describing our district-level estimates.

In Online Appendix E we provide robustness checks for our student-level estimates. We explore robustness to the sample, the functional form of formal experience, our use of controls, the test score units, our measure of university selectivity, whether experience excludes part-time experience, and bootstrapped standard errors.

## V District-Level Estimates

We now turn to a model of district-year mean student outcomes. We sacrifice the precise measurement of an individual teacher's multiple outputs for robustness to possible within-school or within-district assignment of students to teachers' courses depending on the type of

the newly-assigned temporary teacher. We also gain external validity by extending the analysis to the national level.

We derive our district empirical specification from the same empirical model, Equations 8 and 9, by introducing an aggregation matrix,  $A$ . An element  $(A_{r,c})$ , corresponding to row  $r = dt$  and column  $c = j$ , is the fraction of district  $d$ 's teacher deputy work force that teacher  $j$  comprises in year  $t$ . Each column will have 0 in all elements corresponding to year  $t$  except one, as each teacher works in a single district in a given year. If a district has 20 high school deputy teachers in tested subjects, then each of those 20 teachers will have  $1/20$  as their non-zero element.<sup>15</sup>

We perform the within risk-set transformation on Equations 8 and 9, as in the student empirical specification. Then, we left multiply the demeaned equations by matrix  $A$ , which yields our district-level model:

$$(19) \quad y_{dt} = \beta \tilde{U}_{dt} + \xi_{dt}$$

$$(20) \quad \tilde{U}_{dt} = \lambda \tilde{z}_{dt} + \tilde{v}_{dt}$$

with  $\tilde{x}_{dt} = (1/N_{dt}) \sum_{j \in J_{dt}} \tilde{x}_{jt}$  for each variable  $x$  where  $J_{dt}$  is the set of deputy teachers in district  $d$  in year  $t$  and  $N_{dt}$  is the size of this set. As above, we do not demean the outcome,  $y_{dt}$ , so

$$(21) \quad \xi_{dt} = \tilde{\eta}_{dt} + \frac{1}{N_{dt}} \sum_{j \in J_{dt}} \tilde{y}_{m(j)t}$$

Before presenting the results, we revisit the identifying assumptions now that we have a district model. We maintain the prior assumptions that waitlist position is conditionally independent from teacher's unobserved potential productivity and assigned student types. We further assume that within a region assignees' (demeaned) waitlist positions are independent from the region's other assignees' types. Otherwise, the waitlist position might be an appropriate

---

<sup>15</sup>We implicitly assume that teachers in the same district-year have similar numbers of students. We cannot test this on the national sample, but for the 22 schools in the micro data, we find limited within-school variation in class sizes.



instrument at the individual level but due to aggregation, might correlate with the district-level error. We cannot test this directly, though we do not know of any way a district could target teachers from a specific part of the waitlist, controlling for risk set. We offer an indirect test by checking for a statistical relationship between  $\tilde{z}_{dt}$  and the types of risk sets contributing the district’s assignees. Specifically, we test whether  $\tilde{z}_{dt}$  is related to the district-aggregated mean experience, where the mean is calculated over the risk sets that include the district’s assignees. We present the test in Appendix Table A.7, where we fail to reject no statistical relationship.

Given this assumption of independence in assignments, we might worry that the law of large numbers eliminates any variation in the instrument or endogenous regressor across district-years. But the small number of high school temporary teachers (in tested subjects) assigned for most district-years leaves considerable sampling variation. We show the remaining variation in Figure 4. Compared to Figure 2, the aggregation shrinks the dispersion, especially of our instrument. But with controls we still have enough variation left to generate precise causal estimates. We include district and region-year fixed effects to absorb cross-district persistent test score differences and region-level shocks and include the log of the district’s mean class size as a control (demeaned by risk set).<sup>16</sup>

## V.A Baseline Estimates

We present our district-level estimates in Table 7. We provide visual counterparts with binscatters for the first stage (Appendix Figure A.7) and reduced form (Appendix Figure A.8). Our OLS regression in column 1 yields a positive but statistically insignificant association between teacher time without formal employment and students’ twelfth grade test scores. In column 2 we present the district-level first stage regression. Despite the aggregation, we still predict differences in mean years without formal employment across district-years. The F-statistic on the excluded instrument, though, is relatively low at 4.6, raising weak instrument concerns. In our district-level results, we therefore report the Anderson-Rubin 95% confidence set, which is robust to weak instruments (Anderson and Rubin, 1949). We also find nearly identical con-

---

<sup>16</sup>As in the student model, our inclusion of controls that cleave risk sets – in this case, the district and region-year fixed effects – is not fully consistent with our within risk-set transformation. We include robustness checks in Online Appendix E. Further, we find no relationship between a teacher’s waitlist position and her assigned district’s mean class size or mean log class size.

confidence sets when we employ the two-step method from Andrews (2018).

Our IV estimates show significant decreases in student test scores when a district's teachers have spent longer without formal employment. For each year increase in average time without formal employment, we estimate that student test scores fall by 0.34 student standard deviations ( $\sigma$ ), or 8.9 test score percentiles. The interpretation of the effect sizes for the district results differs slightly from the individual analysis, because students' Panhellenic test scores cover multiple subjects while the individual test scores are subject-specific. Twelfth-grade students take five courses in subjects that appear on the standard Panhellenic exams. We thus divide the estimates by 5 and report the per-class estimate. Our individual estimates from Section IV suggest a decrease of  $0.052\sigma$  for students in a given course if a teacher has an additional year out of formal employment, while the district-level estimate is  $0.068\sigma$ . We cannot statistically reject equality.

We find that the effects carry over to the post-secondary outcomes of the district's twelfth graders, though our estimates are less precise (columns 5-8 of Table 7). We estimate that having teachers with one additional year out of formal employment lowers the log university score by 0.024, on a per-class basis. This translates into 1.5 percentage point lower university admissions rates.

In Online Appendix E we provide robustness checks for our district-level estimates. We explore robustness to the sample, the functional form of formal experience, controls, the test score units, our measure of university selectivity, the aggregation matrix, and regression weighting.

## **V.B District Level Balance and Placebo Tests**

As discussed earlier, one possible threat to identification realizes if teachers' waitlist positions are unrelated to their potential productivity but correlated with the characteristics of the students they teach. In Section IV.B we tested whether within-school assignment of students to teachers correlated with our instrument and found no relationship. But the correlation could still occur at a higher geographic level. For instance, if economic conditions worsen, family life may be more stressful and students may perform poorly. Or perhaps some districts invest in smaller class size, which leads to better test outcomes. If these shocks or actions are correlated with changes to a school's demand for temporary teachers, then we might worry that our es-

timates confound the teacher assignment effects with local shocks. We consider these threats unlikely, as schools request temporary teachers on a rolling basis, with fast churn such that it would be nearly impossible to target specific teachers. This is consistent with Table 8, where we regress a school district's twelfth grade cohort size or local unemployment rate on each assigned teacher's waitlist position. We find no relationship between waitlist position and these district characteristics in the cross-section, nor within district, when we compare changes over time by including district fixed effects. And the null effect remains when we add region-year fixed effects. Thus, teachers with different (conditional) waitlist positions do not appear to be assigned to schools in different types of districts.

We can further use our district analysis to implement two placebo tests that provide additional support for our identifying assumptions. We start by exploiting the timing of assignments. Our district-level estimates include district fixed effects, which would control for time-invariant district differences, such as differential teacher attrition rates, that might relate to teacher assignment. Including district fixed effects would be insufficient, however, if districts have particularly high attrition – or face some other time-varying shock – in some years, for reasons that cause test score changes even if new assignments had an average amount of time out of formal employment. If such shocks are serially correlated, and correlate with our IV, then we would expect that the following year's assignments' time without formal employment would correlate with this year's test score outcomes. This would occur even though future assignees cannot have a direct effect as they only show up in the district a year later. We therefore extend our IV specification to include the average time out of the labor force of the following year's assigned teachers to that district and add as a second instrument the mean waitlist rank for the following year's assigned teachers. We present the test in Table 9. The first column runs our baseline regression, on the sample of districts for which we have assignments the following year. The sample reduction does not change our main point estimate. The second column includes the years out of formal employment of the *following* year's assigned deputy teachers. We see no statistical relationship between future assignees' human capital and current test scores, which is consistent with teachers' waitlist positions being unrelated to regional shocks or trends. The coefficient on current assignees' years waiting becomes less precise, but we can still reject equality between the two coefficients.

Our second placebo test examines the role of teachers assigned through the exact same process but who teach subjects that do not appear on the twelfth grade exams. These subjects include music, foreign languages, physical education, and household economics. If there is some district- or school-level confounder that correlates with our instrument, it likely correlates with waitlist percentiles among all assigned teachers, not just those in tested subjects. We present the relationship between years without formal employment, instrumented with waitlist percentiles, and student test scores for teachers in untested subjects in the last two columns of Table 9. We fail to reject no statistical relationship, and we reject that the coefficient is the same as our tested subjects' coefficient (from Table 7). Thus, the effects we find appear specific to the assignments of teachers in tested subjects. Any confounders would have to apply specifically to these subjects, again making it unlikely that our results are driven by confounding regional shocks as those would be unlikely to only affect students in certain subjects.

This final placebo test also evaluates whether there are any spillovers, either across teachers in the same district or from students reallocating effort across classes based on teachers' characteristics. One violation of the exclusion restriction would occur if, despite being randomly assigned to teachers, students with less experienced teachers take actions to compensate for a low human capital teacher. For example, students might receive additional parental tutoring or reallocate study time toward courses with inexperienced teachers. We could consider these actions as part of the effect of teacher experience on output that we seek to identify. But these actions may also incur additional costs to the system (parental time; lower output in other courses) we would not capture with a student's output in a single course. Because we see no cross-course spillover between tested and non-tested subjects, we consider such reallocations unlikely.

## **V.C Policy Implications**

As in the individual model, we estimate large effects of being out of formal employment. The effect of one year waiting for a formal position translates to 30% of the cross-sectional standard deviation in teacher value-added. But the output loss from the deputy assignment system is distributed unevenly. Districts differ in their reliance on deputy teachers, and the percentage of deputy teachers correlates negatively with district mean test scores. Thus, we conduct a back-

of-the-envelope exercise where we estimate the distributional changes from an assignment system that leads to one fewer year of waiting on average as well as a system that eliminates all waiting. Specifically, we take the absolute value of our point estimate from column 3 of Table 7, multiply it by each district's 2010 share of teachers that are deputies, and multiply by either one year or the district's assignees' mean years waiting, depending on the counterfactual.

The actual cross-district standard deviation of test scores is  $0.341\sigma$ . We estimate that eliminating one year of waiting would lower the standard deviation to  $0.328\sigma$ , a slight compression of cross-district differences. In the top panel of Appendix Figure A.9, we plot how this system change would reorder districts in terms of their mean scores. The x-axis shows districts' actual rank based on mean test scores, with a higher rank better, while the y-axis shows the rank we predict from eliminating one year of waiting. Most districts remain in the same part of the distribution they started, though a non-trivial share of districts are higher up the actual test score distribution because they rely less on deputies. We then implement a more extreme counterfactual, where we eliminate all waiting in the system. Because districts differ in both their reliance on deputies and their assignees' time without formal employment, districts' test score ranks change considerably under the counterfactual. In the bottom panel of Appendix Figure A.9, we see that the elimination of waiting would dramatically reorder districts in terms of mean test scores. The output effects of time without formal employment thus potentially explains much of the cross-district variation in test scores.

## VI Mechanisms

Our estimates capture the effect of a year without formal employment on worker output. This is the policy-relevant estimate for policies that shift (formal) employment. But as the production model in Section III highlighted, this effect could be driven by high levels of depreciation or by missing out on large returns to experience, and the specific mechanism matters for some policies. For example, policies that provide unemployed workers with a structured environment that reviews material learned in formal education might be particularly desirable if depreciation is the dominant mechanism.

To demonstrate how we will separate the impact of forgone experience from depreciation,

we return to the motivating example given at the beginning of Section III. We have previously described comparing the two teachers in the same academic year. Consider instead a comparison that controls for experience levels – specifically, suppose we compare teacher 1 in her first year of experience (2004-05) with teacher 2 in her first year of experience (2005-06). We implement this comparison by altering the risk set definition. Previously, a teacher’s risk set ( $m$ ) consisted of the teachers in the same subject whose degrees were conferred in the same year-month. We then interacted  $m$  with time to generate  $\mu_{mt}$  fixed effects for the empirical specification. We now drop the within-year comparison and add a further condition to the risk set – teachers must have the same number of years of prior experience. Denote this new risk set as  $m' = m'(j, t)$ .<sup>17</sup>

We then estimate:

$$(22) \quad y_{it} = \beta^c U_{jt} + \mu_{m'}^c + \eta_{it}^c$$

$$(23) \quad U_{jt} = \lambda^c z_j + \theta_{m'}^c + \nu_{jt}^c$$

where we index parameters with superscript  $c$  to indicate these are conditional on experience. We again do the within-transformation, as above, but with the different risk set definition. Relating Equation 22 to the production model (Equation 6),

$$(24) \quad \beta^c = \ln(1 - \delta)$$

$$(25) \quad \mu_{m'}^c = \ln(1 + \gamma - \delta)E_{jt} + \ln(\bar{H}_{j0})$$

$$(26) \quad \eta_{it}^c = (\ln(H_{j0}) - \ln(\bar{H}_{j0})) + \varepsilon_{it}$$

where the average  $\ln(\bar{H}_{j0})$  is taken over teachers in risk set  $m'$ .

---

<sup>17</sup>We are now comparing teachers who graduated at the same time across different school years. Thus, the new risk set takes both  $j$  and  $t$  as arguments.

We present the results of the student-level model in the first two columns of Table 10. Our estimates of the effect of years waiting, conditional on experience, are large but slightly noisy. For both test score functional forms, we estimate effects considerably larger than our baseline estimates, but the standard errors are large enough that we fail to reject equivalence between the estimates (p-values of 0.62 for both functional forms). Our district-level model generates considerably more precise estimates on a per-class basis, as shown in the last two columns. We estimate a per-class drop in output, conditional on experience, of  $0.043\sigma$ , or 1.40 percentiles. Thus, even controlling for experience levels, we still find that time out of formal employment matters.

To translate these estimates to statements about experience returns or depreciation rates, we rely on our production model's relationship to the empirical model (Equations 10 and 24). Solving for  $\gamma$  and  $\delta$  as functions of the empirical model parameters, we have:

$$(27) \quad \delta = 1 - \exp(\beta^c)$$

$$(28) \quad \gamma = \exp(\beta^c - \beta) - \exp(\beta^c).$$

For our district-level model, where most of non-employment spells are prior to the first assignment, we estimate that  $\delta = 0.043$  (0.019) and  $\gamma = 0.068$  (0.024). The net gain from a year of experience – the depreciation from a year passing plus the experience return associated with working – is 2.5% of initial human capital. Thus, experience returns are just enough to counteract fairly large human capital depreciation rates, and not working leads to large losses in human capital. The teachers in our individual-level model sample are more experienced and more of the non-employment spells come after the first assignment. For these teachers, our estimates are less precise but yield  $\hat{\delta} = 0.172$  (0.225) and  $\hat{\gamma} = 0.044$  (0.016). This higher depreciation rate implies, perhaps counterintuitively, that human capital from prior on-the-job experience may depreciate faster. Further, we estimate that even while working, more experienced teachers' human capital depletes, though we cannot reject a zero change (p-value = 0.705). Of course,  $\delta$  and  $\gamma$  remain composite parameters. In particular,  $\delta$  includes any constant age effects. But provided the age effect is weakly positive, our estimate of  $\delta$  would be a

lower bound on the true human capital depreciation rate.

## VII Concluding Remarks

This paper demonstrates that workers become less productive if they accrue time without formal employment. We show that teachers who are quasi-randomly assigned additional years without formal employment are less productive than teachers with less time without formal employment. The effects are large, as an additional year without formal employment leads to a 5 – 7% of a student standard deviation reduction in student test scores. We estimate that these effects are driven by large skill depreciation rates and that the skills gained during employment just compensate for depreciation among less experienced teachers, while periods of not working lead to large reductions in human capital. Our paper identifies a specific channel that generates structural unemployment duration dependence, which has been widely found in the literature.

These large estimates describe persistent costs of unemployment for early-career workers and highlight a benefit of policies that promote employment or labor force attachment, especially in contexts with high levels of youth unemployment like Spain or Italy. In the absence of such policies, activities that mimic the work environment and provide structure may help non-employed workers maintain their skills. These results are particularly important at the time of writing, given that a large portion of the workforce has been furloughed or is working remotely due to the COVID-19 pandemic. Our estimates suggest that the fact that many workers are not working may have long term effects on productivity, which warrants further study.

While our work demonstrates that human capital depreciates, there remain a number of important avenues for further research. Our analysis has little to say about the exact type of skills that are depreciating. We also caution that the results focus on a particular profession – teachers – in a small European country. Human capital depreciation rates might be larger or smaller in other professions or contexts, particularly in settings that require more or less training and skill, where workers are more readily substitutable, or where there are fewer opportunities for “re-skilling” while unemployed. Additionally, we focus on workers very early in their career. Older workers may have different rates of skill depreciation and returns to



experience. We encourage future work explores the productivity effects of non-employment in other settings.

## References

- Altonji, Joseph G and Nicolas Williams**, “The Effects of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth,” 1992.
- Alvarez, Fernando E, Katarína Borovičková, and Robert Shimer**, “Decomposing Duration Dependence in a Stopping Time Model,” Technical Report, National Bureau of Economic Research 2016.
- Anderson, Theodore W and Herman Rubin**, “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *The Annals of Mathematical Statistics*, 1949, 20 (1), 46–63.
- Andrews, Isaiah**, “Valid Two-Step Identification-Robust Confidence Sets for GMM,” *Review of Economics and Statistics*, 2018, 100 (2), 337–348.
- Angrist, Joshua D**, “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,” *The American Economic Review*, 1990, pp. 313–336.
- and **Jörn-Steffen Pischke**, *Mostly Harmless Econometrics*, Princeton University Press, 2008.
- Autor, David H, Nicole Maestas, Kathleen J Mullen, and Alexander Strand**, “Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants,” Technical Report, National Bureau of Economic Research 2015.
- Baily, Martin**, “Some Aspects of Optimal Unemployment Insurance,” *Journal of Public Economics*, 1978, 10 (3), 379–402.
- Becker, Gary**, “Investment in Human Capital: A Theoretical Analysis,” *Journal of Political Economy*, 1962, 70 (5), 9–49.
- , “Human Capital: A Theoretical and Empirical Analysis,” *Columbia University Press*, 1964.
- Ben-Porath, Yoram**, “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, 1967, 75 (4, Part 1), 352–365.
- Benhenda, Asma**, “Absence, Substitutability and Productivity: Evidence from Teachers,” 2017.
- Bertrand, Marianne and Antoinette Schoar**, “Managing with Style: The Effect of Managers on Firm Policies,” *Quarterly Journal of Economics*, 2003, 118 (4), 1169–1209.
- Blinder, Alan S and Yoram Weiss**, “Human Capital and Labor Supply: A Synthesis,” *Journal of Political Economy*, 1976, 84 (3), 449–472.
- Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw**, “Female Labor Supply, Human Capital and Welfare Reform,” *Econometrica*, 2016, 84 (5), 1705–1753.
- Card, David, Raj Chetty, and Andrea Weber**, “Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market,” *The Quarterly Journal of Economics*, 2007, 122 (4), 1511–1560.
- Centeno, Mário and Álvaro A Novo**, “Excess Worker Turnover and Fixed-Term Contracts: Causal Evidence in a Two-Tier System,” *Labour Economics*, 2012, 19 (3), 320–328.
- Chetty, Raj**, “A General Formula for the Optimal Level of Social Insurance,” *Journal of Public Economics*, 2006, 90 (10), 2351–2356.
- , **John Friedman, and Jonah Rockoff**, “Measuring the Impact of Teachers I: Evaluating Bias in Teacher Value-Added Estimates,” *American Economic Review*, 2014, 104 (9), 2633–79.
- Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis**, “Deputy Teacher

- Survey,” 2019-2020.
- , – , and – , “Human Capital Depreciation and Returns to Experience,” Data Repository, openicpsr-165721 2022.
- Edin, Per-Anders and Magnus Gustavsson**, “Time out of Work and Skill Depreciation,” *ILR Review*, 2008, 61 (2), 163–180.
- Farber, Henry S, Dan Silverman, and Till von Wachter**, “Factors Determining Callbacks to Job Applications by the Unemployed: An Audit Study,” *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2017, 3 (3), 168–201.
- Frisch, Ragnar and Frederick V Waugh**, “Partial Time Regressions as Compared with Individual Trends,” *Econometrica: Journal of the Econometric Society*, 1933, pp. 387–401.
- Giles, David EA**, “Instrumental Variables Regressions involving Seasonal Data,” *Economics Letters*, 1984, 14 (4), 339–343.
- Giroud, Xavier and Holger Mueller**, “Firm Leverage, Consumer Demand and Employment Losses During the Great Recession,” *Quarterly Journal of Economics*, 2017, 132 (1), 271–236.
- Greece High School Archives**, “Student and Teacher Data,” 2001-2011.
- Greece Ministry of Education and Religious Affairs**, “Twelfth Grade Student Outcomes,” 2003-2011.
- , “Deputy Waitlists,” 2003-2012.
- , “Deputy Assignments,” 2004-2011.
- Hanushek, Eric A and Steven G Rivkin**, “Generalizations about Using Value-Added Measures of Teacher Quality,” *American Economic Review*, 2010, 100 (2), 267–71.
- Herrmann, Mariesa A and Jonah E Rockoff**, “Worker Absence and Productivity: Evidence from Teaching,” *Journal of Labor Economics*, 2012, 30 (4), 749–782.
- Hoffmann, Florian and Philip Oreopoulos**, “Professor Qualities and Student Achievement,” *The Review of Economics and Statistics*, 2009, 91 (1), 83–92.
- Imai, Susumu and Michael P Keane**, “Intertemporal Labor Supply and Human Capital Accumulation,” *International Economic Review*, 2004, 45 (2), 601–641.
- Jackson, C Kirabo**, “Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers,” *Review of Economics and Statistics*, 2013, 95 (4), 1096–1116.
- and **Elias Bruegmann**, “Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers,” *American Economic Journal: Applied Economics*, 2009, 1 (4), 85–108.
- Jackson, Kirabo, Jonah Rockoff, and Doug Staiger**, “Teacher Effects and Teacher-Related Policies,” *Annual Review of Economics*, 2014, 6.
- Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan**, “Earnings Losses of Displaced Workers,” *American Economic Review*, 1993, pp. 685–709.
- Jarosch, Gregor**, “Searching for Job Security and the Consequences of Job Loss,” *Manuscript, Stanford University*, 2015.
- Kahn, Lisa B**, “The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy,” *Labour Economics*, 2010, 17 (2), 303–316.
- Keane, Michael and Kenneth Wolpin**, “The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment,” *International Economic Review*, 2001, 42 (4).

- Kehoe, Patrick J, Virgiliu Midrigan, and Elena Pastorino**, “Debt Constraints and Employment,” *Journal of Political Economy*, 2019, 127 (4), 1926–1991.
- Kroft, Kory, Fabian Lange, and Matthew Notowidigdo**, “Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment,” *Quarterly Journal of Economics*, 2013, 128 (3), 1123–1167.
- Lavy, Victor**, “Evaluating the Effect of Teachers Group Performance Incentives on Pupil Achievement,” *Journal of Political Economy*, 2002, 110 (6), 1286–1317.
- , “Performance Pay and Teachers’ Effort, Productivity, and Grading Ethics,” *American Economic Review*, 2009, 99 (5), 1979–2011.
- **and Rigissa Megalokonomou**, “Persistency in Teachers’ Grading Bias and Effects on Longer-Term Outcomes: University Admissions Exams and Choice of Field of Study,” *National Bureau of Economic Research*, 2019, (26021).
- **and** – , “Long Term Effects of Teachers: Evidence from a Teacher Value Added Approach,” 2020.
- Ljungqvist, Lars and Thomas J Sargent**, “The European Unemployment Dilemma,” *Journal of Political Economy*, 1998, 106 (3), 514–550.
- Lovell, Michael C**, “Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis,” *Journal of the American Statistical Association*, 1963, 58 (304), 993–1010.
- Manuelli, Rodolfo E and Ananth Seshadri**, “Human Capital and the Wealth of Nations,” *American Economic Review*, 2014, 104 (9), 2736–62.
- Marx, Matt, Deborah Strumsky, and Lee Fleming**, “Mobility, Skills, and the Michigan Non-Compete Experiment,” *Management Science*, 2009, 55 (6), 875–889.
- Neal, Derek**, “Industry-Specific Human Capital: Evidence from Displaced Workers,” *Journal of Labor Economics*, 1995, 13 (4), 653–677.
- OECD**, “Education for a Bright Future in Greece,” *Reviews of National Policies for Education*, OECD. Publishing, Paris., 2018.
- Opper, Isaac M**, “Does Helping John Help Sue? Evidence of Spillovers in Education,” *American Economic Review*, 2019, 109 (3), 1080–1115.
- Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz**, “The Short-and Long-Term Career Effects of Graduating in a Recession,” *American Economic Journal: Applied Economics*, 2012, 4 (1), 1–29.
- Papay, John P and Matthew A Kraft**, “Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Improvement,” *Journal of Public Economics*, 2015, 130, 105–119.
- Rice, Jennifer King**, “Learning from Experience? Evidence on the Impact and Distribution of Teacher Experience and the Implications for Teacher Policy,” *Education Finance and Policy*, 2013, 8 (3), 332–348.
- Rockoff, Jonah**, “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data,” *American Economic Review*, 2004, 94 (2), 247–252.
- Rockoff, Jonah E, Douglas O Staiger, Thomas J Kane, and Eric S Taylor**, “Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools,” *American Economic Review*, 2012, 102 (7), 3184–3213.
- Rosen, Sherwin**, “A Theory of Life Earnings,” *Journal of Political Economy*, 1976, 84 (4, Part

2), S45–S67.

**Schmieder, Johannes F, Till von Wachter, and Stefan Bender**, “The Effect of Unemployment Benefits and Nonemployment Durations on Wages,” *American Economic Review*, 2016, 106 (3), 739–77.

**Shimer, Robert and Ivan Werning**, “On the Optimal Timing of Benefits with Heterogeneous Workers and Human Capital Depreciation,” Technical Report, National Bureau of Economic Research 2006.

**Stylianidou, Fani, George Bagakis, and Dimitris Stamovlasis**, “Attracting, Developing and Retaining Effective Teachers, Country Background Report for Greece,” *Report, OECD Activity*, 2004.

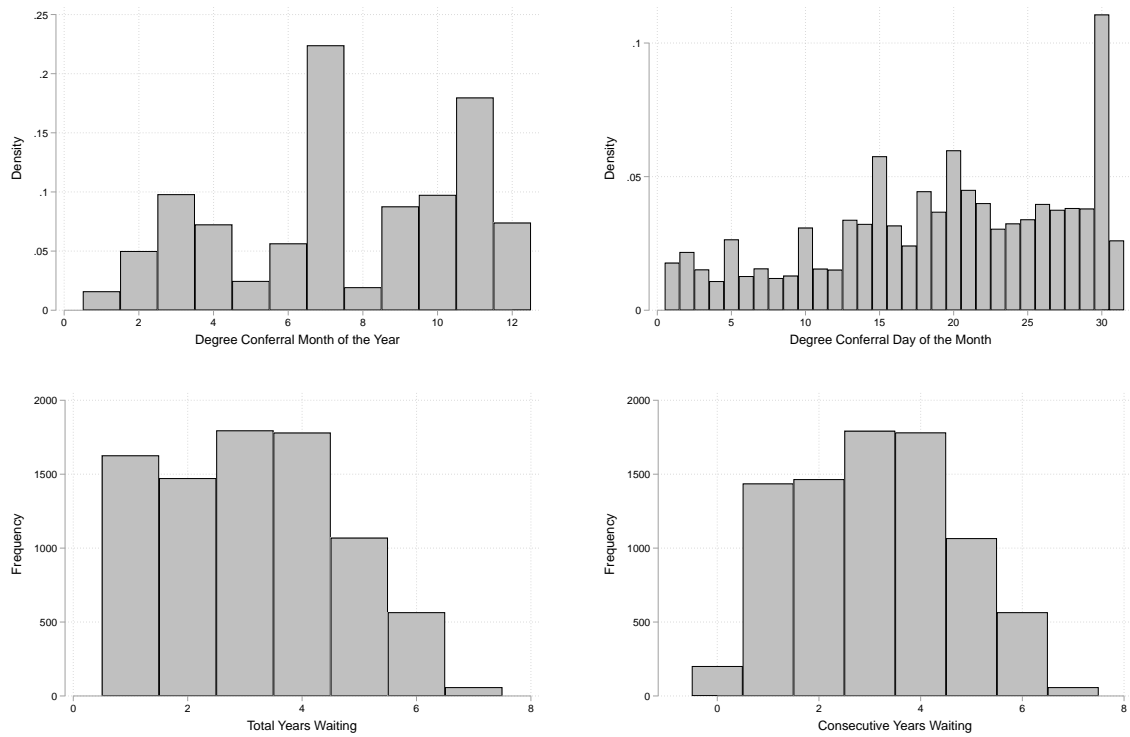
**Taylor, Eric S and John H Tyler**, “The Effect of Evaluation on Teacher Performance,” *American Economic Review*, 2012, 102 (7), 3628–51.

**Tsakloglou, Panos and Ioannis Cholezas**, “Education and Inequality in Greece,” *Discussion Paper No. 1582*, 2005.

**Wiswall, Matthew**, “The Dynamics of Teacher Quality,” *Journal of Public Economics*, 2013, 100, 61–78.

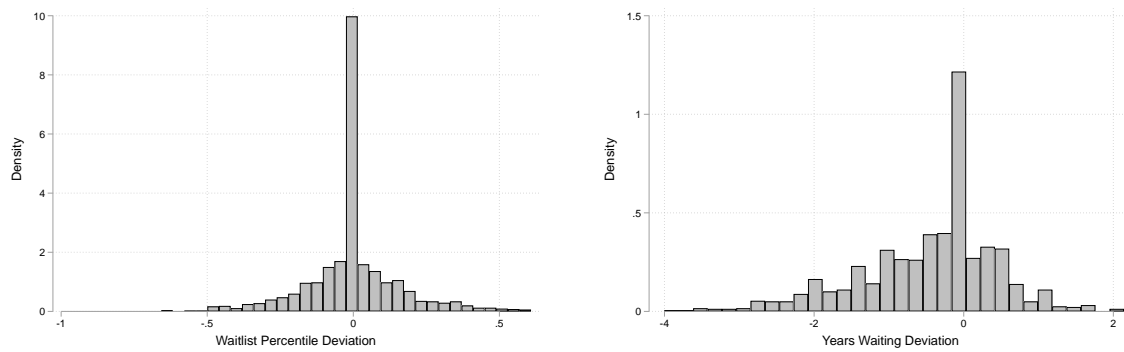
**Zivin, Joshua Graff and Matthew Neidell**, “The Impact of Pollution on Worker Productivity,” *American Economic Review*, 2012, 102 (7), 3652–3673.

**Figure 1: Degree Conferral Months and Days and Years without Formal Employment**



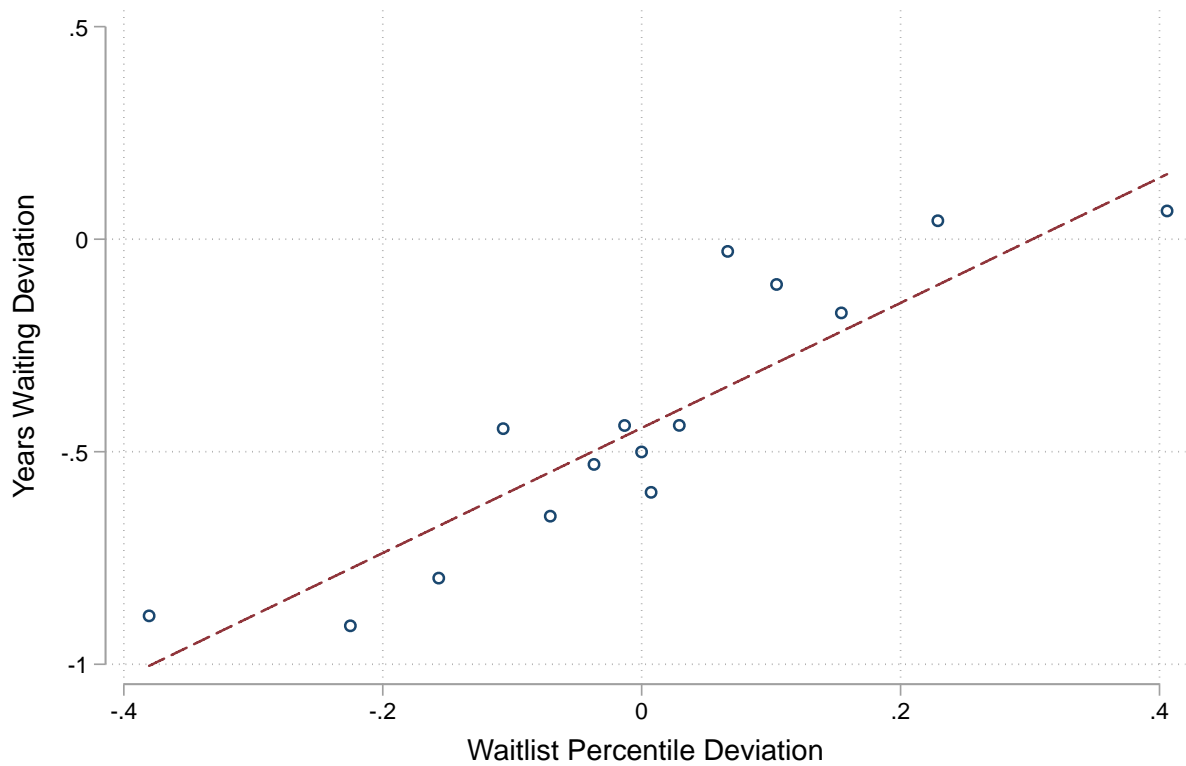
Notes: The top panel of this figure shows the histograms of the degree conferral month of the year (left) and day of the month (right). The bottom panel of this figure shows the histograms of the total years without formal employment (left) and the consecutive years without formal employment (right). An observation is a teacher on a waitlist between 2003 and 2011 whose degree was conferred before 2006.

**Figure 2: Waitlist Rank and Years without Formal Employment – Deviations from Risk Set Mean**



Notes: The figure shows histograms of the demeaned waitlist percentile (left) and years without formal employment (right) where the demeaning is done by risk set-year. A risk set is a degree conferral year-month and subject combination. The waitlist percentile is the initial position on the fresh graduates waitlist, normalized to vary from 0 to 1. An observation is a teacher-year. The sample includes all assigned teachers between 2003 and 2011 whose degree was conferred before 2006.

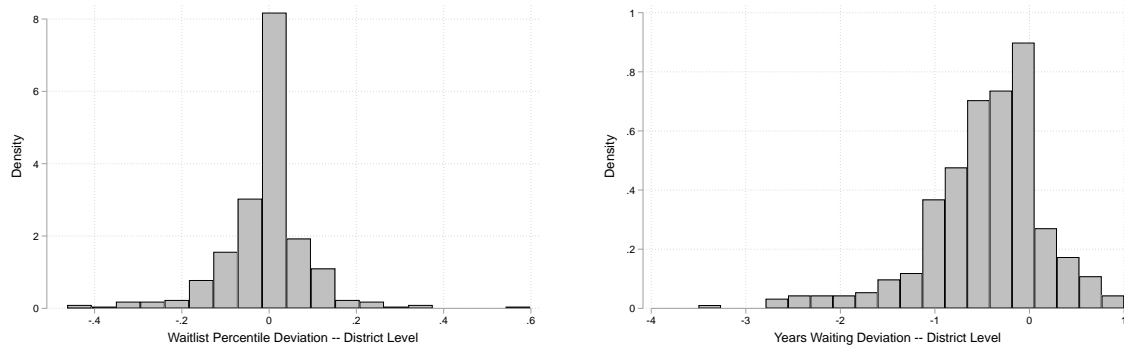
**Figure 3: Years without Formal Employment and Waitlist Position – Teacher Level**



Notes: The binscatter figure shows the relationship, at the teacher-year level, between demeaned waitlist percentile and demeaned years without formal employment where the demeaning is done by risk set-year. A risk set is a degree conferral year-month and subject combination. The waitlist percentile is the initial position on the fresh graduates waitlist, normalized to vary from 0 to 1. The sample includes all teachers on a waitlist between 2003 and 2011 whose degree was conferred before 2006.



**Figure 4: District-Level Waitlist Position and District-Level Years Waiting – Deviations from Risk Set Mean**



Notes: The figure shows histograms of the district-year mean (demeaned) waitlist percentile (left) and (demeaned) years without formal employment (right) where the district-year mean is calculated over the assigned teachers' (demeaned) variables. A risk set is a degree conferral year-month and subject combination. The sample includes all district-years that received temporary teachers between 2004 and 2010.

**Table 1: Summary Statistics**

	Observations	Mean	Standard Deviation	Minimum	Maximum
<b>Panel A: Waitlist Data</b>					
Degree Mark	541,212	6.74	0.69	0	10
Years Experience	541,500	1.75	2.03	0	16
Inexperienced (0 Years)	541,500	0.39	0.49	0	1
Any Private Experience	400,277	0.00	0.06	0	1
Any EU Experience	400,277	0.00	0.04	0	1
Any Permanent Experience	400,277	0.00	0.01	0	1
Greek & History	543,060	0.54	0.50	0	1
Math & Stats	543,060	0.13	0.34	0	1
Physics & Biology	543,060	0.17	0.38	0	1
Economics	543,060	0.06	0.25	0	1
Computer Science	543,060	0.09	0.29	0	1
Degree Month-Year-Subject Cohort Size	22,944	62.18	73.39	1	566
<b>Panel B: Micro School Data</b>					
Grade 10	104,660	0.23	0.42	0	1
Grade 11	104,660	0.24	0.43	0	1
Grade 12	104,660	0.52	0.50	0	1
(Teacher) Subjects Taught	1,101	2.17	1.15	1	6
(Teacher) Classes Taught	784	2.94	1.74	1	11
GPA	103,840	14.68	2.94	0	20
Subject Exam Score	82,333	12.04	5.57	0	20
Admitted	97,203	0.64	0.48	0	1
Admitted to Acad Univ	62,505	0.62	0.48	0	1
Application List Length	75,897	25.27	21.95	1	235
Univ Selectivity Percentile	62,505	48.91	28.27	0	100
<b>Panel C: Assignment Data</b>					
Greek & History	24,906	0.46	0.50	0	1
Math & Stats	24,906	0.21	0.41	0	1
Physics & Biology	24,906	0.26	0.44	0	1
Economics	24,906	0.03	0.16	0	1
Computer Science	24,906	0.05	0.21	0	1

Notes: The table shows summary statistics from the waitlist, micro school, and assignment data. “Degree Mark” refers to the teacher’s university grade-point average, out of 10.

“Degree Month-Year-Subject Cohort Size” indicates the number of teachers who had their degrees conferred in the same year, month, and subject. “Admitted” is an indicator for whether the student was admitted to university, “Application List Length” is the number of degree-institution combinations the student listed in applying to university, and “Univ Selectivity Rank” is the selectivity percentile of the university the student attended, where selectivity is measured by the mean entrance score of the enrolled students.

**Table 2: University Grade Point Average and Attrition by Waitlist Position**

	Teacher GPA	Teacher GPA	Attriter	Attriter	Attriter	Attriter
Waitlist Percentile	0.199 (0.0180)	0.00535 (0.0331)	-0.0546 (0.0129)	0.00533 (0.0263)		
Teacher GPA					-0.0284 (0.00437)	0.00677 (0.00493)
Mean DV	6.852	6.860	0.428	0.346	0.412	0.347
N	21161	20655	19242	16364	24030	21152
Risk Set	No	Yes	No	Yes	No	Yes

Notes: The table shows the relationship between teachers' university grade-point-average (out of 10), whether they attrit from the waitlists, and their waitlist position on their first fresh graduates list. Waitlist position is normalized by the list length to be in percentiles from 0 to 1. "Attriter" is an indicator for whether the teacher left the waitlists before the end of the sample and without accruing any formal experience. "Risk Set" indicates whether risk set fixed effects are included, where a risk set is a set of teachers in a single subject who have their degrees conferred in the same month-year. The sample consists of teachers with degrees conferred before 2006 who are in risk sets that generated at least one assignment during the sample period.

**Table 3: Characterizing Compliers**

	Years Waiting Subjects					Degree Year	
	Full	Greek	Math	Science	Other	Pre 2000	Post 2000
Waitlist Percentile	1.471 (0.136)	0.986 (0.169)	1.279 (0.547)	2.529 (0.251)	0.387 (0.355)	1.508 (0.238)	1.450 (0.150)
F Stat	117.42	34.16	5.47	101.42	1.19	40.26	93.43
N	1653	942	203	401	107	830	823

	Degree Mark		Gender		Degree Conferral		
	Mark>7	Mark<7	Male	Female	Summer	Fall	Other
Waitlist Percentile	1.505 (0.196)	1.437 (0.187)	1.800 (0.255)	1.414 (0.156)	1.589 (0.261)	0.997 (0.280)	1.685 (0.198)
F Stat	58.74	58.80	49.68	81.80	37.11	12.65	72.75
N	855	798	513	1140	393	300	960

	Special Circumstances			
	No Special	Special	Never Special	Ever Special
Waitlist Percentile	1.496 (0.136)	0.874 (1.143)	1.653 (0.149)	0.645 (0.353)
F Stat	120.86	0.58	123.86	3.34
N	1590	63	1453	200
Year FE	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes

Notes: The table shows the first-stage relationship between teachers' waitlist position and years without formal employment ("Years Waiting"), stratified by various subsamples. Waitlist position is normalized by the list length to be in percentiles. Gender is inferred by name. Degree conferral is either in the Fall (November), Summer (July) or other. Teachers may appear in the regressions multiple times if they have received multiple assignments during the sample period. Special circumstances refer to factors that directly affect waitlist positions (having at least three children, having a special needs child, military service, and reading Braille). We split teachers according to whether they currently have special circumstances and whether they ever have special circumstances during our sample period. Standard errors are clustered by teacher.

**Table 4: Effect of Years without Formal Employment on Student Sorting**

	IV	IV	IV
	Prior Year GPA	Students per Grade	Grade 12
Years Waiting	0.00796 (0.0103)	0.179 (1.528)	-0.0221 (0.0212)
Deputy	-0.0549 (0.0485)	1.027 (8.186)	0.0121 (0.0881)
Mean DV	0.0381	63.24	0.586
Clusters	390	390	390
N	54370	54370	54370
Risk Set	Yes	Yes	Yes

Notes: The table presents instrumental variable estimates. An observation is a student-subject-year. The dependent variable is the student's prior year grade point average (column 1), the teacher's number of students per grade (column 2), or whether the student is in grade 12 (column 3). The instrument is the assigned teacher's imputed waitlist position (demeaned by risk set), normalized to run between 0 and 1 within a risk set. Risk sets are teachers in the same subject whose degrees were conferred in the same month-year. The sample includes students taught by deputy and permanent teachers. Permanent teachers each have their own risk set. Standard errors are clustered by teacher.

**Table 5: Effect of Years without Formal Employment on Students' Subject Exam Scores**

	OLS	RF	FS	IV	IV	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Years Waiting	Score ( $\sigma$ )	Score (Percentile)	First Semester ( $\sigma$ )	Second Semester ( $\sigma$ )	Exam ( $\sigma$ )
Years Waiting	-0.0355 (0.0312)			-0.0518 (0.0103)	-1.545 (0.291)	-0.0363 (0.0111)	-0.0443 (0.0113)	-0.0535 (0.0117)
Deputy	-0.0802 (0.0555)	-0.0786 (0.0590)	-0.144 (0.177)	-0.0860 (0.0606)	-2.785 (1.764)	-0.0611 (0.0583)	-0.107 (0.0635)	-0.0758 (0.0671)
Prior Year GPA	0.737 (0.00887)	0.737 (0.00887)	0.000239 (0.00162)	0.737 (0.00886)	21.90 (0.225)	0.599 (0.0124)	0.643 (0.0139)	0.703 (0.00845)
Waitlist Perc		-1.439 (0.300)	27.77 (1.188)					
Mean DV	0.0372	0.0372	-0.0176	0.0372	51.06	0.111	0.0783	-0.0239
Clusters	390	390	390	390	390	390	390	388
F Stat			546.06					
N	54370	54370	54370	54370	54370	54360	54323	54068
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table includes OLS, reduced form (“RF”), first stage (“FS”), and IV regressions. An observation is a student-subject-year. “Years Waiting” is the deputy teacher’s years without formal employment and “Waitlist Perc” is the imputed waitlist position, normalized by the risk set size to be in percentiles from 0 to 1. Both variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. “Score” is the student’s average subject-specific test score during the year. “First Semester” and “Second Semester” are the semester-specific test scores, and “Exam” is the end-of-year exam. This exam is a national exam for 11th graders before 2006 and 12th graders in all years; otherwise it is the school exam. Test results are expressed in student standard deviation units ( $\sigma$ ) or percentiles. Standard errors are clustered by teacher.

**Table 6: Effect of Years without Formal Employment on Students' Post-Secondary Outcomes**

	IV	IV	IV	IV	IV
	Ln University Score	List Length	Admitted	Academic	Selectivity (Admission)
Years Waiting	-0.0192 (0.00407)	1.245 (0.284)	-0.0183 (0.00752)	-0.0126 (0.00557)	-2.166 (0.395)
Deputy	-0.0355 (0.0213)	1.307 (1.227)	-0.0105 (0.0388)	0.0134 (0.0279)	-3.523 (2.032)
Prior Year GPA	0.241 (0.00368)	-3.611 (0.182)	0.293 (0.00405)	0.307 (0.00423)	24.49 (0.246)
Mean DV	9.532	25.23	0.670	0.625	49.37
Clusters	370	370	391	370	370
N	48818	58702	72879	48818	48818
Risk Set	Yes	Yes	Yes	Yes	Yes

Notes: The table includes instrumental variable estimates with (demeaned) imputed waitlist position as the instrument. An observation is a student-subject-year where the outcomes do not vary by subject but the teachers do. “Years Waiting” is the deputy teacher’s years without formal employment and “Waitlist Perc” is the imputed waitlist position, normalized by the risk set size to be in percentiles from 0 to 1. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. “Ln University Score” is the natural log of the student’s university admissions score. “List Length” is the number of institution-programs the student lists on her ordered list for admissions. “Admitted” and “Academic” are whether the student is admitted to any university and an academic university, respectively. The non-academic university option is a technical university. For “Selectivity (Admission)” we calculate the mean university/admissions score for the class of students admitted to each university-program and order university-programs from highest to lowest. The selectivity measure is the percentile of this ordering where 100 is the program whose admits have the highest mean score. Standard errors are clustered by teacher.

**Table 7: Effect of Years without Formal Employment on Districts' Panhellenic Exam Scores**

	OLS	FS	IV	IV	IV	IV	IV	IV
	Score ( $\sigma$ )	Years Waiting	Score ( $\sigma$ )	Score (Percentile)	Ln University Score	List Length	Admitted	Academic
Years Waiting	0.0907 (0.0972)		-0.342 (0.121)	-8.904 (3.000)	-0.120 (0.0481)	1.050 (2.089)	-0.0741 (0.0379)	-0.0727 (0.0468)
Ln Class Size	0.243 (0.163)	-0.0522 (0.231)	0.240 (0.0983)	8.050 (2.027)	0.198 (0.0533)	3.209 (1.674)	0.0230 (0.0357)	0.0913 (0.0392)
Waitlist Perc		1.785 (0.833)						
Per Class	0.0181		-0.0683	-1.7808	-0.0240	0.210	-0.0148	-0.0145
AR LB			-0.7010	-18.0513	-0.2244	-2.5065	-0.1687	-0.1969
AR UB			-0.1548	-4.5081	-0.0189	6.5905	-0.0065	0.0070
F Stat		4.59						
Mean DV	-0.0440	-0.155	-0.0440	49.07	9.518	25.24	0.818	0.610
N	390	390	390	390	390	390	390	390
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table includes OLS, first stage (“FS”), and IV regressions. An observation is a district-year. “Years Waiting” is the deputy teacher’s years without formal employment and “Waitlist Perc” is the waitlist position, normalized by the list length to be in percentiles. Both variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Test score outcomes are measures of student performance on the national twelfth grade exams, in student standard deviation units ( $\sigma$ ) or percentiles (“Percentile”). “Ln University Score” is the natural log of the student’s university admissions score. “List Length” is the number of institution-programs the student lists on her ordered list for admissions. “Admitted” and “Academic” are whether the student is admitted to any university and an academic university, respectively. The non-academic university option is a technical university. “Per Class” indicates the per-class effect, which is the main coefficient divided by 5 for the 5 classes twelfth graders take in tested subjects. “Reg-Yr FE” are region-year fixed effects. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).



**Table 8: Assigned District Characteristics by Waitlist Position**

	Cohort Size	Unemployment Rate	Cohort Size	Unemployment Rate	Cohort Size	Unemployment Rate
Waitlist Percentile	0.645 (5.233)	-0.421 (0.573)	-0.639 (0.912)	-0.324 (0.281)	-1.063 (0.678)	0.000339 (0.00217)
Mean DV	50.14	9.738	50.14	9.738	50.14	9.738
SD of DV	28.76	3.687	28.76	3.687	28.76	3.687
N	1653	1653	1653	1653	1653	1653
Year FE	Yes	Yes	Yes	Yes	No	No
District FE	No	No	Yes	Yes	Yes	Yes
Reg-Yr FE	No	No	No	No	Yes	Yes

Notes: The table shows the relationship between teachers' waitlist position, on the fresh graduates list, and the characteristics of the district the teacher is assigned to. Waitlist position is normalized by the list length to be in percentiles. "Cohort Size" is the average number of 12th grade students at a school in the district, while "Unemployment Rate" is the local unemployment rate. "Reg-Yr FE" are region-year fixed effects where regions typically include several districts. Teachers may appear in the regressions multiple times if they have received multiple assignments during the sample period. Standard errors are clustered by teacher.

**Table 9: Placebo Tests – Future Assignments and Untested Subjects**

	Future	Future	Nontested	Nontested
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score (Percentile)
Years Waiting t	-0.337 (0.0874)	-0.211 (0.214)		
Ln Class Size	0.144 (0.218)	0.268 (0.320)	0.0861 (0.0191)	2.160 (0.501)
Years Waiting t+1		0.0962 (0.132)		
Years Waiting			-0.0158 (0.0861)	0.570 (2.356)
Per Class t	-0.0675	-0.0422		
Per Class t+1		0.0192		
Per Class			-0.0025	0.0904
AR LB	-0.5483		-0.2458	-4.3281
AR UB	-0.1954		0.4187	15.7290
Mean DV	-0.102	-0.102	-0.0173	49.65
Test p-val		0.0031		
N	133	133	281	281
District FE	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes

Notes: The table includes IV regressions that test for placebo effects. The “Future” columns test whether future assignments affect current scores. The sample is teachers in tested subjects and district-years with assignments the following year. The “Nontested” columns test whether teachers in nontested subjects affect scores in tested subjects. The sample in the last two columns is all teachers in subjects that are not included on the twelfth grade exams. An observation is a district-year and the outcome is the mean score on the national twelfth grade exams (in student standard deviation units or percentiles). “Years Waiting” is the deputy teacher’s years without formal employment and the instrument is the waitlist position, in percentiles, of the deputy teachers when they were on the fresh graduates list. Both variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. “Per Class” rows indicate the per-class effect, which is the main coefficient divided by 5 for tested subjects and 6.3 for nontested subjects. “Test p-val” tests whether the coefficients on “Years Waiting t” and “Years Waiting t+1” are equal. “Reg-Yr FE” are region-year fixed effects. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

# **Online Appendix**

## **Human Capital Depreciation and Returns to Experience**

**Michael Dinerstein, Rigissa Megalokonomou, and Constantine Yannelis**

**Table 10: Effect of Years without Formal Employment on Test Scores – Controlling for Experience**

	Individual	Individual	District	District
	Score ( $\sigma$ )	Score (Percentile)	Score ( $\sigma$ )	Score (Percentile)
Years Waiting	-0.189 (0.271)	-5.737 (8.299)	-0.216 (0.0997)	-6.987 (2.561)
Deputy	-0.0373 (0.0791)	-1.316 (2.318)		
Prior Year GPA	0.737 (0.00885)	21.90 (0.225)		
Ln Class Size			0.154 (0.0834)	5.925 (1.606)
Per Class			-0.0433	-1.3973
AR LB			-0.4572	-13.3235
AR UB			-0.0467	-3.1845
Mean DV	0.0372	51.06	-0.0440	49.07
Clusters	390	390		
N	54370	54370	390	390
District FE			Yes	Yes
Reg-Yr FE			Yes	Yes
Risk Set	Incl Exp	Incl Exp	Incl Exp	Incl Exp

Notes: This table shows IV regressions of test score outcomes on years waiting without formal employment, while controlling for experience. “Years Waiting” is the deputy teacher’s years without formal employment and the instrument is the waitlist position. Both variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year and with the same prior years of experience. In the “Individual” columns, the sample is the set of students and teachers in our micro data set, with a student-subject-year as an observation. In the “District” columns, the sample is all school districts in Greece, with a district-year as an observation. Outcomes are subject-specific test scores (expressed in student standard deviation units or percentiles) in the “Individual” model and Panhellenic test scores in the “District” model. In the “Individual” model, standard errors are clustered by teacher. “Per Class” indicates the per-class effect in the district model, which is the main coefficient divided by 5 for the 5 classes twelfth graders take in tested subjects. “Reg-Yr FE” are region-year fixed effects. District-model standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

# A Teacher Assignment Process Appendix

## A.A Types of Teacher Assignments in Greece

This section provides additional institutional details to complement the discussion in Section I. The education system in Greece appoints teachers to either permanent or temporary positions. Permanently appointed teachers (“permanent teachers”) are considered to be civil servants and once hired they enjoy job security. Every year they have the option to remain at their previous school. Teachers appointed to temporary positions (“temporary teachers”) are employed on a contract basis for no more than ten months. Each year they must re-apply through a centralized assignment system for a new short-term appointment. Even if they receive an assignment in the following year, it will almost certainly be at a different school. These temporary teachers, formally called substitute or deputy teachers, can be either *full-time*, teaching 16-23 classroom hours per week at a standardized salary,<sup>18</sup> or *hourly*, teaching up to 4 hours per week at a standardized hourly rate.

The fraction of teachers in temporary assignments has grown considerably over the last two decades. Between 2011 and 2015, there was a 35% increase in the number of deputy teachers who were employed by schools, such that now 15-20% of the teacher workforce are on temporary contracts (OECD, 2018). This share varies by district. Temporary teachers might be the minority in schools in affluent urban neighborhoods, but often dominate in small and remote areas, especially in the islands (OECD, 2018).

There are at least two reasons behind this trend. First, budgetary pressures since the 2008 financial crisis have increased the use of temporary staff to cover teaching needs. As civil servants, permanent teachers count as a long-term liability to the national budget. In an attempt to reduce these committed expenditures, the European Commission agreed that European structural funds could be used to cover the salaries of temporary teaching staff in Greece and other European countries (OECD, 2018). In practice, these expenditures do not represent salaries, but payments for educational services.<sup>19</sup> Second, since 2009 Greece has had a hiring freeze

---

<sup>18</sup>In practice, nearly all full-time deputy teachers teach the maximum 23 hours per week, where hours refer to classroom instruction time. Less commonly, a full-time deputy teacher could agree to work between 5 and 15 hours per week. These teachers get monthly prorated payments. Full-time permanent teachers with fewer than 6 years of experience teach 23-24 hours per week while more experienced teachers cover 20-21.

<sup>19</sup>Thus, temporary teachers do not get paid during the summer, unlike permanent teachers.

for permanent staff. As hiring new temporary teaching staff has become the only way to cover teaching needs in schools, an increasing number of teachers have been hired on a contract basis.

## **A.B How the System Determines First Assignments from Waitlists**

University graduates with a degree in education prior to 2011 were entitled to a teaching position in a Greek public school, and are assigned via waitlists. For each subject, there are two main types of waitlists for teachers depending on teachers' seniority.<sup>20</sup> The first list is for fresh university graduates with no prior teaching experience. Each year fresh graduates are added to the ends of the waiting lists according to their exact date of degree conferral. If graduates share the same date of degree conferral, ties are broken in favor of the teacher with the higher university grades. Unlike other higher education systems where a university's graduates receive their degrees on the same day, in Greece degree conferral occurs once the pivotal course's grade is entered. The pivotal course is often a student-teaching assignment or involves a written thesis or oral defense.

Once a teacher rises to the top of the list, a position is offered to this teacher the next time it is requested by a school. The lists do not distinguish geographically, so the offered position may be anywhere in the country. The teacher has a week to file the requisite paperwork to accept the position. Occasionally, the position will involve teaching at multiple schools in the same district. The time waiting for an assignment depends on the teacher's waitlist position, the length of the list, and the number of openings. We provide more descriptive statistics in Section II, but typical wait times during our sample period were several years (Tsakloglou and Cholezas, 2005). In recent years, the supply of teachers has outpaced demand, as the prospect of eventually receiving a permanent teaching position with a high salary and job security may overcome the need to wait for the position.<sup>21</sup> Furthermore, many university students choose

---

<sup>20</sup>To be eligible for the waitlists, the following conditions must be met: (a) the applicants must be either Greek or from North-Epirus or Greeks from Constantinople/Istanbul and from the islands of Imvros and Tenedos (Law No. 3832 / 1958) or European Union citizens (Law No. 2431/1996), (b) male applicants must present a military certificate that shows that they have served their compulsory military service or a certificate that shows that the applicant has a military exemption, and (c) expatriates from Cyprus, Egypt, Turkey and North-Epirus must submit a birth certificate and a certificate to the Ministry certifying that they are Greeks. There is no age restriction.

<sup>21</sup>See [Kathimerini](#) for a discussion. Stylianidou et al. (2004) discusses the relative attractiveness of teaching to other options.

teaching degrees without expecting such a long wait. In Appendix Table A.8, we show summary statistics from our online survey of 200 current and former teachers. Just 34% of teachers were aware of the centralized assignment system at the point they were deciding on a profession; but once they entered the system, 76% report understanding the assignment process and 73% knew their exact waitlist position. The increasingly long wait times have led to large protests and politicians have recently proposed changing the system.<sup>22</sup>

### **A.C Accumulating Credits for Subsequent Assignments**

After teachers complete their first temporary assignment, via the fresh graduates list, they enter the second subject-specific list, which consists of teachers who have some prior teaching experience or have taken a written assessment (ASEP). Position in the experienced waitlist depends on their teaching credits, which teachers collect based on their prior experience, score in ASEP, and other factors. Once teachers rise to the top of this list, they get assigned a temporary teaching position. When they complete it, they earn additional teaching credits and re-enter the experienced teachers waitlist for the next temporary assignment. This is (like the initial assignment) a contract-based employment for up to 10 months. We plot histograms of the 1995 degree cohort's wait times for first and second assignments in Appendix Figure A.2. While teachers have long waits for their first assignment, second assignments typically occur one year later with some teachers waiting two or three years.

Teachers earn one teaching credit for each month of prior teaching in a school that is located in an urban area, but they earn two teaching credits per month of prior teaching in remote areas and islands. They also collect credits based on their marital status and the number of children that they have. Job performance for prior teaching does not alter credits.

The ASEP examination system was introduced in 1997 (Law No. 2525/1997) to guarantee permanent positions to teachers that scored the highest on assessments that tested subject-specific and general pedagogic knowledge (Stylianidou et al., 2004). The ASEP examination took place every two years until 2008, the last time it was offered (OECD, 2018). For the sample period, ASEP examinations took place in 2002, 2004, 2006, and 2008 and the weights assigned to ASEP performance and waitlist-related teaching credits as well as academic cre-

---

<sup>22</sup>See [Euronews](#) for a discussion.

dentials changed annually.

Credits determine the order for temporary assignments and any potential permanent assignments. In more recent years, budget constraints have reduced the number of permanent positions, with the 2008 – 2020 hiring freeze eliminating new permanent positions entirely. In 2020, around 10,500 new permanent positions were announced to be filled in 2021 and 2022.

## **A.D Prospective Teachers' Actions**

As part of this process, prospective teachers with university degrees in education have several decisions. While they wait for an assignment, teachers may find alternate ways to generate income.<sup>23</sup> Importantly, these activities may not include taking a full-time job, which would remove a teacher's public school teaching eligibility.<sup>24</sup> If the prospective teacher takes full-time formal employment or notifies the government that she no longer wants to be considered for temporary public school teaching positions, then she will no longer appear on the lists. Finally, if offered a position the teacher finds unattractive or unacceptable, a teacher may reject it. Rejection, however, is quite costly. The teacher is placed at the end of the waitlist and becomes ineligible for any assignments in the following two years. Rejections tend to be rare, but if they were more common and selective they might pose identification challenges. Lists are released publicly, and thus there is little to no scope for manipulation or changing one's order once assigned.

---

<sup>23</sup>Waiting teachers are not eligible for unemployment insurance.

<sup>24</sup>This restriction prevents teachers from working at private schools, though the Greek private education sector is small at 7% enrollment share.



## B Data Appendix

### B.A Waitlists

#### B.A.1 Waitlist Types

The first component of the dataset is teachers' waitlists from Greece (Greece Ministry of Education and Religious Affairs, 2003-2012). These waitlists rank teachers who are waiting for a school assignment as a deputy teacher (*thesi anapliroti*). They can be found on the website of the Ministry of Education under [e-aitisi.sch.gr](http://e-aitisi.sch.gr), for data starting in the 2003-04 school year. On the Ministry of Education website, there is a list of several waitlists. Two types of waitlists are relevant for our empirical analysis. These are:

1. Main lists for experienced high school deputy teachers
2. Main lists for inexperienced high school deputy teachers

The Greek translation for each one of those is "*Pinakes katataxis anapliroton deuterovathmias ekpaideusis*" and "*Pinakes katataxis anapliroton deuterovathmias ekpaideusis midenikis proipiresias*," respectively.

#### B.A.2 Waitlist Content

There are up to fifty subject waitlists. Categorizations are based on the subjects that teachers teach. Teachers obtain these specializations during their university undergraduate studies and they have to report their specialization when they enter these lists. For example, PE01 teachers specialize in teaching religion studies, PE02 teachers specialize in teaching Greek language, history and other literature subjects, PE03 teachers teach mathematics, etc. Teachers can only teach subjects that belong to their categorization. In the main analysis, we restrict attention to five subject categorizations that are tested on the national exams (Panhellenic). These are: PE02, PE03, PE04, PE09, and PE19. Subject PE02 is for Greek/History/Ancient Greek teachers (*filologoi*), subject PE03 is for mathematicians (including Algebra, Statistics, and Geometry), PE04 is for science teachers, subject PE09 is for Economics teachers, and subject PE19 is for Computer Science teachers. For PE04, we include four subcategorizations: PE04.01 (physics),

PE04.02 (chemistry), PE04.04 (biology), and PE04.05 (geology). Usually, there is no PE04.03 list.

In each waitlist, we have a list of teachers with information on the teacher's identity (first name, last name, father's name, mother's name, date of birth), the teacher's degree (degree mark, degree date, subject of specialization), whether the teacher has received any other specialized training (Braille, foreign languages, etc.), and a teachers' past experiences, such as whether obligatory military conscription was completed, that affect eligibility. The waitlists also contain information on *moria*, which are points that teachers collect in order to obtain future assignments. *Moria* can be collected through various ways: academic qualifications, professional experience, social criteria, and other criteria like knowing foreign languages. We also observe *moria* experience, which refers to the number of *moria* credits accrued through professional teaching at a *moria*-eligible school after initial assignments.<sup>25</sup>

### **B.A.3 Ranking Teachers**

Waitlist observations are ranked by sum of *moria* on the waitlists. This variable perfectly predicts rank on the waitlist. The only waitlist for which the sum of collected *moria* does not predict a teacher's position on the list is the main waitlist for inexperienced teachers. This is the waitlist for fresh graduates, who are ranked based on the following lexicographic ordering: (oldest) degree year, (oldest) degree month, (oldest) degree day, (highest) degree mark. Due to bureaucratic hurdles and subject-specific exam days, the degree date offers enough variation that tie-breaking with the degree mark is relatively uncommon.

Regulations for teacher assignments are governed by law 1268/1982 (25, par.12) of the Constitution, according to which: "A student is automatically announced a degree holder (thereby ceases to have a student status) following the end of the exam period during which they fulfilled the requirements of their degree completion. According to Law 1268/1982 (clause 25, paragraph 12) and the decision of the Council of State (Decision 366/1994), as well as the ensuing explanation on the relevant document from the Ministry of Education (17-5-2004, 5/45340/B3), the date of degree conferral of the degree holder is the date of

---

<sup>25</sup>By definition, this information is only available for the experienced teacher lists; inexperienced teachers have yet to accrue experience.

announcement of the grade of the last exam by the member of teaching faculty.”

#### **B.A.4 Eligibility**

For someone to be included in these waiting lists, the following conditions must be met: a) the applicants should be either Greek or from North-Epirus or ethnic Greeks from Constantinople and from the islands of Imbros and Tenedos (Law No. 3832 / 1958) or European Union citizens (Law No. 2431/1996); b) male applicants should present a military certificate that shows that they have served their compulsory military service or a certificate that shows that the applicant has a military exemption; and c) expatriates from Cyprus, Egypt, Turkey and North-Epirus should submit a birth certificate and a certificate to the Ministry certifying that they are Greeks. There is no age restriction.

#### **B.B Assignments**

The second component of our dataset is assignments (Greece Ministry of Education and Religious Affairs, 2004-2011a). Assignment data include information on teacher identity (first name, last name, father’s name and, sometimes, mother’s name), the taught subject (PE02, PE03, etc.), and the teacher’s assignment unit. The assignment is typically at the district level and is given by a letter and a region, e.g., A Evrou, B Evrou, A Artas, B Artas, etc. The prefix/letter refers to the regional office of the school authority of the relevant region. For a minority of teachers, the assignment data includes the specific school the teacher was assigned to.

#### **B.C Test Scores**

We have test score data from the Ministry of Education (Greece Ministry of Education and Religious Affairs, 2003-2011b). For the national sample, we have the composite score built from the following subjects: Greek Language, History, Mathematics, Physics, Biology, special modules, and Economics. The test score data is at the individual level, which we aggregate to the school- or district-level in our district-level model.

## **B.D Individual Level Data**

There are 22 public schools for which we have individual test scores and other outcomes for the 10th, 11th, and 12th grade (Greece High School Archives, 2001-2011). These schools maintained an electronic archive, with data at the student-teacher-class level, which we then hand-collected. The student-level data are available for the students of all three grades for these 22 schools. The teacher-level data provides us with a teacher identifier for each grade, class, year, school and subject combination. The years available varies by school.

## C Teacher Survey

We conducted a survey of current and former teachers between December, 2019 and January, 2020 (Dinerstein et al., 2019-2020). The survey was conducted using Qualtrics, and the sample was drawn using advertisements in Facebook groups for Greek teachers, as well as internet forums for Greek teachers. The survey asked respondents about their perceptions and experiences with the waitlist system and process, as well as the activities they engaged in while waiting and once assigned. Respondents were offered a small Amazon gift card for participation, though take-up of the gift card was very low.

We attach the survey, plus an English translation, to this appendix. Appendix Table A.8 shows summary statistics for selected survey variables. Approximately three-quarters of teachers are aware of their waitlist positions and understand the assignment process, but just one-third of teachers understood the assignment process when they decided to become a teacher. Only 12% of teachers ever rejected an assignment. Appendix Figure A.10 shows the distribution of the number of years on waitlists.

The survey examined what teachers did while on the waitlists, and if they continued any activities once assigned. Appendix Table A.1 shows the activities in which individuals participated while waiting. Approximately one-half of teachers gave private lessons. Nearly 40% of teachers worked in a non-education position while 19% worked in the education sector but in a non-teaching position. One-third of teachers continued with studies and 16% started a family. Many teachers reported engaging in multiple activities while waiting.

Appendix Table A.2 explores jobs teachers engaged in while working as a deputy teacher, and whether these activities had started when the teacher was on the waitlist. The top-row shows that around 30% of teachers engage in part-time work, while teaching in the public system. Among these teachers, the most common form of part-time work is offering private lessons, though small fractions of teachers report working in other industries. The fraction of teachers continuing an activity they started while waiting is quite similar. Conditional on continued activities, private lessons is by far the most common.

Appendix Table A.3 regresses whether a teacher participated in an activity or continued part time work on a teacher's years waiting for her first assignment. In the left column the

dependent variable is an indicator variable for whether the teacher participated in an extra activity, and in the right column the dependent variable is an indicator variable for whether the teacher participated in an extra activity, while working as a public school teacher, that had been started while the teacher was waiting for an assignment. The results indicate that there is no statistically significant relationship between years waiting and either outcome. The point estimates are quite small as are the robust standard errors, so we can rule out even a small correlation between years waiting and participating in an activity or continuing an activity once assigned.

## Default Question Block

**The University of Chicago, Electronic Description and Participation Agreement in the survey**

**Number of the Survey: 19-1614**

**Title of the Survey: Human Capital Depreciation**

**Researchers: Michael Dinerstein, Rigissa Megalokonomou, Constantine Yanellis**

**Description:** We are academics at the University of Chicago and Queensland conducting research on Greek teachers. We want to ask you some questions about your experience, and what you did while waiting as a teacher. This will help us with our research. The survey should only take 10 minutes, your responses are completely anonymous.

**Reward:** If you finish the survey, we will offer you a \$5 Amazon gift card that will be sent to the email address that you will provide us with.

**Contact and Questions:** You can only take the survey once. Please answer all the questions. If you have any question about the survey, please email us: [spyridon.kypraios@chicagobooth.edu](mailto:spyridon.kypraios@chicagobooth.edu)

We really appreciate your input!

Participation Agreement

- Yes, I agree to participate in the survey
- No, I don't agree to participate in the survey

**Section 1: Background Info**

**Note:**

At the end of the multiple-choice questions exist a text in order to write anything you want further

1.1 Did you start from the first day as a permanent or deputy teacher?

- Yes, I have started as permanent teacher from the 1st day
- No, I have started as deputy teacher from the 1st day
- I haven't started to work in a school yet
- None of the above

1.2 The position which you have is a full time, part-time or hourly paying position?

- Full time
- Part time
- Hourly paying
- None of the above

1.3 In what year did you get your teaching degree?

1.4 What is/was your teaching categorization (e.g. ΠΕ70, ΠΕ03, ΠΕ02 etc)?

1.5 If you have started to teach in a school, in what year did you receive your first deputy assignment?



1.6 If you ever left teaching, how many years post-degree did you?

1.7 What was the main reason you became a teacher?

## Section 2: Understanding and Expectations of the System

2.1 If you were in the system with the waitlists. Did you know your waitlist position?

Yes, I knew it

No, I didn't know it

None of the above

2.2 Did you understand the assignment process?

Yes, I did

No, I didn't

None of the above

2.3 How long did you expect to wait for an assignment?

2.4 Were you aware of the waitlist and assignment systems when you chose to become a teacher?

Yes, I was

No, I wasn't

None of the above

2.5 When you were on the waitlists, did ASEP performance affect position?

- Yes, it did
- No, it didn't
- None of the above

### Section 3: Time Spent Waiting

3.1 While waiting for assignments, which of the following did you do? Check all that apply.

- Teaching in private lessons
- Occupation in another field of a none education sector
- Further studies
- Teaching in private school
- Create a family
- Occupied in non-teaching position related to education sector
- Other

3.2 What was your main motivation in choosing how to spend time waiting?

- Economic factors
- Improving skills
- Other

3.3 How did you support yourself financially? Check all that apply.

- Support by the family
- Support by spousal income
- Income from part time job
- Income from occupation in another field of a none education sector
- Other

3.4 Did you take any measures to improve your waitlist position (e.g., take an exam)?

#### Section 4: Attrition and Rejecting Assignments

4.1 Did you consider and/or end up leaving the system?

- Yes, I did
- No, I didn't
- None of the above

4.2 (If you left) How long did you consider leaving the system before you did?

- Before I have started to work in the school
- While you were waiting in the waitlist
- While you were working in school as a deputy teacher
- While you were working in school as a permanent teacher
- None of the above

4.3 What were the relevant considerations in making this decision? (more than one choice)

- Economic reasons
- Family reasons
- Physiological reasons
- Distance from the residency
- Uncertainty
- None of the above

4.4 Did you consider rejecting an assignment?

- Yes, I had rejected

No, I hadn't rejected

None of the above

4.5 What were the relevant considerations in making this decision?

Economic reasons

Family reasons

Physiological reasons

Distance from the residency

Uncertainty

None of the above

## Section 5: Job Characteristics Once Assigned

(Answer for your first assignment)

5.1 Did you choose the district you ended up at?

5.2 Did you choose the school(s) you worked at?

5.3 Did you choose the classes(e.g, A1, A2 or A3 etc) you taught?

5.4 Did you choose the grades (e.g. A', B' or Γ' of Primary school or A', B' or Γ' of High school) you taught?

5.5 What categorization did you teach? Was this subject on the national exam?

5.6 How many schools did you work in?

5.7 What were the main skills you developed while teaching (if any)?

5.8 While you were working in a school, simultaneously you were working somewhere else?

Yes,

Please, define the kind of occupation

No

5.9 Did you continue an unofficial work-- which you had started while you were waiting in the waitlist-- when you took a permanent position in a school?

Yes,

Please, define the kind of occupation

No

## Section 6: Assessment of skills most affected

6.1 While waiting, did you feel like you lost skills?

Yes, I think this

No, I don't think this

None of the above

6.2 If so, which types of skills?

6.3 Did you take any steps to maintain skills? (e.g., attend workshops)

Yes, I did

Please, define the kind of steps

No, I didn't

None of the above

## Section 7: Miscellaneous and Open-ended Questions

7.1 Do you think teaching quality matters for national exam scores?

7.2 Do you consider the end of year tests important?

7.3 What are the incentives to do well as a teacher?

7.4 Anything else you want us to know about the system?

## Είσαι Έλληνας/νίδα Εκπαιδευτικός; Κάνε κλικ ώστε να λάβεις μέρος στην έρευνα κα

### Πανεπιστήμιο του Σικάγο, Ηλεκτρονικό Έντυπο Περιγραφής και Συμμετοχής στην Έρευνα

Αριθμός Έρευνας:

Τίτλος Έρευνας: Η Απαξίωση του Ανθρώπινου Κεφαλαίου

Ερευνητής/τές: Μίκαελ Ντίνεστέϊν Ρήγισσα Μεγαλοκονόμου Κωνσταντίνος Γιαννέλης

**Περιγραφή:** Είμαστε Ακαδημαϊκοί από το Πανεπιστήμιο του Σικάγο (The University of Chicago) και το Πανεπιστήμιο του Κουίνσλαντ (The University of Queensland, Australia). Διεξάγουμε μια έρευνα για τους Έλληνες-Ελληνίδες Καθηγητές/τριες και Δασκάλους/λες. Θα θέλαμε να συμπληρώσετε κάποιες ερωτήσεις σχετικά με την εμπειρία σας και το τι κάνετε ή κάνετε από την αποφοίτηση σας από το Πανεπιστήμιο μέχρι να ξεκινήσετε να εργάζεστε σε σχολείο για πρώτη φορά. Η διάρκεια του ερωτηματολογίου δεν θα ξεπεράσει τα 10 λεπτά, και η συμπλήρωση του ερωτηματολογίου γίνεται **τελείως ανώνυμα**.

**Ανταμοιβή:** Αν πληροίτε τα κριτήρια και συμπληρώσετε την έρευνα θα σας δοθεί δωροεπιταγή Amazon αξίας \$5 (\$5 Amazon gift card). Λαμβάνοντας μέρος σε αυτή την έρευνα, ίσως να μη σας ωφελήσει προσωπικά, αλλά εμείς θα πάρουμε χρήσιμες πληροφορίες που στο μέλλον θα βοηθήσουν άλλους εκπαιδευτικούς. Οι απαντήσεις σας είναι εμπιστευτικές και αφότου σας αποστείλουμε τη δωροεπιταγή, θα διαγράψουμε κάθε προσωπική σας πληροφορία.

**Επικοινωνία & Ερωτήσεις:** Μπορείτε να συμπληρώσετε την έρευνα μόνο μια φορά. Παρακαλώ να απαντήσετε σε όλες τις ερωτήσεις. Αν έχετε κάποια ερώτηση σχετικά με την έρευνα, παρακαλώ να μας στείλετε ηλεκτρονικό μήνυμα (email) : [spyridon.kypraios@chicagobooth.edu](mailto:spyridon.kypraios@chicagobooth.edu)

### Συγκατάθεση Συμμετοχής:

Συμφωνώ να συμμετάσχω στην έρευνα

Δεν συμφωνώ να συμμετάσχω στην έρευνα

## Μέρος 1: Βασικές Πληροφορίες

### Επισήμανση:

Τα πλαίσια στο τέλος των ερωτήσεων επιλογής υπάρχουν ώστε να συμπληρώσετε αν θέλετε κάτι παραπάνω

1.1 Ξεκινήσατε απο την πρώτη μέρα να εργάζεστε ως μόνιμος εκπαιδευτικός ή ως αναπληρωτής;

- Ναι, ξεκίνησα να εργάζομαι ως **ΜΟΝΙΜΟΣ** εκπαιδευτικός απο την 1η μέρα
- Όχι, ξεκίνησα να εργάζομαι ως **ΑΝΑΠΛΗΡΩΤΗΣ** εκπαιδευτικός απο την 1η μέρα
- Δεν έχω ξεκινήσει να εργάζομαι προς το παρόν
- Κανένα από τα παραπάνω

1.2 Η θέση που κατέχετε είναι πλήρης, μερικής ή ωρομίσθιας απασχόλησης;

- Πλήρης απασχόλησης
- Μερικής απασχόλησης
- Ωρομίσθιας απασχόλησης
- Κανένα απο τα παραπάνω

1.3 Ποιο έτος λάβατε το πτυχίο σας από το πανεπιστήμιο;

1.4 Ποιά είναι η ειδίκευση σας όπως προσδιορίζεται από τον κωδικό σας ( για παράδειγμα ΠΕ70, ΠΕ03, ΠΕ02 κλπ);



1.5 Αν έχετε ήδη εργαστεί σε σχολείο, ποιο έτος εργαστήκατε για πρώτη φορά ως αναπληρωτής/τρια καθηγητής/τρια;

1.6 Εάν κάνατε σπουδές στα παιδαγωγικά, αλλά αργότερα αποφασίσατε να αλλάξετε κλάδο, πόσα χρόνια αφού πήρατε το πτυχίο σας πήρατε αυτή την απόφαση;

1.7 Ποιος είναι ο κύριος λόγος που ασχοληθήκατε με το λειτούργημα του δασκάλου/ας;

## Μερος 2: Κατανόηση και προσδοκίες από το σύστημα εκπαίδευσης

2.1 Εάν ήσασταν στο σύστημα με τις λίστες αναμονής για τη πρόσληψη αναπληρωτών, γνωρίζατε την θέση που κατείχατε στον πίνακα κατάταξης/αναμονής για τοποθέτηση σε σχολείο μετά από κάθε περίοδο νέων προσλήψεων;

Ναι, γνώριζα

Όχι, δεν γνώριζα

Κανένα απο τα παραπάνω

2.2 Θεωρείται ότι είχατε κατανοήσει την διαδικασία τοποθέτησης σε σχολείο;

Ναι, την είχα κατανοήσει

Όχι, δεν την είχα κατανοήσει

Κανένα από τα παραπάνω

2.3 Πόσο χρόνο περιμένετε στον πίνακα κατάταξης/αναμονής αναπληρωτών μηδενικής προϋπηρεσίας για να αρχίσετε να εργάζεστε σε σχολείο για πρώτη φορά;

2.4 Θεωρείται ότι ήσασταν ενημερωμένος/η για τη διαδικασία με τις λίστες αναμονής και το σύστημα διορισμού όταν επιλέξατε να γίνεται δάσκαλος/α;

Ναι, ήμουν ενημερωμένος/η

Όχι, δεν ήμουν ενημερωμένος/η

Κανένα από τα παραπάνω

2.5 Όταν ήσασταν στον πίνακα κατάταξης/αναμονής αναπληρωτών μηδενικής προϋπηρεσίας, επηρέασαν τα μόρια του ΑΣΕΠ (εάν είχατε) την θέση σας στον πίνακα;

Ναι, την επηρέασαν

Όχι, δεν την επηρέασαν

Κανένα από τα παραπάνω

### Μέρος 3: Χρόνος Αναμονής

3.1 Όσο χρόνο περιμένετε ώστε να εργαστείτε για πρώτη φορά ως αναπληρωτής/τρια, τι από τα παρακάτω κάνατε; (Μπορείτε να σημειώσετε περισσότερες από μια απαντήσεις)

Διδασκαλία σε κάποιο ιδιωτικό φροντιστήριο

Παραδίδατε ιδιαίτερα μαθήματα

Διευρύνετε τις σπουδές σας με κάποιο νέο τίτλο σπουδών

Ξεκινήσατε μια οικογένεια

Εργαστήκατε σε κάτι άλλο σχετικό με τον τομέα της εκπαίδευσης

Εργαστήκατε σε κάτι άλλο, ΜΗ σχετικό με τον τομέα της εκπαίδευσης

Άλλο (Να το αναγράψετε):

3.2 Ποιο ήταν το βασικό κίνητρο σας ώστε να επιλέξετε αυτό που κάνατε όσο χρόνο περιμένατε;

Οι οικονομικές απολαβές

Διεύρυνση των δεξιοτήτων

Άλλο (Να το αναγράψετε):

3.3 Πως υποστηρίζατε τον εαυτό σας οικονομικά όσο χρόνο περιμένατε; ( Μπορείτε να σημειώσετε περισσότερες από μια απαντήσεις)

Με την υποστήριξη της οικογένειας

Με την υποστήριξη του εισοδήματος του/της συζύγου

Με το μισθό από εργασία μερικής απασχόλησης

Με το μισθό από εργασία μη σχετική με την εκπαίδευση

Άλλο (Να το αναγράψετε):

Κάνατε κάποια ενέργεια ώστε να βελτιώσετε τη θέση σας στη λίστα αναμονής και να αυξήσετε τα μόρια σας (για παράδειγμα: δώσατε κάποιες εξετάσεις) ;

#### Μέρος 4: Αποχώρηση ή απόρριψη θέσης

4.1 Σκεφτήκατε να αλλάξετε κλάδο εργασίας και να βρείτε δουλειά σε άλλο κλάδο;

Ναι, το σκέφτηκα

Όχι, δεν το σκέφτηκα

Κανένα από τα παραπάνω

4.2 Αν ναι, σε ποιο στάδιο της διαδικασίας;

- Πριν αρχίσω να εργάζομαι σε σχολείο
- Όσο χρόνο περίμενα στη λίστα αναμονής
- Κατά τη διάρκεια εργασίας μου σε σχολείο ως αναπληρωτής/τρια
- Κατά τη διάρκεια εργασίας μου σε σχολείο ως μόνιμος/η
- Κανένα από τα παραπάνω

#### 4.3 Ποια στοιχεία σας οδήγησαν σε αυτή την απόφαση;

- Οικονομικοί λόγοι
- Οικογενειακοί λόγοι
- Ψυχολογικοί λόγοι
- Απόσταση από την κατοικία
- Αβεβαιότητα
- Άλλοι λόγοι

#### 4.4 Είχατε απορρίψει κάποια προσφορά θέσης αναπληρωτή/τριας;

- Ναι, είχα απορρίψει
- Όχι, δεν είχα απορρίψει
- Κανένα από τα παραπάνω

#### 4.5 Ποια στοιχεία σας οδήγησαν σε αυτή την απόφαση;

- Οικονομικοί λόγοι
- Οικογενειακοί λόγοι
- Ψυχολογικοί λόγοι
- Απόσταση από την κατοικία
- Αβεβαιότητα
- Άλλοι λόγοι

## Μέρος 5: Χαρακτηριστικά της εργασίας ως καθηγητής/τρια

(Απαντήστε σύμφωνα με τις συνθήκες της πρώτης φοράς που εργασθήκατε)

5.1 Είχατε επιλέξει την περιοχή που σας προτάθηκε η θέση εργασίας;

5.2 Είχατε επιλέξει το/τα σχολείο/α που σας προτάθηκε η θέση εργασίας;

5.3 Είχατε επιλέξει τα τμήματα (A1, A2 ή A3) που σας προτάθηκαν να διδάξετε;

5.4 Είχατε επιλέξει τις τάξεις (για παράδειγμα, Α', Β' ή Γ' Δημοτικού, ή Α', Β' ή Γ' Λυκείου) που σας προτάθηκαν να διδάξετε;

5.5 Τι μάθημα/ματα διδάξατε; Ήταν μάθημα/τα που εξετάζονταν στην διαδικασία των Πανελληνίων εξετάσεων;

5.6 Σε πόσα σχολεία εργαζόσασταν ταυτόχρονα;

5.7 Ποιες ήταν οι σημαντικότερες δεξιότητες που αποκτήσατε όσο διδάσκατε και γιατί;

5.8 Ενώ εργαζόσασταν σε κάποιο σχολείο, εργαζόσασταν ταυτόχρονα και κάπου αλλού;

Ναι,

Παρακαλώ προσδιορίστε το είδος της δουλειάς:

Όχι

5.9 Συνεχίσατε να κάνετε κάποια ανεπίσημη δουλειά--την οποία ξεκινήσατε να κάνετε όσο περιμένατε στη λίστα αναμονής--όταν τελικά διοριστήκατε σε κάποιο σχολείο;

Ναι,

Παρακαλώ προσδιορίστε το είδος της δουλειάς:

Όχι

## Μέρος 6 : Αξιολόγηση των δεξιοτήτων

6.1 Όσο χρόνο περιμένατε στους πίνακες/λίστες αναμονής, θεωρείται ότι χάσατε κάποιες από τις δεξιότητες που είχατε αποκτήσει;

Ναι, θεωρώ ότι έχασα

Όχι, θεωρώ ότι δεν έχασα

Κανένα από τα παραπάνω

6.2 Αν ναι, ποιο τύπο δεξιοτήτων;

6.3 Όσο χρόνο περιμένατε κάνατε κάποια ενέργεια ώστε να διατηρήσετε τις δεξιότητές σας;  
Αν ναι, τι κάνατε;

Ναι, έκανα.

Παρακαλώ συμπληρώστε τι ενέργειες κάνατε :

Όχι, δεν έκανα

Κανένα από τα παραπάνω

## Μέρος 7 : Διάφορες Ερωτήσεις Ανοιχτού Τύπου

7.1 Θεωρείτε ότι η ποιότητα της διδασκαλίας διαδραματίζει ρόλο στα αποτελέσματα των μαθητών στις Πανελλήνιες εξετάσεις;

7.2 Θεωρείτε σημαντικά τα διαγωνίσματα στο τέλος της χρονιάς;

7.3 Ποια ήταν τα κίνητρα που σας ώθησαν να αποδώσετε καλύτερα στον ρόλο σας ως εκπαιδευτικός;

7.4 Οτιδήποτε άλλο που θεωρείται σημαντικό για το σύστημα και θα θέλατε να γνωρίζουμε;

## D Construction of Instrument

In this appendix we provide more details on how we construct our instrument, as it varies across our individual- and district-level specifications. We also assess the robustness of our main results to different ways to construct the instrument.

We start by introducing some notation. As before, teacher  $j$  is in risk set  $m$ . Let  $n_{jt}^D$  be  $j$ 's actual waitlist position on the deputy ( $D$ ) list in year  $t$  and  $n_{jt}^H$  be  $j$ 's actual waitlist position on the hourly ( $H$ ) list in year  $t$ .<sup>26</sup> This waitlist position is an integer, where the teacher with position 1 is the next to be assigned. Let  $\tau(j)$  be teacher  $j$ 's first year on the waitlists. Based on the institutional rules, we know – and verify – that in this year,  $n_{j\tau(j)}^D$  and  $n_{j\tau(j)}^H$  are ordered lexicographically in the date in which  $j$  earned her degree and her degree mark.

We face three research design choices in constructing an instrument: 1) how to scale the waitlist position in predicting time to assignment; 2) how to combine information from the deputy and hourly lists; and 3) what sample to use. As our individual- and district-level analysis relies on different types of data, with different sampling frameworks, we make slightly different research design decisions and then assess robustness. Even though it reverses the paper's order, here we start with the district-level instrument as it is derived from the larger sample.

### D.A Instrument for District Analysis

For the district-level analysis, we can use all teacher assignments in Greece, so we have a large number of recent graduates. This helps in two ways. First, recent graduates have all faced a similar waitlist process (and have had a similar amount of time elapsed since graduating) such that differences in initial waitlist position have a somewhat similar effect across cohorts on years not in formal employment. This relative homogeneity in the first stage allows us to have more precision than if we were combining teachers from cohorts from different decades. Second, we observe these recent graduates' initial waitlist positions in our waitlist data, which start in 2003. This allows us more flexibility on how to scale the waitlist position and in combining deputy and hourly list data. In the district-level analysis, we thus construct our instrument,  $z_j$ , as follows. Let  $\bar{n}_{\tau(j)}^D$  and  $\bar{n}_{\tau(j)}^H$  be the list lengths of the deputy and hourly lists,

---

<sup>26</sup>For simplicity, we describe teachers in a single subject and thus suppress notation for subject.



respectively, in  $j$ 's first year on the waitlists. This length is the highest waitlist position, across all risk sets, on the inexperienced teachers' list in that year. We then convert  $j$ 's waitlist position into a list percentile by dividing by the list length:

$$(29) \quad z_j^D = n_{j\tau(j)}^D / \bar{n}_{\tau(j)}^D$$

and

$$(30) \quad z_j^H = n_{j\tau(j)}^H / \bar{n}_{\tau(j)}^H.$$

Converting position to percentiles is a specific choice of how to scale the waitlist variation. We choose this scaling because the main source of list length heterogeneity is subject, and subjects with longer waitlists tend to assign more teachers annually. Thus, by normalizing by list length, we control for such differences and isolate the useful variation. But because there are other ways to scale the variation, we include robustness checks below.

With these two measures, we combine them into a single instrument by taking the minimum:  $z_j = \min\{z_j^D, z_j^H\}$ . We do this because a teacher may accrue experience from either type of assignment and thus the “better” waitlist position is the one that matters for accruing any experience. We considered other choices such as the mean, the max, and the deputy list value. We show below that our results are robust to these choices.

Finally, in the district-level analysis, we include all teachers but only contract teachers who ever appear on an inexperienced list 2003 or later have non-singleton risk sets. We focus on variation among these teachers in identifying our causal estimates because these are the teachers for whom we can use the waitlist information for scaling and combining hourly and deputy positions, two design choices that yield statistical power. We aggregate across teachers, each with instrument  $z_j$ , to get to the district-level instrument.

## D.B Instrument for Individual Analysis

For the individual-level analysis, our sample of contract teachers is much smaller and dominated by more experienced teachers.<sup>27</sup> We thus run out of statistical power if we restrict our sample to teachers who appear on an inexperienced list 2003 or later, which means that we cannot exploit the waitlist information for scaling and hourly-deputy combining purposes. Instead, we construct our instrument as follows. For each year  $t$ , let  $\check{n}_{jt}$  be teacher  $j$ 's waitlist position on a pseudo-waitlist made up only of the teachers who are on the actual waitlist in year  $t$  (denoted by  $\mathcal{J}_t$ ). As with the actual inexperienced teacher waitlist, we determine  $\check{n}_{jt}$  lexicographically based on date of degree conferral.<sup>28</sup> We then calculate the highest pseudo-waitlist position for each risk set:

$$(31) \quad n_{mt}^* = \max_{m(j)=m} \check{n}_{jt}$$

and form our instrument as:

$$(32) \quad z_{jt} = \check{n}_{jt} / n_{mt}^*.$$

The instrument is thus normalized by the highest waitlist position in a risk set, where the risk set is composed of all teachers in the same degree conferral year-month and subject who are still teaching in year  $t$ .

Several research design choices are worth further comment. First, while we use a “pseudo-waitlist,” it is ordered identically to how the actual waitlist is ordered (up to the inconsequential tie-breaking). Provided our assumption of no selective attrition holds, the pseudo-waitlist position contains the exact same ordinal information as the actual waitlist position. The use of an imputed waitlist position thus primarily matters for (1) scaling purposes and (2) our inability to combine information from deputy and hourly lists. Both of these issues are primarily about combining extra information to generate a more powerful first-stage. But because

---

<sup>27</sup>The small sample is in part driven by our ability to match teachers between datasets. Some schools record only teachers' first name initial, which leads to many cases where we cannot uniquely link the teacher to the waitlist data.

<sup>28</sup>We could also incorporate degree mark to break degree conferral date ties, as the actual waitlists do. Because degree mark could be related to teacher ability, we choose to break ties randomly in our pseudo-waitlist, but results are very similar when we incorporate degree mark.

our individual-level analysis has outcomes at the teacher level, the link between a teacher’s (pseudo-)waitlist position and experience is much stronger than the one in the aggregated district-level analysis. Thus, the missing information is potentially not as important to the individual-level analysis. Second, we still make a specific scaling choice by normalizing by the risk set maximum position. The full set of teachers still teaching in  $t$  includes many whose degrees came well after some of our focal teachers and thus were not relevant for initial assignments. Incorporating these teachers into the list length would therefore just add noise. We can test whether this decision matters in our district-level analysis and will show our results are robust to normalizing by highest waitlist position on the full list or within risk set. We implement each research choice uniformly for inexperienced and experienced teachers.

Before showing the robustness checks, we confirm that ordering teachers by degree conferral date produces a similar assignment process as ordering by actual waitlist position. In Appendix Figure A.1 we take the 2003 mathematics waitlists and show binscatter plots for the year of first assignment. In Panel A we use the actual waitlist position (in the data) for the x-axis and in Panel B we use the order implied by degree conferral date. We see very similar patterns of time until first assignment across the two panels.

## **D.C Robustness to Research Design Choices**

We now explore the robustness of our results to the research design choices. As the district-level analysis uses more waitlist information (the actual list length and the combination of hourly and deputy lists), we conduct most robustness checks using the district-level specification.

We first investigate how the scaling of the waitlist variation affects the results. We specify several variations of our instrument. “Baseline” is our baseline model’s instrument, which normalizes waitlist position by waitlist length and takes the minimum across deputy and hourly lists. “Mean across Lists” and “Max across Lists” make the same normalization but take the mean and maximum, respectively, across deputy and hourly lists. “Deputy Position” also uses the same normalization but takes the value from the deputy list, ignoring the hourly list. These instruments vary how we combine the variation from the two waitlists a teacher may be assigned from.

We show how the instruments affect the student test scores in percentiles (Appendix Table

A.9) and standard deviation units (Appendix Table A.10) where we do not control for teacher prior experience. We then present the results for student test scores in percentiles, where we control for teacher prior experience (Appendix Table A.11). We find that in the first four columns the estimates are very stable across instrument choices.

We then examine how the scaling of the waitlist position affects the estimates. “Risk Set Norm” normalizes the waitlist position by the highest waitlist position in the same risk set. This scaling matches how we construct our individual-analysis instrument and thus demonstrates whether this scaling matters for the results. “Raw Position” uses the actual waitlist position, without any normalization. “Log Position” takes the log of the waitlist position rather than normalizing by list length. We see in the tables that these research design choices make minimal difference for our qualitative conclusions.

Finally, we investigate the importance of sample selection by estimating our individual-level model on deputy teachers only. We present the results in Appendix Table A.12 and find that the estimates are nearly identical. This is unsurprising as the non-deputy teachers contributed no variation in the instrument and thus were valuable for their effect on the controls. We also show that are results are stable to restricting to the small set of teachers for whom we have the most confidence in our matching procedures across data sets (columns 3 and 6).

## E Robustness Checks

In this appendix we provide a variety of robustness checks around our individual- and district-level estimates of the causal effect of a year without formal employment on student outcomes.

### E.A Individual-Level Estimates

#### E.A.1 Controls

In Appendix Table A.13, we present our main results without any lagged GPA controls (column 1). We find similar point estimates to our baseline regression, but with reduced precision. In column 2 we estimate the baseline specification but instead of controlling for lagged GPA, we control for demeaned lagged GPA where the teacher-year mean is removed (column 2). Because we do not observe students' lagged GPA for every teacher in the country, we cannot demean the control by risk set, as would be consistent with our model. But we find that demeaning by a finer level – the teacher-year mean – leaves our main point estimate essentially unchanged.

#### E.A.2 Functional Forms

As discussed in Section III, the economics of education literature sometimes argues that returns to experience are declining at higher levels of experience. Further, test score units do not have a standard conversion rate to teacher human capital measures. We thus offer variations on our main specification where we include log test scores and log years without formal employment. We present the results in Appendix Table A.14. We find strong effects regardless of the functional form. One year waiting leads to a 2.5% drop in students' test scores (column 2), while a 10% increase in time waiting corresponds to a  $0.08\sigma$  effect on students' test scores (column 3). For the log-log specification, we estimate an elasticity of student test scores with respect to years waiting of -0.41 (column 4).

### **E.A.3 Sample**

We argue that within-month variation in degree conferral is orthogonal to teacher type and plotted the distribution in Figure 1. But the distribution is not uniform either, with a peak on the 30th of the month. We also see a peak in within-year degree conferral in July. We confirm that our results are not sensitive to these degree months and days by dropping teachers with degree conferrals on the 30th of the month and then by dropping teachers with degree conferrals in the month of July. We present the results in Appendix Table A.15.

### **E.A.4 Outcomes**

The results are robust to different functional forms of our outcomes. In Appendix Table A.16 we show causal effects on unstandardized test scores, log test scores, and raw university score. We also include several variations in calculating an institution-program's selectivity. In the main analysis, we calculated selectivity based on enrollees' mean university scores. Here we show selectivity based on enrollees' mean national exams scores, which are a different weighting than the university admissions scores. In both cases, these selectivity measures are means across multiple years, including years in our sample. To avoid any concerns of our sample affecting the selectivity measures, we also include selectivity measures derived from 2003 admissions outcomes only.

### **E.A.5 Full-Time versus Part-Time Employment**

The main results include years of part-time work in the experience measure. We can instead define a year of experience as working and accruing above-median work credits. We rerun our main analysis with this alternate definition of years without formal *full-time* employment and present the results in Appendix Table A.17. The estimated effects are very similar to our main results.

### **E.A.6 Standard Errors**

When we demean using risk sets, the mean has a sampling distribution. We do not account for this in our main estimates because risk sets are large enough that sampling variation in

the mean is likely to be second order. Here we incorporate such sampling variation by bootstrapping our estimates. First, we sample from the full sample of teachers, to calculate risk set means. Then we run 500 wild clustered bootstrap iterations, where we sample in the instrumental variable analysis according to the Rademacher distribution and use the same draw for all observations in a cluster and for first stage and second stage residuals. We construct a bootstrapped standard error estimates for the estimates in Tables 5 and 6 and present them in Appendix Table A.18. We find that the sampling error from the risk set means is tiny, as 6 of the 13 estimates have smaller bootstrapped than non-bootstrapped standard errors, and the average ratio of bootstrapped to non-bootstrapped standard errors is 1.01.

## **E.B District-Level Estimates**

### **E.B.1 Controls**

We present estimates that vary our use of controls in Appendix Table A.19. In the first column, we show estimates where we residualize all fixed effects by risk sets. The point estimate is very similar to our main result. The second and third columns add additional (demeaned) district-level controls. The point estimates are similar and move closer to our individual-model estimates.

### **E.B.2 Functional Forms**

As with the individual-level analysis, we vary the functional forms and present the results in Appendix Table A.20.

### **E.B.3 Sample**

In Appendix Table A.21, we present the results dropping teachers with degree conferrals on the 30th of the month and then dropping teachers with degree conferrals in the month of July.

### **E.B.4 Outcomes**

In Appendix Table A.22, we show the causal effects on the other test or selectivity outcomes.

### **E.B.5 Scaling and Weighting**

In the district-level model, we use an aggregation matrix, which included all deputy teachers, regardless of whether we observe them on an inexperienced list in our sample period. For these teachers, we consider them part of their own risk sets so that their years not working do not identify the causal estimates. In Appendix Table A.23, we explore the sensitivity of our results to including these additional deputy teachers. We find similar estimates that are slightly smaller in magnitude to our main estimates. Given all deputy teachers factor into a district's teaching, the attenuation toward zero is consistent with this specification including classical measurement error.

As the aggregation of an individual-level model, we estimate our district-level model weighting by the number of students in each district-year. In Appendix Table A.24, we show how the results change with different weightings. The alternate weightings lead to somewhat more negative point estimates.



## References

- Altonji, Joseph G and Nicolas Williams**, “The Effects of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth,” 1992.
- Alvarez, Fernando E, Katarína Borovičková, and Robert Shimer**, “Decomposing Duration Dependence in a Stopping Time Model,” Technical Report, National Bureau of Economic Research 2016.
- Anderson, Theodore W and Herman Rubin**, “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *The Annals of Mathematical Statistics*, 1949, 20 (1), 46–63.
- Andrews, Isaiah**, “Valid Two-Step Identification-Robust Confidence Sets for GMM,” *Review of Economics and Statistics*, 2018, 100 (2), 337–348.
- Angrist, Joshua D**, “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,” *The American Economic Review*, 1990, pp. 313–336.
- and **Jörn-Steffen Pischke**, *Mostly Harmless Econometrics*, Princeton University Press, 2008.
- Autor, David H, Nicole Maestas, Kathleen J Mullen, and Alexander Strand**, “Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants,” Technical Report, National Bureau of Economic Research 2015.
- Baily, Martin**, “Some Aspects of Optimal Unemployment Insurance,” *Journal of Public Economics*, 1978, 10 (3), 379–402.
- Becker, Gary**, “Investment in Human Capital: A Theoretical Analysis,” *Journal of Political Economy*, 1962, 70 (5), 9–49.
- , “Human Capital: A Theoretical and Empirical Analysis,” *Columbia University Press*, 1964.
- Ben-Porath, Yoram**, “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, 1967, 75 (4, Part 1), 352–365.
- Benhenda, Asma**, “Absence, Substitutability and Productivity: Evidence from Teachers,” 2017.
- Bertrand, Marianne and Antoinette Schoar**, “Managing with Style: The Effect of Managers on Firm Policies,” *Quarterly Journal of Economics*, 2003, 118 (4), 1169–1209.
- Blinder, Alan S and Yoram Weiss**, “Human Capital and Labor Supply: A Synthesis,” *Journal of Political Economy*, 1976, 84 (3), 449–472.
- Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw**, “Female Labor Supply, Human Capital and Welfare Reform,” *Econometrica*, 2016, 84 (5), 1705–1753.
- Card, David, Raj Chetty, and Andrea Weber**, “Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market,” *The Quarterly Journal of Economics*, 2007, 122 (4), 1511–1560.
- Centeno, Mário and Álvaro A Novo**, “Excess Worker Turnover and Fixed-Term Contracts: Causal Evidence in a Two-Tier System,” *Labour Economics*, 2012, 19 (3), 320–328.
- Chetty, Raj**, “A General Formula for the Optimal Level of Social Insurance,” *Journal of Public Economics*, 2006, 90 (10), 2351–2356.
- , **John Friedman, and Jonah Rockoff**, “Measuring the Impact of Teachers I: Evaluating Bias in Teacher Value-Added Estimates,” *American Economic Review*, 2014, 104 (9), 2633–79.
- Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis**, “Deputy Teacher

- Survey,” 2019-2020.
- , – , and – , “Human Capital Depreciation and Returns to Experience,” Data Repository, openicpsr-165721 2022.
- Edin, Per-Anders and Magnus Gustavsson**, “Time out of Work and Skill Depreciation,” *ILR Review*, 2008, 61 (2), 163–180.
- Farber, Henry S, Dan Silverman, and Till von Wachter**, “Factors Determining Callbacks to Job Applications by the Unemployed: An Audit Study,” *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2017, 3 (3), 168–201.
- Frisch, Ragnar and Frederick V Waugh**, “Partial Time Regressions as Compared with Individual Trends,” *Econometrica: Journal of the Econometric Society*, 1933, pp. 387–401.
- Giles, David EA**, “Instrumental Variables Regressions involving Seasonal Data,” *Economics Letters*, 1984, 14 (4), 339–343.
- Giroud, Xavier and Holger Mueller**, “Firm Leverage, Consumer Demand and Employment Losses During the Great Recession,” *Quarterly Journal of Economics*, 2017, 132 (1), 271–236.
- Greece High School Archives**, “Student and Teacher Data,” 2001-2011.
- Greece Ministry of Education and Religious Affairs**, “Twelfth Grade Student Outcomes,” 2003-2011.
- , “Deputy Waitlists,” 2003-2012.
- , “Deputy Assignments,” 2004-2011.
- Hanushek, Eric A and Steven G Rivkin**, “Generalizations about Using Value-Added Measures of Teacher Quality,” *American Economic Review*, 2010, 100 (2), 267–71.
- Herrmann, Mariesa A and Jonah E Rockoff**, “Worker Absence and Productivity: Evidence from Teaching,” *Journal of Labor Economics*, 2012, 30 (4), 749–782.
- Hoffmann, Florian and Philip Oreopoulos**, “Professor Qualities and Student Achievement,” *The Review of Economics and Statistics*, 2009, 91 (1), 83–92.
- Imai, Susumu and Michael P Keane**, “Intertemporal Labor Supply and Human Capital Accumulation,” *International Economic Review*, 2004, 45 (2), 601–641.
- Jackson, C Kirabo**, “Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers,” *Review of Economics and Statistics*, 2013, 95 (4), 1096–1116.
- and **Elias Bruegmann**, “Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers,” *American Economic Journal: Applied Economics*, 2009, 1 (4), 85–108.
- Jackson, Kirabo, Jonah Rockoff, and Doug Staiger**, “Teacher Effects and Teacher-Related Policies,” *Annual Review of Economics*, 2014, 6.
- Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan**, “Earnings Losses of Displaced Workers,” *American Economic Review*, 1993, pp. 685–709.
- Jarosch, Gregor**, “Searching for Job Security and the Consequences of Job Loss,” *Manuscript, Stanford University*, 2015.
- Kahn, Lisa B**, “The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy,” *Labour Economics*, 2010, 17 (2), 303–316.
- Keane, Michael and Kenneth Wolpin**, “The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment,” *International Economic Review*, 2001, 42 (4).

- Kehoe, Patrick J, Virgiliu Midrigan, and Elena Pastorino**, “Debt Constraints and Employment,” *Journal of Political Economy*, 2019, 127 (4), 1926–1991.
- Kroft, Kory, Fabian Lange, and Matthew Notowidigdo**, “Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment,” *Quarterly Journal of Economics*, 2013, 128 (3), 1123–1167.
- Lavy, Victor**, “Evaluating the Effect of Teachers Group Performance Incentives on Pupil Achievement,” *Journal of Political Economy*, 2002, 110 (6), 1286–1317.
- , “Performance Pay and Teachers’ Effort, Productivity, and Grading Ethics,” *American Economic Review*, 2009, 99 (5), 1979–2011.
- **and Rigissa Megalokonomou**, “Persistency in Teachers’ Grading Bias and Effects on Longer-Term Outcomes: University Admissions Exams and Choice of Field of Study,” *National Bureau of Economic Research*, 2019, (26021).
- **and –** , “Long Term Effects of Teachers: Evidence from a Teacher Value Added Approach,” 2020.
- Ljungqvist, Lars and Thomas J Sargent**, “The European Unemployment Dilemma,” *Journal of Political Economy*, 1998, 106 (3), 514–550.
- Lovell, Michael C**, “Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis,” *Journal of the American Statistical Association*, 1963, 58 (304), 993–1010.
- Manuelli, Rodolfo E and Ananth Seshadri**, “Human Capital and the Wealth of Nations,” *American Economic Review*, 2014, 104 (9), 2736–62.
- Marx, Matt, Deborah Strumsky, and Lee Fleming**, “Mobility, Skills, and the Michigan Non-Compete Experiment,” *Management Science*, 2009, 55 (6), 875–889.
- Neal, Derek**, “Industry-Specific Human Capital: Evidence from Displaced Workers,” *Journal of Labor Economics*, 1995, 13 (4), 653–677.
- OECD**, “Education for a Bright Future in Greece,” *Reviews of National Policies for Education*, OECD. Publishing, Paris., 2018.
- Opper, Isaac M**, “Does Helping John Help Sue? Evidence of Spillovers in Education,” *American Economic Review*, 2019, 109 (3), 1080–1115.
- Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz**, “The Short-and Long-Term Career Effects of Graduating in a Recession,” *American Economic Journal: Applied Economics*, 2012, 4 (1), 1–29.
- Papay, John P and Matthew A Kraft**, “Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Improvement,” *Journal of Public Economics*, 2015, 130, 105–119.
- Rice, Jennifer King**, “Learning from Experience? Evidence on the Impact and Distribution of Teacher Experience and the Implications for Teacher Policy,” *Education Finance and Policy*, 2013, 8 (3), 332–348.
- Rockoff, Jonah**, “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data,” *American Economic Review*, 2004, 94 (2), 247–252.
- Rockoff, Jonah E, Douglas O Staiger, Thomas J Kane, and Eric S Taylor**, “Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools,” *American Economic Review*, 2012, 102 (7), 3184–3213.
- Rosen, Sherwin**, “A Theory of Life Earnings,” *Journal of Political Economy*, 1976, 84 (4, Part

2), S45–S67.

**Schmieder, Johannes F, Till von Wachter, and Stefan Bender**, “The Effect of Unemployment Benefits and Nonemployment Durations on Wages,” *American Economic Review*, 2016, 106 (3), 739–77.

**Shimer, Robert and Ivan Werning**, “On the Optimal Timing of Benefits with Heterogeneous Workers and Human Capital Depreciation,” Technical Report, National Bureau of Economic Research 2006.

**Stylianidou, Fani, George Bagakis, and Dimitris Stamovlasis**, “Attracting, Developing and Retaining Effective Teachers, Country Background Report for Greece,” *Report, OECD Activity*, 2004.

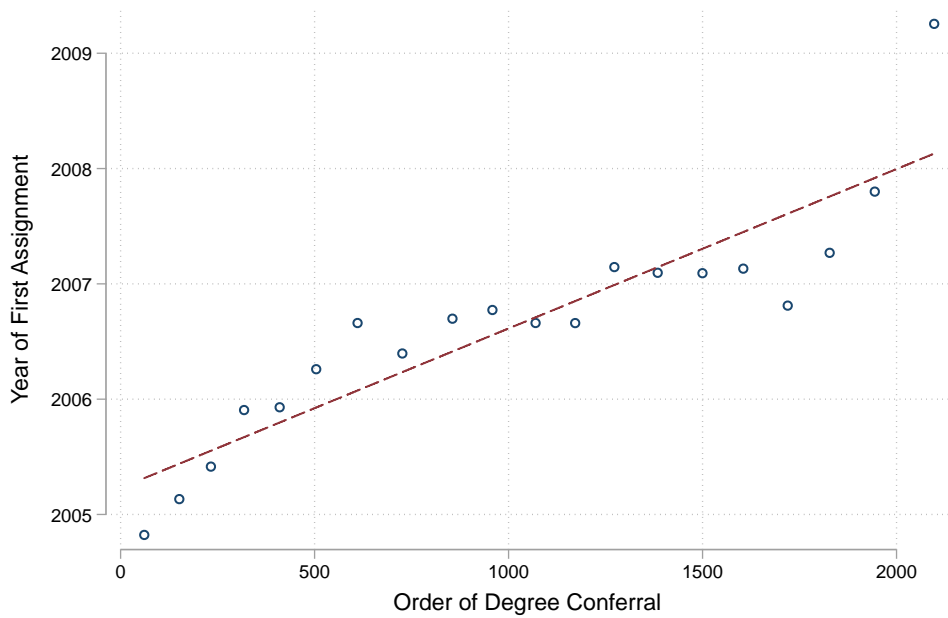
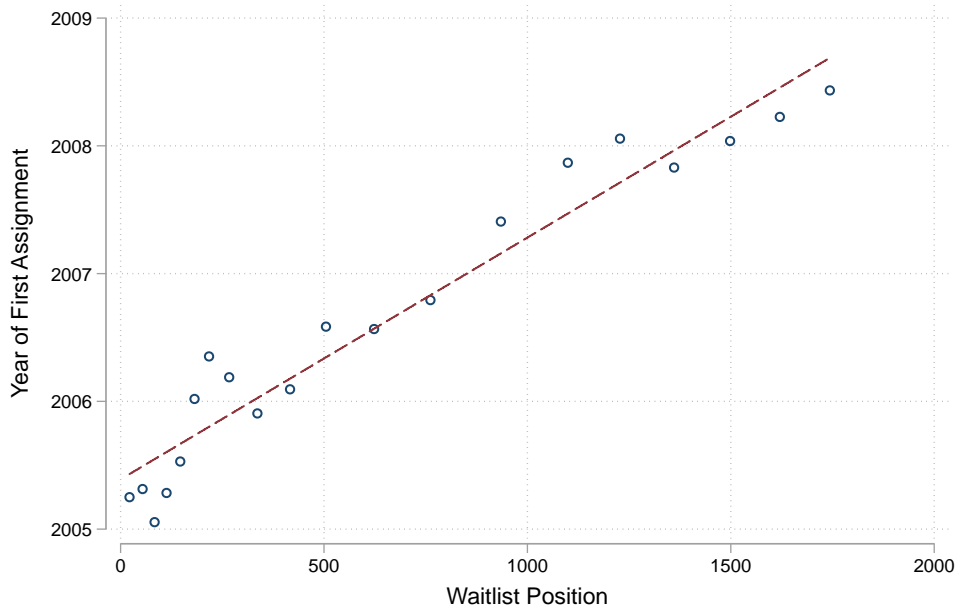
**Taylor, Eric S and John H Tyler**, “The Effect of Evaluation on Teacher Performance,” *American Economic Review*, 2012, 102 (7), 3628–51.

**Tsakloglou, Panos and Ioannis Cholezas**, “Education and Inequality in Greece,” *Discussion Paper No. 1582*, 2005.

**Wiswall, Matthew**, “The Dynamics of Teacher Quality,” *Journal of Public Economics*, 2013, 100, 61–78.

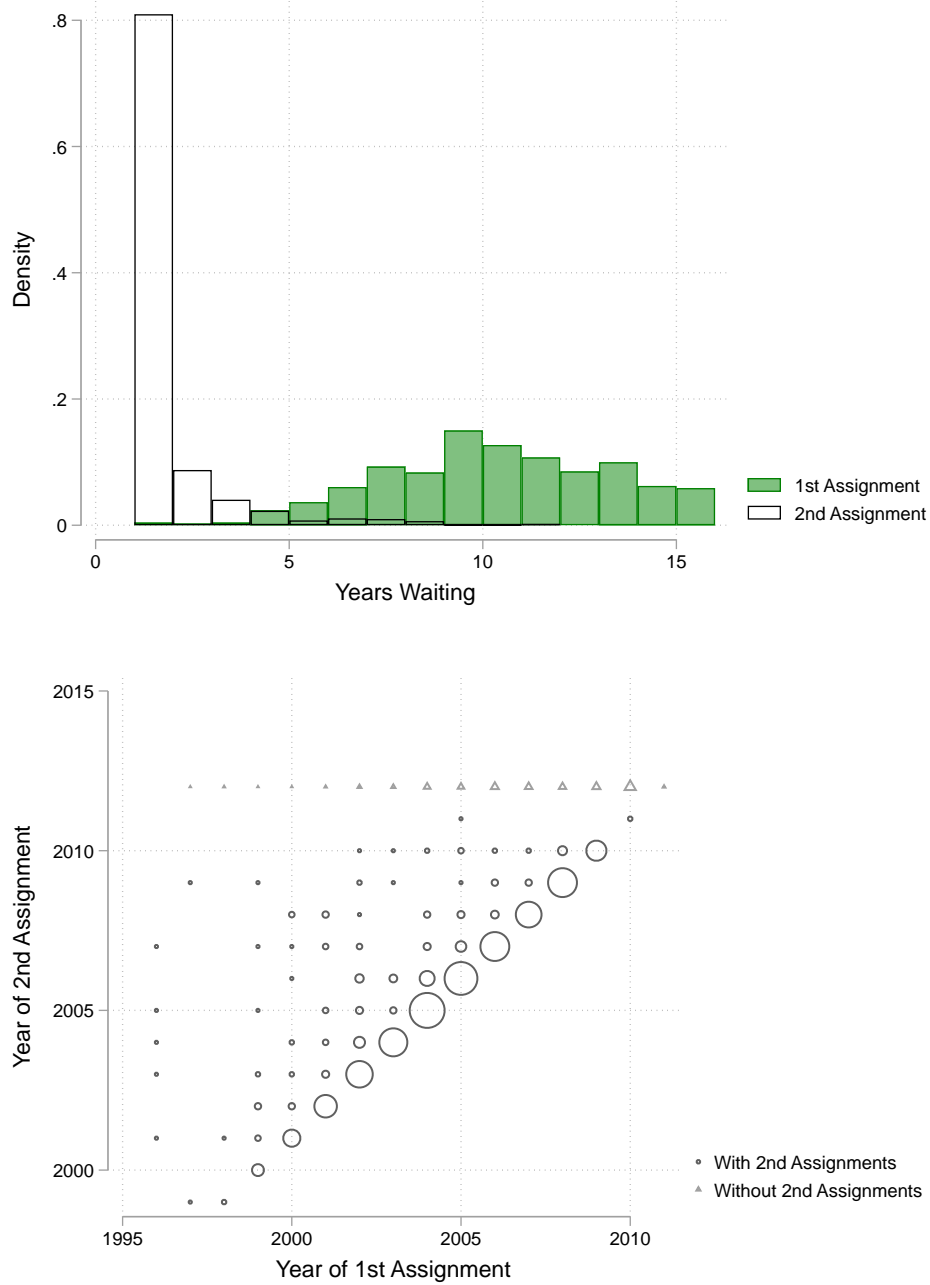
**Zivin, Joshua Graff and Matthew Neidell**, “The Impact of Pollution on Worker Productivity,” *American Economic Review*, 2012, 102 (7), 3652–3673.

**Figure A.1: Assignment Year by Waitlist Position**



Notes: The top figure is a binscatter showing a teacher's 2003 waitlist position on the mathematics list (x-axis) and the year of first assignment (y-axis). The bottom figure is a binscatter with the same sample but changing the x-axis to the imputed order of degree conferral we use in the individual-level empirical model.

**Figure A.2: Timing of First and Second Assignments**



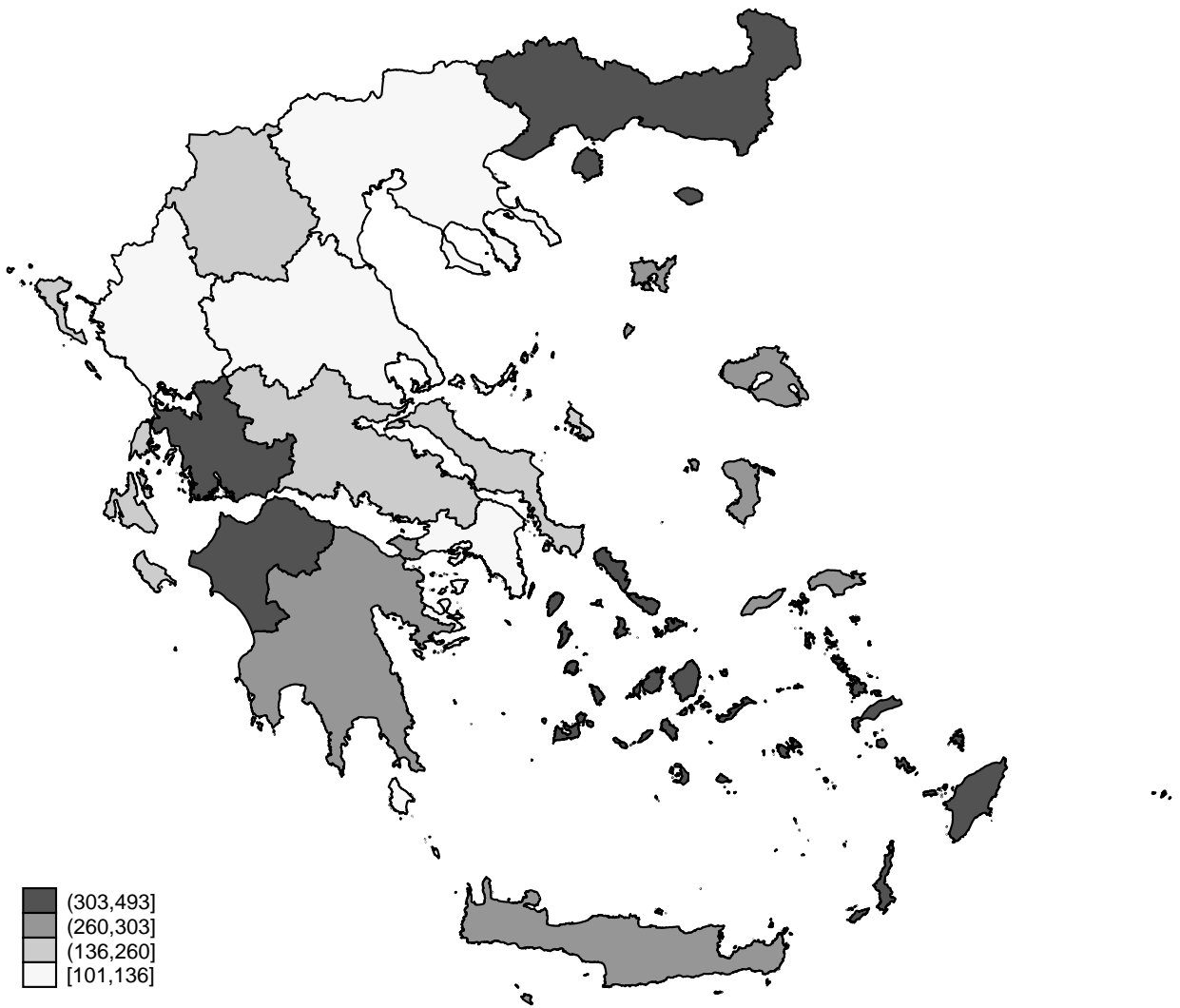
Notes: The figures show the relationship between first and second assignments for teachers in the 1995 degree conferral cohort. The top figure shows the distribution of wait times prior to first (solid) and second (clear) assignments. The bottom figure is a scatterplot with the year of the first assignment on the x-axis and the year of the second assignment on the y-axis. The size of the dots are proportional to the number of teachers in each group. The triangles at the top of the y-axis indicate teachers who never received a second assignment.

**Figure A.3: Counties Containing Schools with Student Level Test Score Data**



Notes: This figure shows the counties with schools for which we have individual student test score data.

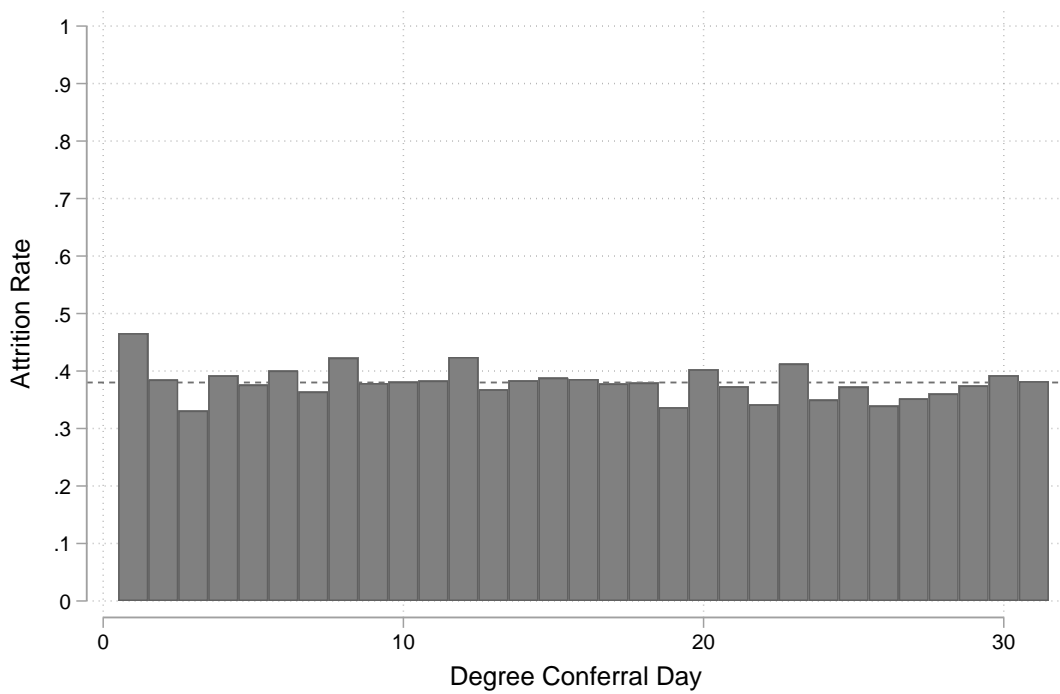
Figure A.4: Number of Deputy High-School Teachers Assigned, 2006



Notes: This figure shows the number of deputy high-school teachers assigned in each Greek region in 2006.

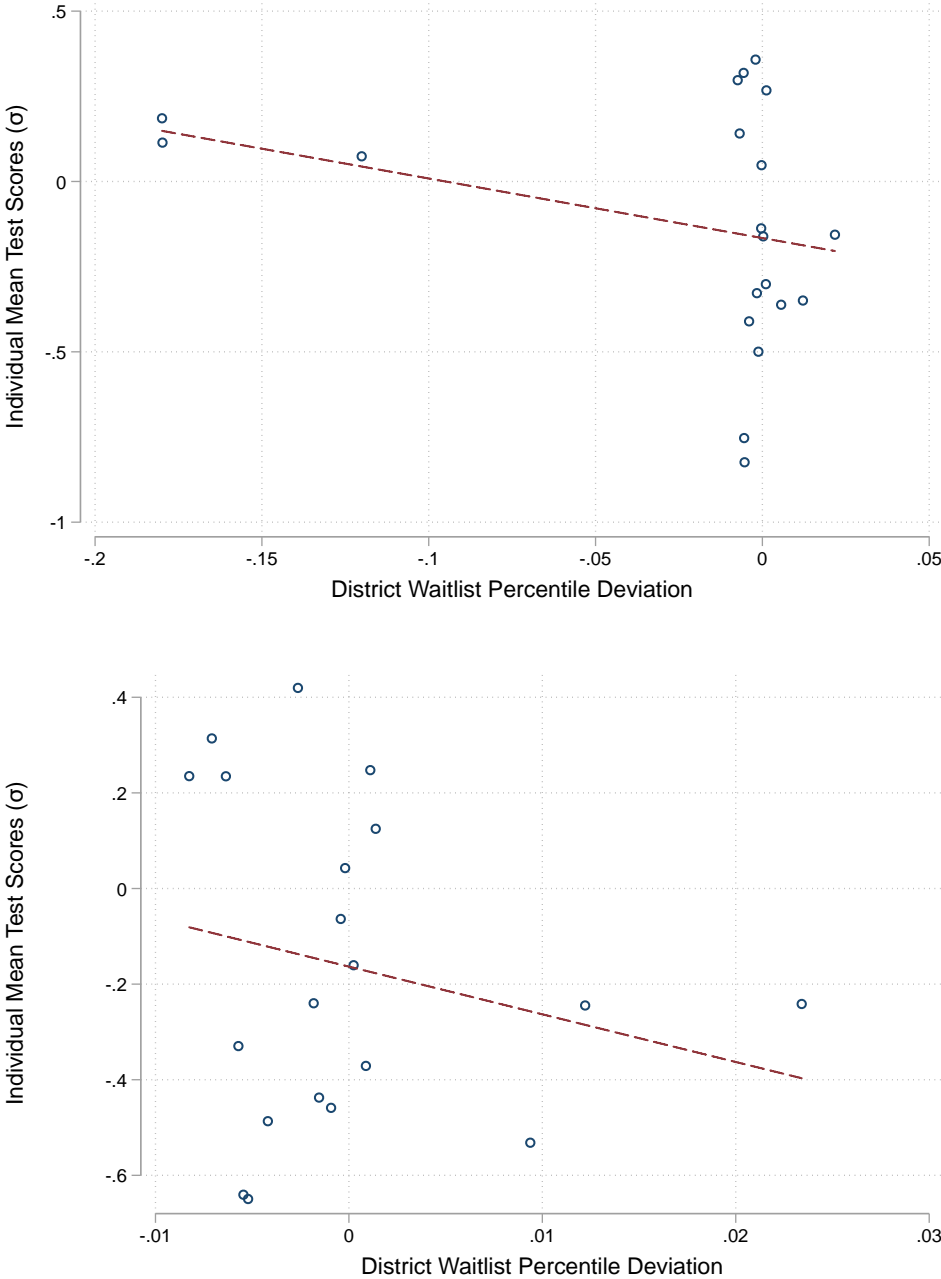


**Figure A.5: Attrition and Degree Conferral Day**



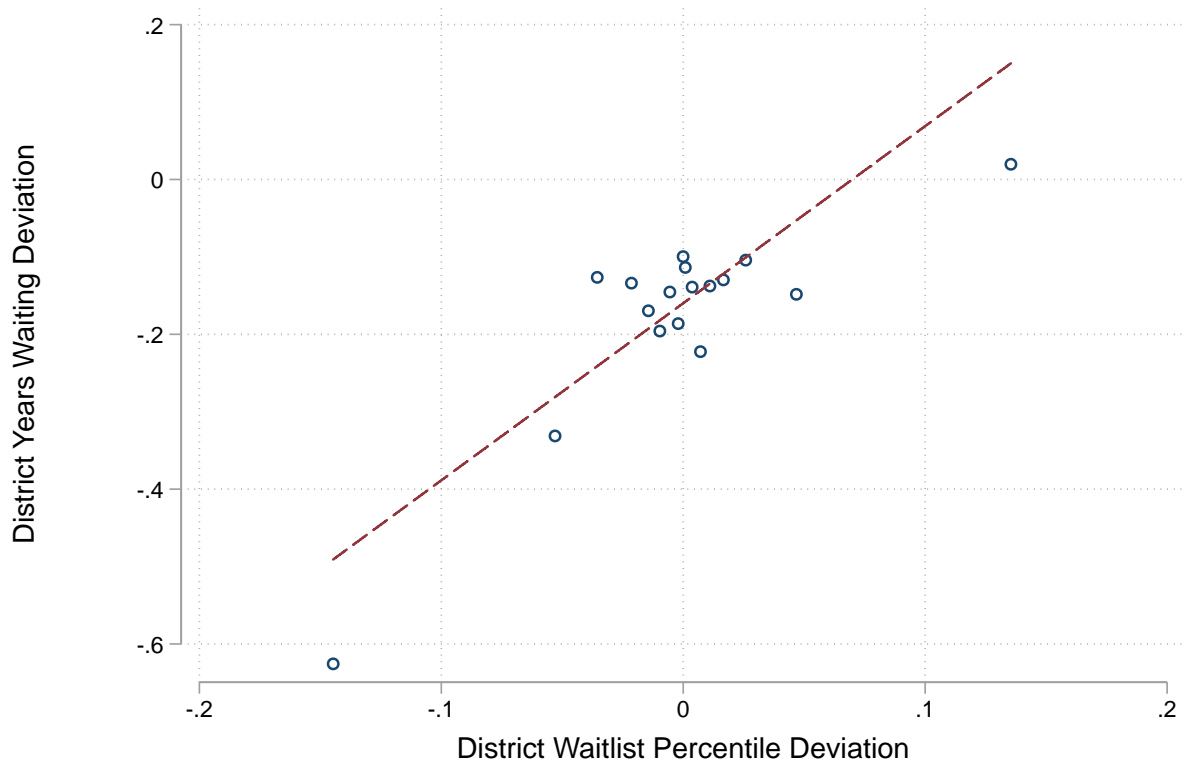
Notes: This figure shows attrition rates by the day of the month in which teachers' university degrees were conferred. This day of the month variation is our within risk-set timing variation that identifies our causal effects. Attrition is defined as leaving the waitlists before the end of our sample period and without ever having accrued experience.

**Figure A.6: Test Scores ( $\sigma$ ) and Waitlist Position – Individual Level**



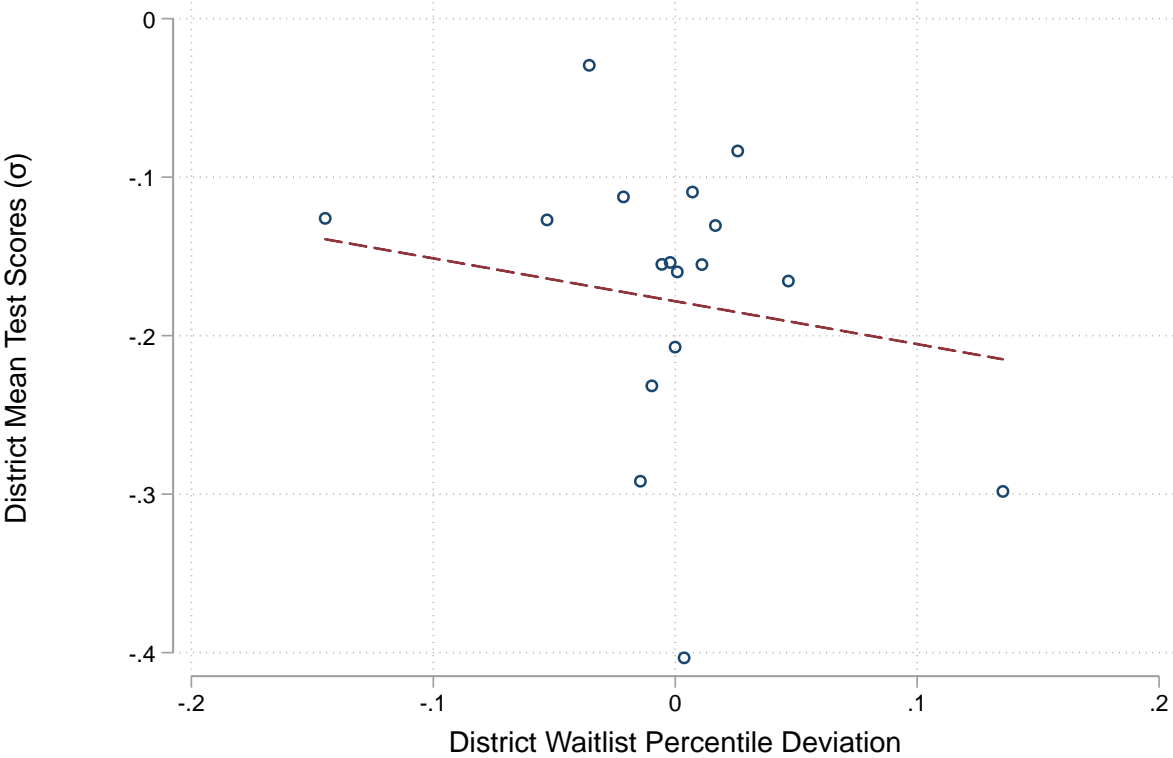
Notes: The binscatter figures show the relationship, at the teacher-year level, between demeaned waitlist percentile and student test scores (in student standard deviation units) where the demeaning is done by risk set-year. A risk set is a degree conferral year-month and subject combination. The binscatter also controls for student lagged grade-point-average. The top figure shows the full deputy teacher sample while the bottom figure zooms in on the middle of the distribution by excluding the points with instrument less than -0.05.

**Figure A.7: Years without Formal Employment and Waitlist Position – District Level**



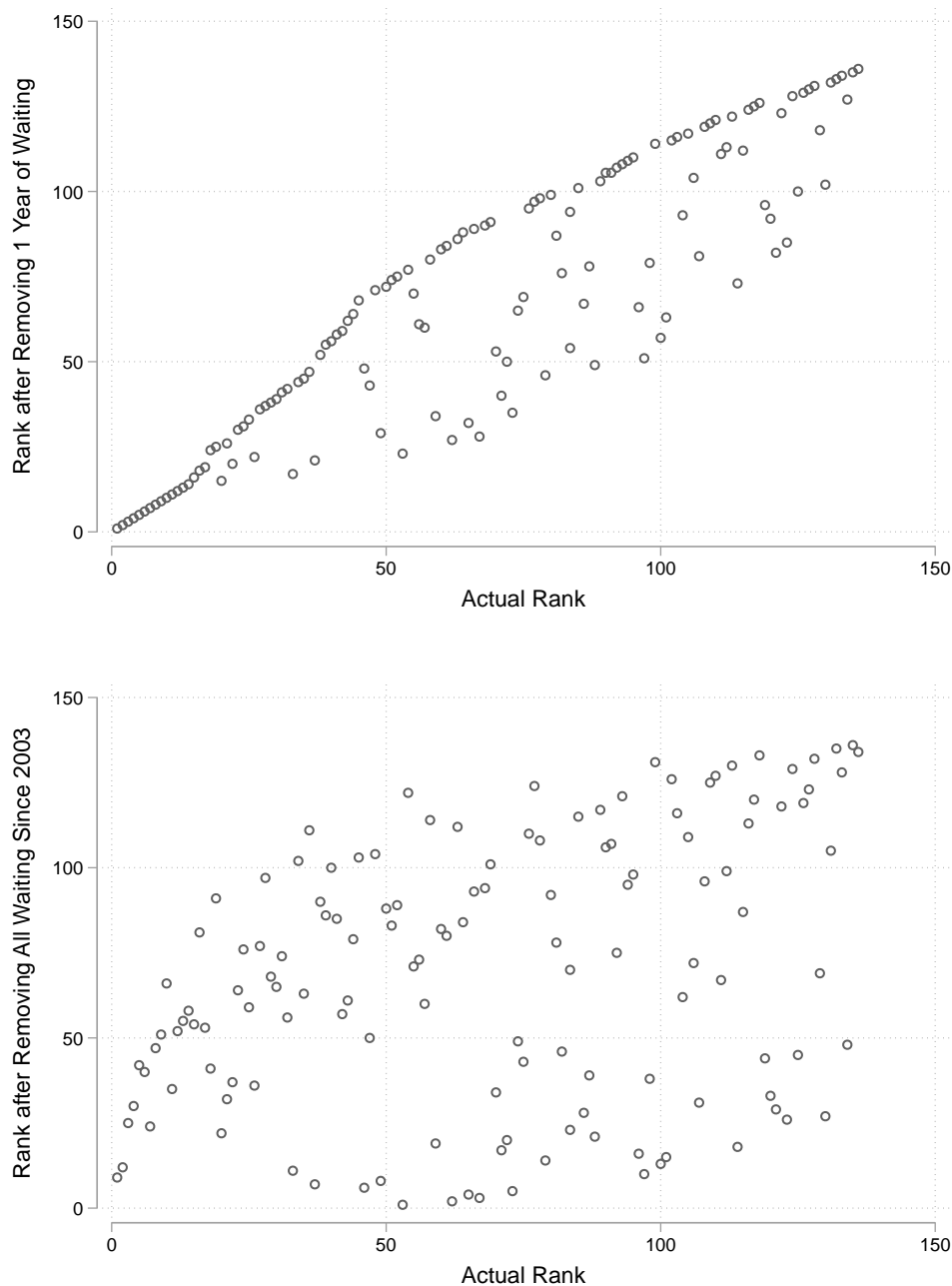
Notes: The binscatter figure shows the relationship, at the district-year level, between district demeaned waitlist percentile and district demeaned years without formal employment where the demeaning is done by risk set-year. A risk set is a degree conferral year-month and subject combination. The waitlist percentile is the initial position on the fresh graduates waitlist, normalized to vary from 0 to 1. The sample includes all teachers on a waitlist between 2003 and 2011 whose degree was conferred before 2006.

Figure A.8: Panhellenic Test Scores ( $\sigma$ ) and Waitlist Position – District Level



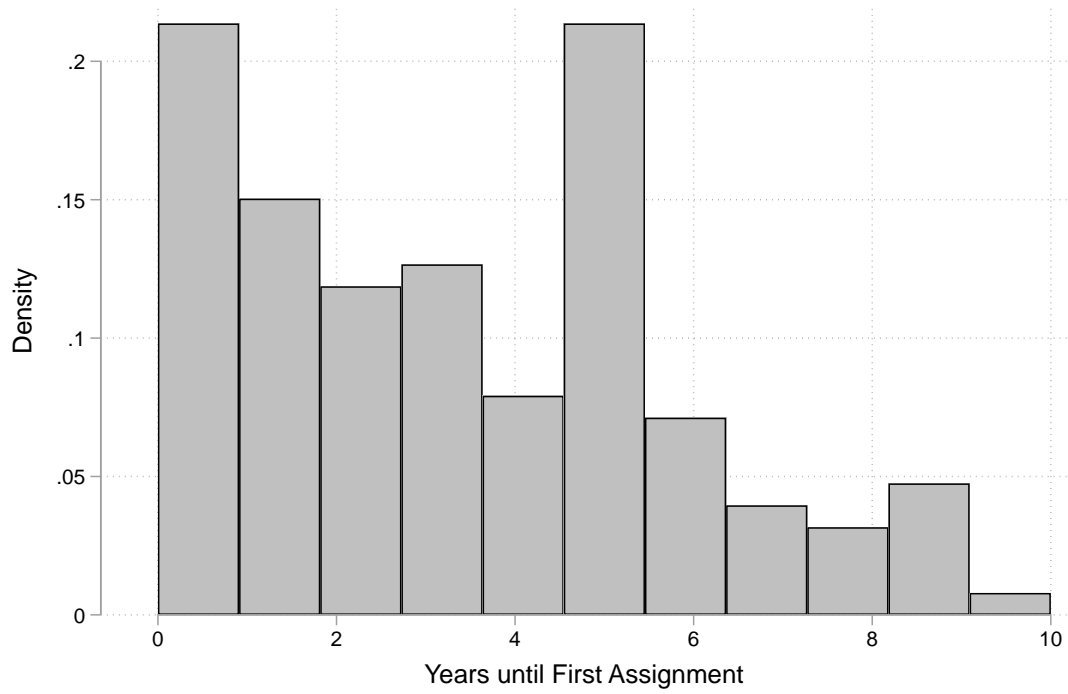
Notes: The binscatter figure shows the relationship, at the district-year level, between district demeaned waitlist percentile and district Panhellenic test scores (in student standard deviation units) where the demeaning is done by risk set-year. A risk set is a degree conferral year-month and subject combination. The waitlist percentile is the initial position on the fresh graduates waitlist, normalized to vary from 0 to 1. The sample includes all teachers on a waitlist between 2003 and 2011 whose degree was conferred before 2006.

**Figure A.9: Cross-District Effects of Eliminating Waiting**



Notes: The figures show how changing the time out of formal employment affects districts' test score ranks. The top figure reduces the number of years deputy teachers wait without formal employment by 1 year; the bottom figure reduces the number of years deputy teachers wait to 0. We calculate each district's test score rank, under the actual scores and under counterfactuals, where the district with the lowest mean test scores has the rank 1. For each counterfactual, we take the absolute value of our point estimate from the district-level model, multiply by each district's heterogeneous exposure to the deputy assignment system, and then multiply by the number of years waiting reduced in the counterfactual.

**Figure A.10: Years Waiting for First Assignment – Survey**



Notes: This figure shows the distribution of the years spent waiting between degree conferral and first teaching assignment. The sample is the teachers who took our online survey. Responses with implied waiting times that are negative or more than 10 years have been excluded.

**Table A.1: Activities while Waiting**

	Fraction
Private Lessons	0.54
Further Studies	0.33
Started a Family	0.16
Non-Teaching Work in Education Sector	0.19
Work in Non-Education Sector	0.39
Other	0.19

Notes: This table shows the fraction of survey respondents in various activities during the time spent waiting for an assignment. Respondents in the online survey could choose multiple activities.

**Table A.2: Teachers Working in Part-Time Jobs**

	Part-Time Work while Teaching	Continued Part-Time Work while Teaching
Yes	0.28	0.33
Private Lessons	0.09	0.17
Private School	0.02	0.01
Tourism Industry	0.01	0.01
Work in Other Private Sector	0.04	0.03
Other	0.04	0.03
No	0.72	0.67
<i>N</i>	157	158

Notes: This table shows the part-time jobs teachers report having, while teaching, in the online survey. “Yes” indicates the teacher held a part-time job while teaching. “Part-Time Work while Teaching” is the distribution of part-time jobs while “Continued Part-Time Work while Teaching” is the distribution of part-time jobs that continue activities started while the teacher was waiting for an assignment. Respondents in the online survey could choose multiple responses.



**Table A.3: Relationship between Years Waiting and Activities during Teaching**

	Activity while Teaching	Continued Activity
Years until First Assignment	0.00706 (0.0143)	0.0126 (0.0147)
Constant	0.249 (0.0598)	0.292 (0.0608)
Mean DV	0.273	0.333
N	132	132

Notes: This table shows regressions of activity while teaching on years spent waiting until first assignment. Years spent waiting is calculated as the difference between the year of the first teaching assignment and the year of degree conferral. “Activity while Teaching” is a dummy variable for whether the teacher participated in an extra activity – e.g., private lessons – while working as a public school teacher. “Continued Activity” is a dummy variable for whether the teacher participated in an extra activity, while working as a public school teacher, that had been started while the teacher was waiting for an assignment. The sample is the teachers responding to our online survey.

**Table A.4: OLS and IV across Samples**

	OLS	IV	OLS	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.0355 (0.0312)	-0.0518 (0.0103)	-0.0595 (0.0211)	-0.0499 (0.0121)
Deputy	-0.0802 (0.0555)	-0.0860 (0.0606)	-0.0575 (0.0513)	-0.0546 (0.0526)
Prior Year GPA	0.737 (0.00887)	0.737 (0.00886)	0.737 (0.00888)	0.737 (0.00887)
Mean DV	0.0372	0.0372	0.0387	0.0387
Clusters	390	390	389	389
N	54370	54370	54282	54282
Risk Set	Yes	Yes	Yes	Yes
Sample	All Teachers	All Teachers	No Special Cases	No Special Cases

Notes: The table presents OLS and instrumental variable estimates. An observation is a student-subject-year. The dependent variable is the student's subject-specific test score, in standard deviation units. The instrument is the assigned teacher's imputed waitlist position (demeaned by risk set), normalized to run between 0 and 1 within a risk set. Risk sets are teachers in the same subject whose degrees were conferred in the same month-year. The sample includes all teachers ("All Teachers") or only teachers without special circumstances that affect waitlist position ("No Special Cases"). Special circumstances refer to factors that directly affect waitlist positions (having at least three children, having a special needs child, military service, and reading Braille). Permanent teachers each have their own risk set. Standard errors are clustered by teacher.

**Table A.5: Effect of Years without Formal Employment on Students' Tests by Subject**

	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.0518 (0.0103)	-0.0653 (0.0173)	-2.206 (6.968)
Deputy	-0.0860 (0.0606)	-0.180 (0.101)	-0.376 (1.044)
Prior Year GPA	0.737 (0.00886)	0.735 (0.0117)	0.739 (0.0129)
Mean DV	0.0372	0.0536	0.0265
Clusters	390	161	278
N	54370	21450	32920
Risk Set	Yes	Yes	Yes
Sample	All	STEM	Non-STEM

Notes: The table includes instrumental variable estimates with (demeaned) imputed waitlist position as the instrument. An observation is a student-subject-year. "Years Waiting" is the deputy teacher's years without formal employment, normalized by the risk set size to be in percentiles. "Years Waiting" and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. The outcome is the student's average full-year score in student standard deviation units ( $\sigma$ ). STEM fields are algebra, geometry, mathematics, biology, physics, technology, and science. Standard errors are clustered by teacher.

**Table A.6: Effect of Years without Formal Employment on Students' Grade Point Average**

	IV	IV	IV
	GPA ( $\sigma$ )	GPA	Log GPA
Years Waiting	-0.0402 (0.00543)	-0.119 (0.0392)	-0.00786 (0.00291)
Deputy	-0.0926 (0.0299)	-0.0400 (0.186)	-0.00413 (0.0136)
Prior Year GPA	0.930 (0.00366)	2.589 (0.0147)	0.178 (0.00146)
Mean DV	-0.00171	14.95	2.686
Clusters	390	393	393
N	54099	74555	74499
Risk Set	Yes	Yes	Yes

Notes: The table includes instrumental variable estimates with (demeaned) imputed waitlist position as the instrument. An observation is a student-subject-year where the outcomes do not vary by subject but the teachers do. "Years Waiting" is the deputy teacher's years without formal employment, normalized by the risk set size to be in percentiles. "Years Waiting" and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. The outcome is the student's grade-point-average (out of 20), expressed in student standard deviation units ( $\sigma$ ), levels, or logs. Standard errors are clustered by teacher.

**Table A.7: District (Demeaned) Waitlist Position and Mean Expected Years without Formal Employment**

	Mean $\mathbb{E}$ [Years Waiting]	Mean $\mathbb{E}$ [Years Waiting]
Mean Waitlist Perc	-0.207 (0.644)	-0.760 (0.635)
Mean DV	3.556	3.556
N	394	394
District FE	No	Yes
Reg-Yr FE	No	Yes

Notes: This table tests the identifying assumption behind our aggregation to a district-level model. An observation is a district-year. The dependent variable is the mean expected years without formal employment where the expected years without formal employment is the mean over a teacher's risk set and the first mean is taken over the teachers assigned to the district in a given year. The explanatory variable is the district-year's mean of its assignees' (demeaned) waitlist position.

**Table A.8: Survey Responses**

	Mean	Std. Dev.	Obs
Degree Conferral Year	2007	8	187
Year of First Assignment	2010	9	159
Num Schools Worked in	1.42	0.87	93
Aware of System when Choosing Teaching	0.34	0.47	184
Understand the Assignment Process	0.76	0.43	184
Knew Waitlist Position	0.73	0.44	184
Considered Attriting	0.49	0.50	200
Rejected an Assignment	0.12	0.33	185
Has Left Teaching	0.20	0.40	200
Believes Skills Depreciated while Waiting	0.18	0.38	187
Invested in Skill Maintenance	0.46	0.50	200

Notes: This table shows summary statistics for selected responses to the online survey of teachers. “Num Schools Worked in” indicates the number of different schools a teacher worked in at the same time.

**Table A.9: District-Level Instrument Robustness: Test Percentiles**

	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)
Years Waiting	-8.904 (3.000)	-10.02 (3.061)	-11.28 (3.174)	-9.698 (3.360)	-7.395 (2.849)	-11.75 (4.778)	-14.23 (4.445)
Ln Class Size	8.050 (2.027)	8.044 (2.131)	8.037 (2.252)	11.15 (2.920)	8.059 (1.896)	11.39 (3.135)	8.021 (2.561)
IV	Baseline	Mean across Lists	Max across Lists	Deputy Position	Risk Set Norm	Raw Position	Log Position
Per Class	-1.7808	-2.0035	-2.2565	-1.9397	-1.4790	-2.3509	-2.851
AR LB	-18.1107	-18.8050	-20.3924	-20.0102	-20.6474	-38.7139	-33.1431
AR UB	-4.4487	-5.4719	-6.5700	-4.7088	-3.7292	-5.6059	-8.5058
Mean DV	49.07	49.07	49.07	49.28	49.07	49.28	49.07
N	390	390	390	337	390	337	390
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table includes IV regressions. An observation is a district-year. “Years Waiting” is the deputy teacher’s years without formal employment. The different forms of the instruments are listed with labels indicating our baseline model (“Baseline”), an instrument using the mean position across deputy and hourly lists (“Mean across Lists”), an instrument using the maximum position (“Max across Lists”), an instrument using the position from the deputy list (“Deputy Position”), our baseline instrument scaled by risk set list length (“Risk Set Norm”), the actual waitlist position (“Raw Position”), and the log of the waitlist position (“Log Position”). All variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Test score outcomes are student performance on the national twelfth grade exams, in percentiles (“Perc”). “Reg-Yr FE” are region-year fixed effects. “Per Class” indicates the per-class effect, which is the main coefficient divided by 5 for the 5 classes twelfth graders take in tested subjects. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.10: District-Level Instrument Robustness: Test Score in Standard Deviations**

	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.342 (0.121)	-0.373 (0.120)	-0.408 (0.123)	-0.208 (0.115)	-0.242 (0.102)	-0.319 (0.164)	-0.419 (0.160)
Ln Class Size	0.240 (0.0983)	0.240 (0.101)	0.240 (0.104)	0.305 (0.144)	0.241 (0.0906)	0.317 (0.153)	0.240 (0.105)
IV	Baseline	Mean across Lists	Max across Lists	Deputy Position	Risk Set Norm	Raw Position	Log Position
Per Class	-0.0683	-0.0746	-0.0815	-0.0416	-0.0484	-0.0638	-0.0838
AR LB	-0.6891	-0.7187	-0.7361	-0.5374	-0.6755	-1.1476	-1.0387
AR UB	-0.1620	-0.1940	-0.2250	-0.0373	-0.0907	-0.1077	-0.1808
Mean DV	-0.0440	-0.0440	-0.0440	-0.0367	-0.0440	-0.0367	-0.0440
N	390	390	390	337	390	337	390
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table includes IV regressions. An observation is a district-year. “Years Waiting” is the deputy teacher’s years without formal employment. The different forms of the instruments are listed with labels indicating our baseline model (“Baseline”), an instrument using the mean position across deputy and hourly lists (“Mean across Lists”), an instrument using the maximum position (“Max across Lists”), an instrument using the position from the deputy list (“Deputy Position”), our baseline instrument scaled by risk set list length (“Risk Set Norm”), the actual waitlist position (“Raw Position”), and the log of the waitlist position (“Log Position”). All variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Test score outcomes are student performance on the national twelfth grade exams, in student standard deviation units. “Reg-Yr FE” are region-year fixed effects. “Per Class” indicates the per-class effect, which is the main coefficient divided by 5 for the 5 classes twelfth graders take in tested subjects. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).



**Table A.11: District-Level Instrument Robustness: Test Percentiles, Controlling for Experience**

	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)	Score (Percentile)
Years Waiting	-6.987 (2.561)	-7.618 (2.623)	-8.243 (2.842)	-7.306 (3.397)	-5.632 (2.336)	-4.742 (2.710)	-9.625 (3.312)
Ln Class Size	5.925 (1.606)	5.813 (1.617)	5.702 (1.635)	8.089 (2.076)	6.165 (1.618)	8.416 (2.080)	5.458 (1.701)
IV	Baseline	Mean across Lists	Max across Lists	Deputy Position	Risk Set Norm	Raw Position	Log Position
Per Class	-1.3973	-1.5237	-1.6487	-1.4611	-1.1265	-0.9484	-1.9249
AR LB	-13.3235	-14.1084	-15.2772	-16.3860	-10.9511	-10.9129	-17.8213
AR UB	-3.1845	-3.7244	-4.0232	-2.2610	-1.7013	-0.1808	-4.7066
Mean DV	49.07	49.07	49.07	49.28	49.07	49.28	49.07
N	390	390	390	337	390	337	390
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Risk Set	Incl Exp	Incl Exp	Incl Exp	Incl Exp	Incl Exp	Incl Exp	Incl Exp

Notes: The table includes IV regressions that control for experience. An observation is a district-year. “Years Waiting” is the deputy teacher’s years without formal employment. The different forms of the instruments are listed with labels indicating our baseline model (“Baseline”), an instrument using the mean position across deputy and hourly lists (“Mean across Lists”), an instrument using the maximum position (“Max across Lists”), an instrument using the position from the deputy list (“Deputy Position”), our baseline instrument scaled by risk set list length (“Risk Set Norm”), the actual waitlist position (“Raw Position”), and the log of the waitlist position (“Log Position”). All variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year and the same number of years of prior experience. Test score outcomes are student performance on the national twelfth grade exams, in percentiles (“Perc”). “Reg-Yr FE” are region-year fixed effects. “Per Class” indicates the per-class effect, which is the main coefficient divided by 5 for the 5 classes twelfth graders take in tested subjects. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.12: Student-Level Analysis – Deputy Teachers Only**

	IV	IV	IV	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )	Admitted	Admitted	Admitted
Years Waiting	-0.0518 (0.0103)	-0.0518 (0.0103)	-0.0652 (0.0153)	-0.0183 (0.00752)	-0.0184 (0.00780)	-0.0169 (0.00534)
Deputy	-0.0860 (0.0606)			-0.0105 (0.0388)		
Prior Year GPA	0.737 (0.00886)	0.735 (0.0225)	0.718 (0.0368)	0.293 (0.00405)	0.296 (0.0198)	0.305 (0.0202)
Mean DV	0.0372	-0.0675	-0.136	0.670	0.654	0.658
Clusters	390	42	23	391	42	23
N	54370	2686	1116	72879	3839	1565
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All Teachers	Deputy Teachers	Confident Matches	All Teachers	Deputy Teachers	Confident Matches

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the sample across columns to show robustness to excluding non-deputy teachers. “Confident Matches” include only deputy teachers for whom the name match between the micro school data and the administrative waitlist data is unambiguous. An observation is a student-subject-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are clustered by teacher.

**Table A.13: Student-Level Analysis, Varying Controls**

	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.0426 (0.0247)	-0.0375 (0.0111)
Deputy	-0.111 (0.0469)	-0.130 (0.0618)
Demeaned Prior GPA		0.746 (0.0103)
Mean DV	-0.00708	0.0372
Clusters	466	390
N	82837	54370
Risk Set	Yes	Yes

Notes: The table presents instrumental variable estimates. An observation is a student-subject-year. The dependent variable is the student's subject-specific test score, in standard deviation units. The instrument is the assigned teacher's imputed waitlist position (demeaned by risk set), normalized to run between 0 and 1. Risk sets are teachers in the same subject whose degrees were conferred in the same month-year. Demeaned prior GPA is demeaned at the teacher-year level. The sample includes students taught by deputy and permanent teachers. Permanent teachers each have their own risk set. Standard errors are clustered by teacher.

**Table A.14: Student-Level Analysis – Different Functional Forms**

	IV	IV	IV	IV
	Score ( $\sigma$ )	Ln Score	Score ( $\sigma$ )	Ln Score
Years Waiting	-0.0518 (0.0103)	-0.0255 (0.0101)		
Deputy	-0.0860 (0.0606)	-0.0462 (0.0424)	-0.0961 (0.0700)	-0.0512 (0.0448)
Prior Year GPA	0.737 (0.00886)	0.204 (0.00450)	0.737 (0.00885)	0.204 (0.00450)
Ln Years Waiting			-0.834 (0.164)	-0.411 (0.145)
Mean DV	0.0372	2.639	0.0372	2.639
Clusters	390	390	390	390
N	54370	54359	54370	54359
Risk Set	Yes	Yes	Yes	Yes

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the functional form of the test outcome (student standard deviation units or log scores) and the measure of years spent not working formally (levels or log). An observation is a student-subject-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are clustered by teacher.

**Table A.15: Student-Level Analysis – Different Samples**

	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.0518 (0.0103)	-0.0532 (0.0129)	-0.0537 (0.0126)
Deputy	-0.0860 (0.0606)	-0.107 (0.0776)	-0.0960 (0.0701)
Prior Year GPA	0.737 (0.00886)	0.737 (0.00896)	0.737 (0.00891)
Mean DV	0.0372	0.0376	0.0379
Clusters	390	376	386
N	54370	53430	53997
Risk Set	Yes	Yes	Yes
Sample	All	Month Not July	Day Not 30th

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the sample across columns to show robustness to excluding common degree conferral months (July) or days (the 30th). An observation is a student-subject-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are clustered by teacher.

**Table A.16: Student-Level Analysis – Different Outcome Definitions**

	IV	IV	IV	IV	IV	IV	IV
	Score	Ln Score	Univ Score	Selec (Adm)	Selec (Natl)	Selec (2003 Adm)	Selec (2003 Natl)
Years Waiting	-0.284 (0.112)	-0.0255 (0.0101)	-233.7 (50.55)	-2.166 (0.395)	-2.100 (0.393)	-2.174 (0.345)	-2.080 (0.339)
Deputy	-0.490 (0.489)	-0.0462 (0.0424)	-348.2 (265.7)	-3.523 (2.032)	-2.954 (2.035)	-3.825 (1.819)	-3.617 (1.790)
Prior Year GPA	2.737 (0.0471)	0.204 (0.00450)	3017.3 (33.97)	24.49 (0.246)	23.74 (0.250)	22.41 (0.301)	21.55 (0.301)
Mean DV	14.60	2.639	14320.8	49.37	49.19	52.61	52.20
Clusters	390	390	370	370	370	370	370
N	54370	54359	48818	48818	48818	35345	35345
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the functional form of the outcome across columns. An observation is a student-subject-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. “Score” is the student’s average subject-specific test score during the year. This is the raw score that is not standardized to be in student standard deviation units. “Ln Score” is the natural log of the raw score. “Univ Score” is the student’s university admissions score, in levels. For the selectivity measures (“Selec”), we order the university-programs according to their enrollees’ mean statistic, defined below, and rank them from highest to lowest. The measure is the percentile of this ordering where 100 is the program whose admits have the highest mean score. “Adm” uses the university admissions score for ranking while “Natl” uses the national Panhellenica score, which is an alternate weighting. The “2003” measures use the 2003 cohort to construct the measures to avoid overlap with our analysis sample. Standard errors are clustered by teacher.

**Table A.17: Effect of Years without Full-Time Formal Employment on Students' Subject Exam Scores**

	OLS	RF	FS	IV	IV	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Years Waiting	Score ( $\sigma$ )	Score (Perc)	First Sem ( $\sigma$ )	Second Sem ( $\sigma$ )	Exam ( $\sigma$ )
Years Waiting	-0.0188 (0.0386)			-0.0458 (0.00943)	-1.365 (0.269)	-0.0320 (0.0102)	-0.0391 (0.0106)	-0.0473 (0.0104)
Deputy	-0.0778 (0.0562)	-0.0786 (0.0590)	-0.304 (0.144)	-0.0925 (0.0628)	-2.977 (1.831)	-0.0656 (0.0600)	-0.113 (0.0663)	-0.0824 (0.0692)
Prior Year GPA	0.737 (0.00887)	0.737 (0.00887)	0.000870 (0.00109)	0.737 (0.00886)	21.90 (0.225)	0.599 (0.0124)	0.643 (0.0139)	0.703 (0.00845)
Waitlist Perc		-1.439 (0.300)	31.43 (0.745)					
Mean DV	0.0372	0.0372	-0.0269	0.0372	51.06	0.111	0.0783	-0.0239
Clusters	390	390	390	390	390	390	390	388
N	54370	54370	54370	54370	54370	54360	54323	54068
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table includes OLS, reduced form (“RF”), first stage (“FS”), and IV regressions. An observation is a student-subject-year. “Years Waiting” is the deputy teacher’s years without *full-time* formal employment and “Waitlist Perc” is the imputed waitlist position, normalized by the risk set size to be in percentiles. Both variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. We define full-time formal employment as teachers who accrue above-median credits in a year. “Score” is the student’s average subject-specific test score during the year. “First Sem” and “Second Sem” are the semester-specific test scores, and “Exam” is the end-of-year exam. This exam is a national exam for 11th graders before 2006 and 12th graders in all years; otherwise it is the school exam. Test results are expressed in student standard deviation units ( $\sigma$ ) or percentiles. Standard errors are clustered by teacher.

**Table A.18: Bootstrapped Standard Errors**

	Point Estimate	Standard Error	Bootstrap Standard Error
OLS, Score ( $\sigma$ )	-0.0355	0.0312	0.0306
RF, Score ( $\sigma$ )	-1.4386	0.2997	0.3039
FS, Years Waiting	27.7714	1.1884	1.1734
IV, Score ( $\sigma$ )	-0.0518	0.0103	0.0111
IV, Score (Perc)	-1.5922	0.3414	0.3708
IV, First Sem ( $\sigma$ )	-0.0363	0.0111	0.0110
IV, Second Sem ( $\sigma$ )	-0.0443	0.0113	0.0111
IV, Exam ( $\sigma$ )	-0.0535	0.0117	0.0127
IV, Ln Univ Score	-0.0192	0.0041	0.0035
IV, List Length	1.2448	0.2835	0.3173
IV, Admitted	-0.0183	0.0075	0.0076
IV, Acad Univ	-0.0126	0.0056	0.0057
IV, Selectivity (Admissions)	-2.1662	0.3951	0.3718

Notes: The table shows non-bootstrapped and bootstrapped standard error estimates for our main individual-level analysis results. Standard errors are clustered by teacher. “RF” and “FS” indicate reduced form and first stage regressions, respectively. Bootstrapped standard errors are produced using 500 iterations of a wild clustered bootstrap with a Rademacher distribution. We bootstrap the entire process, including the calculation of the risk set means.



**Table A.19: District-Level Analysis – Different Controls**

	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.374 (0.160)	-0.413 (0.157)	-0.414 (0.158)
Ln Class Size	-0.712 (0.184)	-0.810 (0.198)	-0.806 (0.198)
Num Teachers		0.0369 (0.0120)	
Num Students			0.00164 (0.000559)
Per Class	-0.0748	-0.0827	-0.0828
AR LB	-0.7347	-0.7562	-0.7580
AR UB	-0.0891	-0.1330	-0.1321
Mean DV	-0.0442	-0.0442	-0.0442
N	394	394	394
District FE	Resid	Resid	Resid
Reg-Yr FE	Resid	Resid	Resid
Risk Set	Yes	Yes	Yes

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. The columns vary our use of controls. All columns use (demeaned) fixed effects. The second and third columns include additional district-level controls. An observation is a district-year. All variables are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.20: District-Level Analysis – Different Functional Forms**

	Score ( $\sigma$ )	Ln Score	Score ( $\sigma$ )	Ln Score
Years Waiting	-0.342 (0.121)	-0.129 (0.0477)		
Ln Class Size	0.240 (0.0983)	0.132 (0.0352)	0.271 (0.0920)	0.143 (0.0334)
Ln Years Waiting			-0.711 (0.238)	-0.268 (0.0991)
Per Class	-0.0683	-0.0258	-0.1423	-0.0537
AR LB	-0.7010	-0.2478	-1.3430	-0.4997
AR UB	-0.1548	-0.0439	-0.3247	-0.0918
Mean DV	-0.0440	2.504	-0.0440	2.504
N	390	390	390	390
District FE	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the functional form of the test outcome (student standard deviation units or log scores) and the measure of years spent not working formally (levels or log). An observation is a district-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.21: District-Level Analysis – Different Samples**

	IV	IV	IV
	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.342 (0.121)	-0.192 (0.0844)	-0.297 (0.103)
Ln Class Size	0.240 (0.0983)	0.252 (0.0849)	0.240 (0.0922)
Per Class	-0.0683	-0.0384	-0.0593
AR LB	-0.7010	-0.3826	-0.5875
AR UB	-0.1548	-0.0485	-0.1370
Mean DV	-0.0440	-0.0440	-0.0440
N	390	390	390
District FE	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes
Sample	All	Month Not July	Day Not 30th

Notes: This table shows IV regressions with (demeaned) imputed wait-list position as the instrument. We vary the sample across columns to show robustness to excluding common degree conferral months (July) or days (the 30th). An observation is a district-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.22: District-Level Analysis – Different Outcome Definitions**

	IV	IV	IV	IV	IV	IV	IV
	Score	Ln Score	Univ Score	Selec (Adm)	Selec (Natl)	Selec (2003 Adm)	Selec (2003 Natl)
Years Waiting	-1.350 (0.452)	-0.129 (0.0477)	-1152.0 (454.7)	-5.638 (2.571)	-5.091 (2.649)	-3.686 (2.967)	-2.823 (3.136)
Ln Class Size	1.351 (0.330)	0.132 (0.0352)	1728.1 (425.5)	9.170 (2.254)	9.569 (2.228)	7.822 (1.826)	8.068 (1.862)
Per Class	-0.2700	-0.0258	-230.4	-1.128	-1.018	-0.737	-0.565
AR LB	-2.5858	-0.2478	-2.3e+03	-12.6628	-12.3293	-12.7330	-12.3841
AR UB	-0.6158	-0.0439	-3.4e+02	-1.4645	-0.7895	0.6606	1.7711
Mean DV	13.10	2.504	14274.0	48.77	48.58	52.40	52.12
N	390	390	390	390	390	385	385
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows IV regressions with (demeaned) imputed waitlist position as the instrument. We vary the functional form of the outcome across columns. An observation is a district-year. “Years Waiting” and the instrument are demeaned by risk set, where a risk set is the cohort of teachers in the same subject with degrees conferred in the same month-year. “Score” is the student’s average subject-specific test score during the year. This is the raw score that is not standardized to be in student standard deviation units. “Ln Score” is the natural log of the raw score. “Univ Score” is the student’s university admissions score, in levels. For the selectivity measures (“Selec”), we order the university-programs according to their enrollees’ mean statistic, defined below, and rank them from highest to lowest. The measure is the percentile of this ordering where 100 is the program whose admits have the highest mean score. “Adm” uses the university admissions score for ranking while “Natl” uses the national Panhellenica score, which is an alternate weighting. The “2003” measures use the 2003 cohort to construct the measures to avoid overlap with our analysis sample. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in “AR LB” and “AR UB” (Anderson and Rubin, 1949).

**Table A.23: Effect of Years without Formal Employment on Districts' Panhellenic Exam Scores without Scaling**

	Score ( $\sigma$ )	Score (Percentile)	Ln University Score	Admitted
Years Waiting	-0.121 (0.0601)	-3.147 (1.491)	-0.0578 (0.0458)	-0.0441 (0.0222)
Ln Class Size	0.106 (0.0534)	2.193 (1.290)	0.0233 (0.0412)	0.0104 (0.0166)
Per Class	-0.0241	-0.6294	-0.0116	-0.0088
AR LB	-0.2755	-6.9844	-0.1611	-0.0959
AR UB	-0.0231	-0.7262	0.0274	-0.0045
Mean DV	-0.0440	49.07	9.518	0.818
N	390	390	390	390
District FE	Yes	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes	Yes

Notes: The table include the main IV regressions without incorporating the deputies we lack inexperience waitlist positions for. An observation is a district-year and the IV outcomes are measures of student performance on the national twelfth grade exams, as mean level (in student standard deviation units) or mean percentile, and university admissions outcomes. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in "AR LB" and "AR UB" (Anderson and Rubin, 1949).

**Table A.24: Effect of Years without Formal Employment on Districts' Panhellenic Exam Scores with Alternate Weighting**

	Score ( $\sigma$ )	Score ( $\sigma$ )	Score ( $\sigma$ )
Years Waiting	-0.342 (0.121)	-0.454 (0.172)	-0.468 (0.132)
Ln Class Size	0.240 (0.0983)	0.502 (0.116)	0.598 (0.0882)
Per Class	-0.0683	-0.0908	-0.0937
AR LB	-0.7010	-0.9650	-0.8392
AR UB	-0.1548	-0.2023	-0.2645
Mean DV	-0.0440	-0.150	-0.178
N	390	390	390
District FE	Yes	Yes	Yes
Reg-Yr FE	Yes	Yes	Yes
Risk Set	Yes	Yes	Yes
Weighting	Num Students	Num Deputies	None

Notes: The table includes the main IV regressions using alternate weighting. An observation is a district-year and the IV outcomes are measures of student performance on the national twelfth grade exams, in student standard deviation units. Weighting is shown in the last row. Standard errors are heteroskedasticity-robust. Anderson-Rubin weak instrument-robust 95% confidence intervals are reported in "AR LB" and "AR UB" (Anderson and Rubin, 1949).