Quality Signals and the Wage Returns to Accreditation: Evidence from Chile

Cristhian Molina* Rocío Valdebenito†

March 27, 2026

[Preliminary draft, do not cite]

[Click Here for the Latest Version]

Employers cannot directly observe the quality of the degree programs completed by job candidates, and institutional reputation serves as a reliable signal for only a small fraction of graduates. We study whether program-level accreditation in Chile resolves this uncertainty and adjusts wages. Using a staggered difference-in-differences design applied to administrative earnings records for 2,572 programs, we exploit first-time accreditation decisions issued between 2007 and 2018. Accreditation raises log annual wages by approximately 3 percent on average, with the premium becoming most pronounced four to five years after the initial award. This average masks significant heterogeneity: the wage premium reaches 9 percent for graduates of lower-quality institutions but is null or negative for those from high-reputation programs. This pattern is consistent with a Bayesian employer-learning model in which public certification is most informative where prior reputations are weakest. Our findings suggest that accreditation acts as a corrective signal that redistributes labor market returns toward graduates whose credentials would otherwise be discounted.

Keywords: Higher Education, Market Signaling, College Choice, Quality Disclosure

JEL: I21, I23, I28, D82, L15

*Department of Economics, University of Illinois at Urbana-Champaign. cmolina3@illinois.edu

†Department of Economics, Universidad Alberto Hurtado. rvaldebeni@uahurtado.edu

1 Introduction

Consider the information problem facing a recruiter at a mid-size firm in Santiago. A resume arrives from a graduate of a program the recruiter has never encountered: a business degree from a regional university, perhaps, or an engineering program at one of the dozens of institutions that expanded rapidly after Chile deregulated higher education in the 1980s. The recruiter cannot inspect the curriculum, cannot easily verify the faculty, and institutional reputation is, at best, a rough signal that distinguishes a handful of elite universities from everyone else (MacLeod and Urquiola, 2015; MacLeod et al., 2017). In this environment, a third-party certification that publicly ranks program quality can do real work: it compresses an information gap that neither applicants nor employers can bridge on their own (Akerlof, 1970; Dranove and Jin, 2010). The theoretical prediction that follows from this logic is precise. Certification matters most where uncertainty is highest. At institutions whose graduates already carry a recognizable brand, the certification label adds little the employer did not already know. At institutions where the employer's prior is weak or dispersed, the label resolves genuine uncertainty, and the wage responds. Whether real-world accreditation systems operate this way is an empirical question. This paper answers it.

We test this prediction using program-level accreditation by Chile's National Accreditation Commission (CNA). When a degree program is accredited for the first time, its certified status is publicly disclosed in government registries that employers routinely consult during candidate screening. The defining feature of our setting is that voluntary accreditation generates substantial cross-program variation in the *timing* of this disclosure, which we exploit for identification. Our two questions map directly onto the theoretical framework. Does first-time accreditation raise the wages of graduates? And is any wage return concentrated among graduates of lower-reputation institutions, the programs where the signal is most informative, as the signaling model predicts?

To answer these questions, we use administrative records from Chile's Ministry of Finance linking higher education graduation files to private-sector earnings reported to the Internal Revenue Service (SII). These records yield average annual wages of graduates one, two, and three years after graduation at the program–cohort level. Our estimation sample covers 2,572 degree programs: 1,725 that received first-time accreditation between 2007 and 2018, and 847 that remained unaccredited throughout.

While programs endogenously choose *when* to seek accreditation, the final CNA decision (both the binary outcome and the number of years awarded) is made by an independent panel of commissioners and is largely unanticipated by employers at the time of hiring. We exploit the staggered timing of these public announcements in a difference-in-differences framework, using the imputation estimator of Borusyak et al. (2024) to avoid the negative-weight problems of two-way fixed effects in staggered settings (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021). The pre-trend tests pass with $p > 0.6$ across all three wage horizons, supporting the parallel trends assumption. With identification in hand, the results are striking.

First-time accreditation raises graduate wages by approximately 3 percent on average, with the effect accumulating over the post-accreditation event window and reaching statistical significance by the second year after graduation. That average, however, is doing exactly what averages do in the presence of heterogeneous treatment effects: it obscures the mechanism. When we separate programs by their institution's accreditation standing, which proxies for how much the employer already knows about graduate quality, the heterogeneity is sharp and follows the signaling prediction with unusual precision. Graduates of programs at Baseline institutions (1–3 year institutional accreditation term) experience a wage premium of 8.9 percent in year 1, declining to 4.6 and 2.6 percent in years 2 and 3 as employer learning accumulates. Graduates of programs at Top-tier institutions (6–7 year term) experience a wage *penalty* of 5.0 percent. The same pattern holds when we stratify by the program's own first-accreditation term, which proxies for treatment intensity: programs receiving a Low-intensity first accreditation (1–3 years) show a wage premium of 10.7 percent in year 1, while programs receiving a High-intensity accreditation show no gain. The reversal at the top is not a noise artifact; it reflects the fact that accreditation provides no new information about graduates from already-recognized elite programs, and may even trigger employer skepticism about why such a program needed external validation. Accreditation generates the largest wage premium precisely where employers had the least prior information about graduate quality.

One threat to a signaling interpretation is institutional gaming. CNA evaluations assess observable metrics including graduation rates, creating incentives for programs to accelerate graduation around the evaluation window (Dranove and Jin, 2010). If accreditation draws additional graduates into the labor market, our wage estimates could reflect a compositional shift rather than employer updating. A timing argument disciplines the most direct version of this concern: graduates observed at event times $\kappa = 1$ through $\kappa = 4$ were enrolled *before* accreditation was awarded and could not have selected into the program on the basis of its certified status. Their wage outcomes are a clean test of the employer-updating channel. We document that total graduation counts do rise following accreditation, and the pattern shows pre-trend violations consistent with anticipation effects during the preparation period, which limits our ability to make clean causal claims about graduation gaming using the same research design. The wage result is unaffected: pre-trend tests for log wages pass with $p > 0.6$ across all three wage horizons, and the wage premium reflects higher earnings conditional on formal employment, not a change in the size or selectivity of the employed pool.

Taken together, these findings speak to three questions the existing literature has not been able to answer cleanly. The most direct is causal: does program-level accreditation raise wages? The closest predecessor, González-Velosa et al. (2015), estimates that an additional year of CNA institutional accreditation is associated with a 9.2 percentage-point wage premium using cross-sectional OLS, but that design cannot separate the accreditation effect from pre-existing differences across institutions. By using quasi-experimental variation in the timing of the information shock, we contribute to the literature on causal returns to college quality (e.g., Dale and Krueger, 2002; Hoekstra, 2009; Reyes et al., 2016) with estimates that are not confounded by selection. The answer is yes: accreditation raises wages. But the magnitude and even the sign of the effect depend entirely on *where* it occurs.

That dependence is the second question, and it is where the paper makes its sharpest contribution. The signaling model (Spence, 1973) predicts that a credential's value decreases with the precision of employer priors. The literature on employer learning (Altonji and Pierret, 2001) has documented that information asymmetries about worker quality diminish over time, but the institutional mechanism through which third-party certification accelerates that learning has received little direct empirical attention. Our heterogeneity results supply that evidence: accreditation functions as a substitute for reputation, most valuable where reputation is weakest (Hörner, 2002; MacLeod et al., 2017). The institutional gradient is not a secondary robustness check. It is the primary result, because without it the 3 percent average is consistent with many stories, including simple quality improvement. The gradient pins down the channel.

The third question is distributional, and it may be the most consequential for policy. Chile's accreditation system was designed, in part, to reduce information asymmetries in a market with enormous institutional heterogeneity, a challenge shared by higher education regulators across Latin America and much of the developing world (Hastings et al., 2013; Reyes et al., 2016). Our findings reveal an implication of that design that is not obvious ex ante: the wage benefits of accreditation flow to graduates of lower-quality institutions, who are exactly the population that employer reputation alone fails to serve. Accreditation redistributes information rents toward graduates whose credentials would otherwise be discounted. For policymakers designing accreditation systems, this means that extending voluntary accreditation to lower-tier institutions is not merely an exercise in quality assurance; it is a mechanism for expanding labor market access to the graduates who need it most.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background. Section 3 presents the data and sample construction. Section 4 describes the identification strategy and causal interpretation. Section 5 presents the main results and heterogeneity analysis. Section 6 investigates the signaling mechanism. Section 7 concludes.

2 Institutional Background

2.1 An Expansion Without a Map

Chile's higher education system grew faster than anyone's ability to read it. In 1990, roughly 200,000 students were enrolled across the country. By 2018, that number had crossed 1.1 million. The growth was not concentrated in a handful of established universities that employers already knew how to rank. It was dispersed across an ecosystem of new providers, including autonomous universities, professional institutes (*institutos profesionales*), and technical training centers (*centros de formación técnica*), many of which had no institutional history, no observable track record, and no reputation to speak of. A recruiter in Santiago trying to evaluate a graduate of a new regional engineering program in 2005 had access to almost no useful information. The transcript told them a degree had been awarded. It said nothing about whether the program was any good.

The information problem was not symmetrical. Graduates of programs at the Universidad de Chile or the Pontificia Universidad Católica, the country's longstanding elite institutions, arrived in the labor market with a reputation that preceded them. Employers had decades of experience with their graduates, and the institutional brand compressed the uncertainty into something manageable. For graduates of the dozens of programs that had opened since deregulation, that compression was absent. The graduate knew what their training was worth. The employer did not. Economists have a name for this gap. In Akerlof (1970)'s formulation, the lemons problem produces a predictable consequence: without a way to distinguish quality, the market treats all graduates from unrecognized programs as interchangeable, and wages pool downward toward the average. Programs that genuinely invested in teaching and curriculum improvement could not capture the return, because employers could not observe the investment. The incentive to invest was accordingly weak.

2.2 The Commission as an Answer

The 2006 Quality Assurance Act (*Ley de Aseguramiento de la Calidad*) created the National Accreditation Commission (CNA) specifically to attack this problem. The design choice that makes CNA useful for our purposes, and that distinguishes it from institutional-level rankings or government oversight regimes in other countries, is that accreditation operates at the program level. Not the university. Not the faculty. Each undergraduate degree program must apply and be evaluated independently. A university could have a highly regarded law program and an unaccredited business program enrolled in the same building. The signal is specific enough to be informative.

When a program is accredited, the Commission grants a term of two to seven years depending on the quality assessment. A short term signals marginal quality; a longer term signals stronger performance across the Commission's evaluation criteria. Programs must reapply before the term expires to maintain their status; accreditation that lapses becomes visible absence, which employers can also observe. The system thus generates not just a binary certified/uncertified distinction but a graduated signal that is updated at every renewal cycle. Employers who consult the registry consistently can track a program's trajectory over time.

2.3 The Mechanics of the Decision

How the accreditation decision is made matters for identification, so it is worth being precise. The process unfolds in three stages. In the first stage, the program conducts an internal self-evaluation: a structured audit of its own curriculum, faculty qualifications, graduation rates, and graduate outcomes. This document is prepared by the program and submitted to the Commission; it reflects what the program knows and chooses to disclose. In the second stage, an external peer review committee (a panel of subject-matter experts who visit the program, interview faculty and students, and examine facilities) produces an independent report with findings and recommendations. Programs see this report. Faculty discuss it. Administrators adjust their expectations.

The third stage is where the clean variation enters. A panel of CNA commissioners, a standing body with no direct involvement in the peer review visit, reads both the self-evaluation and the external report, deliberates, and issues a ruling: accredited or not, and if accredited, for how many years. This ruling is neither a rubber stamp of the peer review committee's findings nor a mechanical function of the scores in the self-evaluation. Commissioners exercise judgment, weigh competing considerations, and sometimes reach conclusions that differ from what the external panel implied. A program that received a cautiously positive peer review report might be awarded a short accreditation term; a program that received a mixed report might receive a longer one if the commissioners found the self-evaluation evidence of sustained improvement compelling. The gap between what the peer review predicts and what the Commission ultimately decides is precisely where our identifying variation lives. The binary outcome and the length of the term awarded are not perfectly anticipated.

2.4 The Registry as Signal Delivery

A ruling that sits in a government filing cabinet does nothing for wages. What makes the CNA decision a labor market signal is the public registry through which it is delivered. CNA publishes all accreditation decisions (program name, institution, term awarded, expiration date) in a database that is integrated into the official government information portal for higher education (*Sistema de Información de Educación Superior*, SIES). Employers screening candidates can access it directly. University admission platforms surface it to prospective students. Career counselors and guidance offices refer to it when advising students. The registry is not a niche document; it is infrastructure. When a program receives its first CNA accreditation, the public announcement changes what employers can know about its graduates, immediately and at low cost.

This is the mechanism through which an administrative decision becomes a wage shock. The Commission's ruling on a Tuesday afternoon in Santiago determines what a recruiter can find on a Wednesday morning when they search for a program name. Programs that had no certified status before the ruling, whose graduates were, from the employer's perspective, indistinguishable from those of unrecognized programs, now carry a label that is verifiable, standardized, and publicly available. The information asymmetry narrows not because the graduates changed, but because the employer's information set expanded.

2.5 Voluntary Participation and the Source of Timing Variation

For most fields, accreditation is a choice. The 2006 law mandates accreditation only for programs in medicine, dentistry, and education, where professional licensing requirements create a direct legal link between accredited status and practice rights. Our analysis excludes these three fields entirely and focuses on the voluntary sector, where the decision to seek accreditation, and its timing, is driven by program-level calculations about readiness, cost, and strategic positioning rather than by legal obligation.

This voluntary structure is the source of the identifying variation we exploit. Programs in the voluntary sector apply when they believe they are ready to pass, or when competitive pressure from newly accredited rivals makes the absence of a credential increasingly costly. Because programs make this choice sequentially and at different points in time, the first-time accreditation events are staggered across the sample period. The Commission's ruling then introduces a near-exogenous shift in public information about that program's quality. The shift is exogenous not because programs apply randomly, but because the final decision, once an application is filed, is made by an independent panel that the program cannot control. What programs can manage is when they enter the process. What they cannot manage is what the Commission decides when they do. That asymmetry (endogenous entry timing, largely unanticipated ruling) is the foundation of the identification strategy developed in Section 4.

3 Data

3.1 Labor Market Outcomes: SII Administrative Earnings Records

The signaling model studied in this paper is a story about what employers believe and how they update those beliefs. That narrative requires an outcome measure observed on the employer side of the labor market, not the employee side. Formal private-sector employers in Chile report annual earnings to the Internal Revenue Service (SII) for every worker on payroll. These administrative records, linked to higher education graduation files by the Ministry of Finance (*Ministerio de Hacienda*), provide exactly that vantage point. The recruiter whose decision the model describes is in the formal sector by construction; an informal employer making an off-the-books wage offer is not the agent responding to a government-published accreditation registry. The SII records therefore capture the universe of wages that the signaling mechanism can plausibly affect.

For each program–graduation year cell, we observe the number of graduates matched to formal employment and their mean annual taxable earnings one, two, and three years after graduation. This three-horizon structure is a feature, not a convenience: the signaling model predicts that the information advantage of the accreditation label should decay as employers accumulate direct experience with graduates. Measuring wages at multiple post-graduation horizons allows us to trace that learning curve. Graduates in the public sector, the self-employed, and informal employment fall outside the SII universe and are not observed; Appendix A details the linkage procedure, variable definitions, and sample construction criteria.

The estimation dataset matches these wage records to program-level accreditation timing. The full wage panel contains 25,054 program–graduation year observations from 3,429 programs. After the sample restrictions described in Section 4, the primary estimation sample contains 18,826 observations from 2,572 programs: 1,725 programs that received first-time accreditation between 2007 and 2018, and 847 that remained unaccredited throughout. That panel structure sets the stage for the second data layer, which supplies the treatment variation the design requires.

3.2 Accreditation Timing: The CNA Registry

The key to using accreditation as an information shock, rather than as a proxy for institutional quality, is knowing precisely when the shock occurred. Survey data and institutional self-reports cannot provide that precision; programs have strategic incentives to describe their accreditation status generously, and the relevant date for employers is the Commission's ruling, not the moment the program decided to apply. The CNA maintains a public registry that records, for each program, the exact date of the Commission's decision, the binary outcome, and the duration of the term awarded. It is the ruling date that determines when accreditation status becomes publicly verifiable on the SIES portal that employers consult during candidate screening.

For each accredited program in the registry, we record the year of first CNA accreditation as the treatment date G_p. We then construct event time $\kappa_{pt} = t - G_p$ to measure each graduation cohort's position relative to the moment the signal became public. This timing precision is what makes the research design possible. Without it, we could say that accredited programs pay higher wages, which is the cross-sectional finding in González-Velosa et al. (2015), but we could not distinguish the effect of the information shock from the pre-existing quality advantage of programs that chose to seek and successfully obtained accreditation. First-time accreditations are distributed across all years from 2007 to 2018, providing the cross-cohort staggering that the imputation estimator exploits.

The registry also records the length of each accreditation term, which ranges from one to seven years. This variation in term length is not incidental. The Commission's guidelines treat the term as a graded quality signal, with longer terms reserved for programs demonstrating more comprehensive compliance with its standards. We use modal institutional term length to construct the three-tier quality proxy (Baseline: one to three years; Enhanced: four to five years; Top-tier: six to seven years) that drives the heterogeneity analysis in Section 5. Programs' first-time term length serves a parallel role at the program level. Both proxies derive directly from the Commission's own records, making them hard measures of quality standing rather than survey-based perceptions.

3.3 Student Composition: The DEMRE Link

The wage effects estimated in this paper could, in principle, reflect two distinct phenomena: employers updating their beliefs about a program's quality after observing the accreditation signal, or high-ability students selecting into newly accredited programs and mechanically raising average wages through their own human capital. Separating these channels requires data on who attends each program in each cohort. That information comes from DEMRE, Chile's centralized admissions authority, which assigns students to programs through a deferred acceptance algorithm based on standardized test scores (*Prueba de Selección Universitaria*, PSU) and secondary school GPA.

Because all students enter through this centralized system, the DEMRE records contain pre-enrollment characteristics (family income, parental education, and PSU scores) that are observationally clean and available at the program–admission cohort level. Linking admission cohorts to graduation cohorts via SIES allows us to construct program-level measures of student body composition in the pre- and post-accreditation periods. This link does two things for the paper. First, it provides the descriptive statistics that characterize treated and control programs before any treatment occurs, documenting baseline comparability along the student composition dimension. Second, and more consequentially, it enables the composition test in Section 6: do higher-SES or higher-scoring students begin selecting into programs after they receive a long-duration accreditation ($\kappa \geq 0$)? If the wage premium is driven by employer belief updating rather than student sorting, we should not find compositional shifts of that kind, at least not among the cohorts enrolled before the accreditation was announced. The DEMRE link is not a supplementary data source. It is the instrument for ruling out the most credible alternative to the signaling channel.

3.4 Descriptive Statistics

Table 1 compares pre-accreditation cohort characteristics across treated and control programs, restricting to observations before each program's first accreditation event ($\kappa < 0$). The comparison covers 1,193 of the 1,725 treated programs in the estimation sample; the remaining 532 programs received first-time accreditation in 2007 and have no pre-treatment wage observations in the panel. All 847 never-treated programs are included.

The table tells a story that motivates the research design. The most striking difference is in pre-period wages: control programs pay approximately 25 percent more per year than treated programs (CLP 7.5 million versus CLP 6.0 million). A naive reading would treat this as evidence that control programs are simply higher quality. On average they are, but not because of accreditation. The gap reflects the composition of the control group, which is disproportionately drawn from Top-tier institutions. Programs at well-established universities whose graduates carry a recognizable institutional brand were among the last to seek voluntary accreditation, because the marginal informational value of a CNA stamp is low when employers already have strong priors about the institution. These programs are overrepresented in the never-treated group (16 percent Top-tier in the control group versus 10 percent in the treated group), which mechanically inflates the control group's pre-period wage baseline.

Treated programs, by contrast, are concentrated in Enhanced institutions (52 percent): programs at universities with solid but not nationally recognized reputations, for whom the accreditation signal closes a genuine information gap. They are also larger on average, graduating approximately 32 students per cohort compared to 24 for control programs, and they enroll a substantially higher share of women (59 percent versus 46 percent), reflecting field-of-study differences in which sectors pursued voluntary accreditation early. Student body composition along academic preparation is modestly higher in treated programs, as measured by secondary school GPA.

These baseline differences are not a problem for identification; they are the reason for it. The event-study design absorbs all time-invariant differences between programs through unit fixed effects and identifies the accreditation effect from within-program wage changes around the treatment event. What the descriptive table establishes is that treated and control programs occupy different positions in the market hierarchy, and cross-sectional comparisons between them would be confounded by exactly the reputational differences the signaling model places at the center of the analysis. The within-program research design sidesteps that confound, and the DEMRE composition data provide the tools to test whether changes in student sorting are doing any of the work the employer-updating story claims for itself.

4 Empirical Strategy

4.1 Identification and Research Design

The central identification challenge is not establishing that accreditation happened; it is establishing that its timing was unanticipated. If programs knew years in advance exactly when they would receive their first accreditation, they could invest strategically in the pre-accreditation period, and any post-accreditation wage gains would conflate the quality signal with earlier compositional improvements. The staggered difference-in-differences design we use below is built precisely to handle this timing uncertainty.

We estimate the causal effect of first-time program accreditation on graduate wages using a staggered difference-in-differences design. The treatment is defined as a binary indicator for the year a program receives its initial accreditation from the National Accreditation Commission (CNA). Our primary outcomes are the log annual earnings of graduates measured one, two, and three years after labor market entry.

The signaling mechanism implies that the relevant treatment timing depends on the graduation cohort relative to the public announcement. We define event time κ_{pt} as the difference between graduation year t and the year of first accreditation G_p:

$$\kappa_{pt} = t - G_p \tag{1}$$

where $G_p = \infty$ for never-treated programs. This specification captures the information available to employers at the time of hiring. Because firms observe the accreditation status of the degree-granting program, the wage effects should accrue to cohorts graduating on or after the date the accreditation is publicly awarded ($\kappa \geq 0$). Crucially, all graduates in the early post-treatment window ($\kappa \in \{1, 2, 3, 4\}$) were enrolled before accreditation was announced, so any wage gains they experience can only reflect employer updating on the new signal, not selection of higher-ability students into the newly accredited program. This feature of the design isolates the pure signaling channel before composition effects can appear.

4.2 Sample Construction

Comparability between treated and control units is not automatic; it must be built into the sample before estimation. A program at an unaccredited institution competes in a different financial-aid environment than one at an accredited institution; a mandatory-accreditation program faces a different incentive structure than a voluntary one. Both differences would confound the wage comparison even if treatment timing were otherwise clean.

We therefore refine the estimation sample on three dimensions. First, we restrict the analysis to programs at institutions that achieved institutional-level accreditation. This ensures that all programs in our sample are eligible for state-funded student loans, removing confounding shifts in student composition driven by financial aid eligibility. Second, we exclude programs in fields where accreditation is legally mandated (education, medicine, and dentistry) to ensure that treatment assignment reflects program-level quality signals rather than regulatory requirements. Finally, we include only programs that received their first accreditation between 2007 and 2018, ensuring a sufficient window to observe post-treatment labor market outcomes. The resulting panel comprises 18,826 observations from 2,572 programs. These restrictions leave a sample where variation in treatment timing is driven by voluntary program choices and commission decisions, not by institutional or regulatory constraints.

4.3 Estimation and Inference

With a clean sample in hand, the remaining challenge is estimating the effect of accreditation without importing the biases that plague the standard two-way fixed effects estimator in staggered settings. When treatment effects are heterogeneous across cohorts, as the signaling model predicts they will be, the traditional estimator uses early-treated units as implicit controls for later-treated ones, producing estimates that can be negatively weighted and misleading (Sun and Abraham, 2021; **?**). We avoid this entirely by using the imputation estimator proposed by Borusyak et al. (2024).

The imputation approach proceeds in two steps. We first estimate a unit-and-period fixed effects model on the sample of never-treated and not-yet-treated observations:

$$Y_{pt} = \alpha_p + \lambda_t + \varepsilon_{pt} \tag{2}$$

We then use the estimated parameters to impute counterfactual outcomes $\hat{Y}_{pt}^{(0)}$ for treated observations. The average treatment effect on the treated (ATT) for each event-time κ is the mean difference between observed and imputed outcomes:

$$\widehat{\text{ATT}}(\kappa) = \frac{1}{N_\kappa} \sum_{(p,t):t-G_p=\kappa} \left[Y_{pt} - \hat{Y}_{pt}^{(0)} \right] \tag{3}$$

This estimator provides a robust measure of the wage premium even if the returns to accreditation evolve over time. We cluster standard errors at the program level to account for serial correlation in wages within degrees.

The event-study profile traces the ATT at each relative period, but the headline estimate reported in Table 2 is a scalar ATT that pools across all cohorts and post-treatment event times. Following Borusyak et al. (2024), this scalar is a weighted average of the cohort-by-event-time specific estimates $\widehat{\text{ATT}}(\kappa)$, where the weights are proportional to the number of treated observations in each cohort-by-event-time cell. This aggregation rule down-weights cells with few graduates and up-weights those with more, so the scalar ATT reflects where most of the identifying variation lies. It is this pooled estimate, rather than any single event-time coefficient, that we interpret as the average early-career wage premium attributable to first-time accreditation.

4.4 Identifying Assumptions

4.4.1 Parallel Trends

The internal validity of any difference-in-differences design rests on parallel trends, but parallel trends is an assumption that cannot be directly tested; it can only be made plausible. The event study pre-period provides the most direct evidence. If treated and control programs were already on diverging wage trajectories before accreditation was announced, the pre-period coefficients would be nonzero and trending. If instead the pre-period is flat, it supports, though does not prove, that the post-period divergence is attributable to the accreditation shock rather than a pre-existing difference.

We test this assumption by examining the pre-treatment coefficients for $\kappa \in \{-4, -3, -2\}$. Across all wage outcomes, we fail to reject the null hypothesis of zero pre-trends ($p > 0.60$), confirming that treated and control programs were on parallel wage trajectories prior to the accreditation shock. We normalize the event study to $\kappa = -1$, the final graduation cohort unaffected by the public signal. The institutional details of the CNA process further support a causal interpretation: the final accreditation decision is determined by an independent commission whose ruling is not fully predicted by the preceding peer-review reports, introducing a degree of exogenous variation in the precise timing of the signal.

4.4.2 No Anticipation

Parallel trends alone is not sufficient for identification. The design also requires that neither programs nor the relevant labor market actors behave differently in anticipation of a treatment date they know in advance, a condition violated if accreditation generates behavioral responses before the formal announcement.

On the program side, the no-anticipation condition is straightforward to motivate. Programs endogenously decide *when* to apply for accreditation, but the final outcome (both whether they are accredited and for how many years) is determined by the CNA Commission after an independent review. A program that submits an application cannot know with certainty whether it will be approved, nor can it anticipate the specific term length. Because the award date is uncertain at the time of application, programs have no sharply anticipated treatment date around which to reorganize hiring, curriculum, or marketing. Any gradual quality investments that programs may undertake in preparation for the peer-review process are precisely what the parallel trends assumption absorbs; the flat pre-trends across wage outcomes provide direct evidence that no such differential investment was occurring.

On the employer side, a subtler version of the no-anticipation concern arises. CNA site visits involve external evaluators drawn from academia and industry, and employers who serve on these panels observe the program under review before the Commission issues its ruling. In principle, such employers could update their wage-setting for graduates of evaluated programs before the formal announcement is public. Two features of the Chilean labor market rule this out as a threat to identification. First, labor markets are large relative to any single employer: the wage of a program's graduates is determined by market-wide beliefs, and no single firm can unilaterally shift equilibrium wages by updating its own priors from a peer visit. Second, employers who serve as CNA evaluators do so to provide input to a national quality commission; the peer visit is a regulatory exercise, not a recruiting opportunity. These evaluators have no systematic incentive to favor or penalize specific programs in their own hiring. The absence of wage movement during the peer-review window, the period between site visits and the public announcement, confirms that the labor market responds to the formal CNA ruling rather than to informal information from the evaluation process.

5 Main Results

5.1 Average Effects of Accreditation on Wages

Initial program accreditation generates a wage premium that emerges immediately and grows as successive cohorts of graduates enter the labor market. Figure 1 presents the event study estimates for log annual earnings one, two, and three years after graduation. Pre-period coefficients are statistically flat, with $p > 0.60$, confirming that treated and control programs were on parallel trajectories before the accreditation announcement. What follows that flat baseline is not a modest drift upward: it is a sustained divergence that compounds across the post-accreditation window, reaching 6–15 percent between five and eight years after the event. For a graduate earning at the control mean, that range represents a five-figure annual difference in wages by mid-career.

The mechanism has a natural two-phase structure. In the early window ($\kappa \in \{1, 2, 3, 4\}$), all graduates were enrolled before accreditation was awarded, so any premium reflects employer updating on the signal alone. The program's quality is unchanged; only what employers know about it has changed. At longer horizons, the growing premium is consistent with a second channel: accreditation attracting higher-ability applicants who enrolled post-decision, gradually shifting the composition of the graduate pool. Table 2 reports the pooled ATT. Averaged across wage horizons, the estimated premium is approximately 3 percent, reaching statistical significance at conventional levels in the second and third years of employment (coefficients of 0.029 and 0.024, respectively). But these pooled averages are the least interesting number in the paper; they conceal a pattern that the signaling model predicted and the data confirm with unusual precision.

5.2 Heterogeneity by Institutional Quality

The central prediction of the signaling framework is that accreditation should matter most where employers know the least. We test this by estimating subgroup ATTs stratified by the institution's accreditation tier. The result in Table 3 is not a gradient; it is a reversal.

Graduates of *Baseline* institutions realize an 8.9 percent wage premium in their first year of employment. The certification converts a noisy employer prior into a verified quality floor, and the market responds immediately. For graduates of *Enhanced* institutions, the premium shrinks to -2.7 percent. For graduates of *Top-tier* institutions, it reaches -5.0 percent. The signal does not merely fade as institutional prestige rises; it flips sign.

This is precisely what a Bayesian model of employer learning implies. At *Top-tier* institutions, employers already hold precise beliefs about graduate quality, beliefs formed through years of direct hiring and institutional reputation. When a program at such an institution finally receives CNA certification, the announcement carries almost no information about quality that employers did not already have, and it may carry unwanted information about the program's relative standing. A *Top-tier* institution that needed to go through the formal accreditation process, and received a term shorter than its peers, may signal to employers that something was previously wrong. Accreditation, for these graduates, is not a credential that opens doors; it is evidence that a door was once closed. Figure 2 traces the diverging event study profiles: the *Baseline*-tier effect grows monotonically over the post-accreditation window, while the *Top-tier* effect remains negative or null throughout the entire horizon.

The synthesis is blunt: the wage return to accreditation is not a fixed feature of the certification; it is a function of what the certification tells an employer that they did not already know. Where prior uncertainty is high, the signal is worth nearly 9 percent in wages. Where prior uncertainty is low, the signal costs 5 percent. The CNA label is not inherently valuable; its value is entirely in the information it resolves.

5.3 Heterogeneity by Program-Level Treatment Intensity

If the informational content of the signal drives returns, then programs whose first accreditation resolves deeper uncertainty, those that received shorter terms signaling a tighter quality margin at the evaluation threshold, should generate the largest wage response. The duration of the accreditation term awarded at first evaluation maps directly onto this prior: a short term certifies quality at the minimum standard; a long term certifies quality well above it. Table 4 tests whether this gradient in informational content produces a gradient in wage returns.

Programs receiving a *Low*-intensity first accreditation (1–3 years) show a first-year wage premium of 10.7 percent ($p < 0.01$), which remains statistically significant through the third year of employment (5.4 percent). Programs receiving *Mid*-intensity accreditation (4–5 years) show a smaller and less persistent premium. Programs receiving the *High*-intensity term (6–7 years) show no significant wage gain. The monotone decline across intensity tiers is exactly what the theory requires.

Together, the institutional and program-level stratifications tell a single, consistent story: the CNA registry functions as a corrective informational tool that disproportionately benefits graduates from programs where employer priors are weakest. A low-intensity first accreditation moves an employer from a prior of high uncertainty to a state of positive, if minimal, certification, and the market prices that resolution immediately and durably. A high-intensity accreditation at a program embedded in a well-regarded institution adds nothing that reputation had not already communicated. Figure 3 presents the corresponding event study profiles.

6 Mechanism: Signaling versus Human Capital

6.1 Enrollment Timing and the Composition of Graduates

A concern with our wage estimates is that accreditation may change the composition of graduates observed in the employment data, rather than raising wages for the existing graduate pool. If institutions game their graduation records around accreditation, producing more graduates to improve quality metrics, the additional graduates might have different wage prospects, biasing our estimates.

More broadly, even absent gaming, accreditation may attract higher-ability students who enrolled *after* learning about the program's accredited status.

A straightforward timing argument disciplines the selection concern. The composition of graduates entering the labor market at event time κ was determined by enrollment decisions of students who entered the program at $\kappa - D$, where D is program duration, approximately four to five years for university programs in Chile. Students who enrolled *because* they observed an accredited program can therefore only appear in the labor market at event times $\kappa \geq D$. At $\kappa \in \{1, 2, 3, 4\}$, the window in which we observe significant wage effects, the graduates were enrolled *before* the program received accreditation and could not have been selected by it. That the event study delivers precisely here, and not earlier, is not incidental: $\kappa \in \{1, 2, 3, 4\}$ is exactly the window the theory predicts should reflect pure employer updating on a new public signal, and the estimates confirm it.

6.2 Graduation Counts, Employment, and On-Time Completion

We estimate event studies for three supplementary outcomes that characterize institutional responses around accreditation: (i) the log of total graduates per program–graduation year cell (log(total grads)); (ii) the share of all SIES graduates found in private-sector employment records within two years of graduation (emp rate = alum/total grads); and (iii) the on-time graduation rate among the enrollment cohort expected to complete in each year (on-time rate). Figure 4 presents the three event studies side by side.

Total graduates. Treated programs have *fewer* graduates than the never-treated control group in the four years preceding accreditation, a pattern that constitutes a clear violation of the parallel trends assumption. The careful reader should pause here: something is happening before treatment that breaks parallel trends for this outcome, and that something has a name. Programs seeking CNA accreditation face strong incentives to demonstrate completion rates and graduate output during the preparation window, creating pressure to accelerate completions before the review. Students who had delayed their thesis defense or final examinations are encouraged, informally or otherwise, to finish in time for the program's self-evaluation report. This is anticipation effects in action, and they contaminate the pre-period for graduation counts in a way that makes this outcome unsuitable for causal inference under standard parallel trends. The wage result survives for a simple reason: employers observe wages, not graduation counts, and the CNA signal reaches them only after the formal decision is announced.

Employment rate. Before accreditation, treated programs exhibit *higher* formal employment rates relative to controls, with statistically significant positive estimates at $\kappa = -3$ and $\kappa = -2$. This reflects positive selection in the pre-accreditation graduate pool: when total graduation counts are depressed, the students who complete tend to be the most able, and they find formal employment at above-average rates. After accreditation, employment rates fall below those of the control group. The post-accreditation decline is not evidence that accreditation fails graduates in the labor market. It is evidence that accreditation induces completion among students who might not otherwise have finished, students who, on average, enter occupations or sectors that do not appear in SII formal-sector payroll records, whether because they are self-employed, work informally, or move abroad. The denominator expands faster than the numerator, reducing the rate. Crucially, the wage sample is drawn from the employed subsample: workers found in formal-sector records. As accreditation pulls marginal completers into the graduate pool who sort away from formal employment, the wage sample becomes increasingly selected toward the kinds of workers who were always likely to find formal jobs. This selection, if anything, makes the wage estimates conservative; the marginal completer induced by accreditation is not represented in the wage data.

On-time graduation rate. The on-time graduation rate, defined as the share of an enrollment cohort completing within the nominal program duration, spikes in the periods immediately preceding accreditation and then flattens after the accreditation decision is issued, remaining at that level thereafter. This pattern is consistent with institutional gaming during the CNA preparation window: programs accelerate completions to improve evaluation metrics, producing a temporary spike that subsides once the evaluation is resolved. The post-accreditation on-time rate does not recover, which implies that the additional graduates observed in the left panel are predominantly older enrolled cohorts choosing to finish rather than new, higher-ability cohorts completing on schedule. A companion analysis documents parallel pre-trends in student composition (test scores, secondary school grades, and family income) and finds no evidence of selective enrollment before accreditation, directly supporting this interpretation.[1]

Synthesis. The three panels tell a coherent story. Before accreditation, treated programs produce fewer graduates, but those graduates are positively selected and find formal employment readily. As programs prepare for CNA evaluation, enrolled students, including delayed completers, are pushed to finish, temporarily raising both total graduation counts and on-time rates. After accreditation, total graduates remain elevated but the on-time rate flattens: the marginal graduate is not a strong on-schedule completer. These marginal graduates dilute the employment rate without contributing proportionally to formal-sector payrolls. The graduates at event times $\kappa \in \{1, 2, 3, 4\}$ who generate the wage effects we estimate belong to cohorts enrolled before accreditation; they are not the product of compositional improvement. The wage increase is attributable to employer updating on a new public signal, not to a change in who is graduating, and the supplementary evidence leaves no other interpretation standing.

6.3 What the Wage Result Tells Us

The wage event studies satisfy the parallel trends assumption ($\chi^2(3)$ p-values of 0.73, 0.61, and 0.65 for log wages at years 1, 2, and 3 after graduation, respectively) and yield a clean causal interpretation. The key distinction from the graduation and employment outcomes is the transmission mechanism: employers update their wage offers only *after* the formal accreditation decision is publicly announced, because the signal is credible and publicly verifiable only at that point. There is no analogous employer-side channel operating in the preparation window, which is why the wage pre-trends are flat while the graduation and employment pre-trends are not.

A residual concern is that the growing graduate pool progressively selects the employment sample over time. As later post-accreditation cohorts, who enrolled knowing the program's accredited status, begin to graduate (at $\kappa \geq D$, where D is program duration), the composition of the labor market sample improves. This compositional shift could, in principle, explain part of the rising wage profile at longer horizons ($\kappa \geq 5$). Our preferred estimates at $\kappa \in \{1, 2, 3, 4\}$ predate this channel: the graduates

[1] Molina and Valdebenito, in progress. Available from the authors upon request.

were enrolled before accreditation, could not have been selected by it, and the on-time graduation evidence confirms they are not systematically stronger completers than earlier cohorts. The wage effect is real, the mechanism is employer updating, and the gaming threat has been handled.

7 Conclusion

We provide causal evidence that first-time program accreditation by Chile's CNA raises graduates' wages by approximately 3 percent on average, with the effect growing over the years following graduation. The key finding, however, is the sharp heterogeneity: the wage premium is large and significant for graduates of lower-quality institutions (approximately 9 percent) but is negative for graduates of the highest-quality institutions. This pattern aligns precisely with the predictions of the Spence signaling model: accreditation carries the most informational content, and thus the largest wage effect, where employer priors are least informative.

The wage effects are causal under the parallel trends assumption, which is supported by joint pre-trend tests ($p > 0.6$ for all three wage horizons). The design exploits the timing of CNA's public accreditation decision as an information shock to the labor market; the endogenous decision by programs to apply does not confound the estimates because we condition on the event of treatment and measure responses relative to the announcement date. Notably, the wage effects are concentrated in the post-accreditation window $\kappa \in \{1, 2, 3, 4\}$, when all graduates in the data were enrolled *before* accreditation was awarded. Compositional selection, meaning better students choosing an accredited program, cannot have affected these cohorts, which rules out the most direct alternative to the employer-updating interpretation. The supplementary analysis of graduation counts and employment rates reveals important nuance. Total graduation counts rise in the post-accreditation period, but the pre-trend violations in this outcome are consistent with anticipation effects: programs preparing for CNA evaluation likely accelerate graduation before the formal decision, contaminating the pre-period window and preventing a clean causal attribution of the graduation increase to the accreditation event itself. The formal employment rate falls as the denominator grows faster than the numerator, a mechanical consequence of graduation inflation rather than an independent result about labor market absorption. The wage premium is unaffected by these dynamics: it reflects higher earnings conditional on formal employment, established using pre-trends that are statistically flat, and concentrated in the early post-accreditation window when graduates were enrolled before the program's certified status was established.

These findings carry specific implications for how accreditation systems should be designed, not just evaluated. Chile's CNA operates on a voluntary basis: programs choose whether to seek certification, and the regulatory agency has limited capacity to pursue programs that do not apply. This architecture creates a predictable imbalance. High-quality institutions enter the system early; they are the easiest to certify and stand to gain the most reputationally from the label in contexts where reputation is already contested. Lower-tier institutions, precisely those for which accreditation generates the largest wage returns to graduates, have the most to gain from certification and, on average, the least organizational capacity to prepare for external evaluation. The result is that a voluntary system, left unsubsidized, may deliver quality assurance primarily where it is least needed. Policymakers who want accreditation to function as a signaling subsidy, shifting information rents toward graduates who cannot rely on institutional prestige alone, must actively encourage or subsidize lower-tier programs to seek evaluation. The alternative, treating accreditation as a pure quality assurance exercise for the already-strong, forfeits the distributional benefit that our heterogeneity results show is the system's largest return. Whether tiered certification schemes can capture this benefit without becoming capture-prone administrative burdens is an open design question, but the evidence here clarifies what is at stake in answering it.

A natural extension of this work is to decompose the total wage effect into three distinct channels. The first is the pure signal effect: the wage premium attributable to the accreditation label alone, holding graduate composition fixed, identified in the early post-accreditation window when enrolled cohorts predate the certification decision. The second is the composition response: as accreditation attracts higher-ability applicants to newly certified programs, the quality of the graduate pool improves, and employers respond to this improvement. This channel can be isolated by tracking the evolution of graduate characteristics (family income, parental education, and prior academic preparation) over the event window; a rise in these measures at $\kappa \geq 5$ but not at $\kappa \in \{1, 2, 3, 4\}$ directly identifies when compositional change begins and how much of the wage growth it explains. The DEMRE applicant microdata, linked at the individual level to the CNA accreditation timeline, already permit this test.

The third channel is program quality improvement: the wage premium that remains after compositional change is accounted for, reflecting improvements to curriculum, faculty, or resources that the accreditation process incentivized. This component can be identified by comparing late-window wage growth between programs whose post-accreditation graduate composition is stable and those whose composition shifts, a contrast that holds constant the signal mechanism and isolates the institutional investment response. Together, the three channels answer a question that the reduced-form estimate cannot: whether accreditation functions primarily as a signal, as a selection device, or as a direct stimulus to institutional investment, a distinction with substantial implications for the optimal design of quality assurance policy.

Table 1: Descriptive Statistics: Pre-Accreditation Cohorts

	Treated	Control	Difference
Panel A: Labor Market Outcomes			
Graduation cohort size	32.395	23.659	8.736***
	(37.949)	(38.795)	(0.826)
Employment rate, year 1	0.821	0.798	0.023***
	(0.233)	(0.245)	(0.005)
Annual earnings, yr. 1 (CLP)	5959.493	7459.020	-1499.526***
	(3613.547)	(4572.991)	(89.660)
Panel B: Student Characteristics			
PSU score (math-verbal)	520.714	516.325	4.389**
	(89.589)	(72.766)	(1.832)
GPA score	582.142	559.617	22.525***
	(66.470)	(61.832)	(1.448)
Male	0.408	0.539	-0.131***
	(0.320)	(0.310)	(0.007)
High income	0.203	0.180	0.024***
	(0.263)	(0.223)	(0.005)
College-educated parents	0.451	0.402	0.049***
	(0.288)	(0.264)	(0.006)
Panel C: Institutional Quality			
Baseline Inst.	0.355	0.347	0.009
	(0.479)	(0.476)	(0.010)
Enhanced Inst.	0.516	0.420	0.095***
	(0.500)	(0.494)	(0.011)
Top-tier Inst.	0.103	0.160	-0.056***
	(0.304)	(0.366)	(0.007)
N (Programs)	1193	847	

Note: This table displays the means of program-level characteristics for the Control (programs never accredited during the sample period) and Treated (eventually-accredited programs, observed in pre-accreditation cohorts) groups. Standard deviations are reported in parentheses beneath the means. The unit of observation is the academic program. The "Difference" column reports the coefficient from an OLS regression of the variable on a treatment status indicator. The N (Programs) row counts programs with at least one pre-accreditation wage observation. Institution quality tiers are defined by CNA accreditation duration: Baseline (1–3 years), Enhanced (4–5 years), Top-tier (6–7 years). Significance levels are denoted as follows: * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Effect of First-Time Accreditation on Graduate Wages

	Log wage yr 1	Log wage yr 2	Log wage yr 3
Accreditation effect	0.030	0.029*	0.024*
	(0.022)	(0.017)	(0.014)
Num. Obs	18672	18707	18708
Control mean	15.609	15.905	16.099

Note: Standard errors (clustered at the program level) are in parentheses. The control outcome mean represents the average log wage in the control group for each outcome. All programs are offered by accredited institutions. The treatment group includes programs accredited for the first time. Average treatment effects on the treated follow Borusyak et al. (2024). * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Heterogeneous Effects by Institutional Accreditation Quality

	Log wage yr 1	Log wage yr 2	Log wage yr 3
Accreditation × Baseline inst.	0.089***	0.046***	0.026***
	(0.017)	(0.011)	(0.009)
Accreditation × Enhanced inst.	-0.027***	-0.003	0.008
	(0.010)	(0.009)	(0.008)
Accreditation × Top-tier inst.	-0.050***	-0.028***	-0.024***
	(0.011)	(0.010)	(0.008)
Num. Obs	18672	18707	18708
Control Mean (CLP thousand)			
Baseline inst.	6057.3	7936.2	9592.7
Enhanced inst.	5996.4	8124.3	9966.6
Top-tier inst.	7135.4	9316.7	11080.8

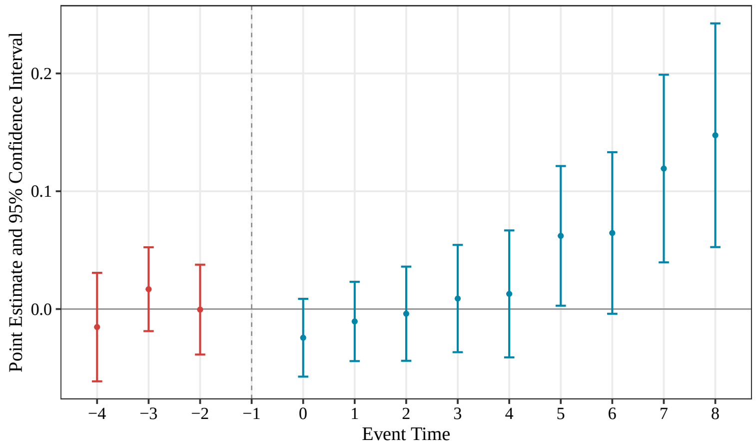
Note: This table displays heterogeneous effects of first-time accreditation on graduate wages by institutional quality. Standard errors (clustered at the program level) are in parentheses. The unit of observation is the academic program. Each row reports the average treatment effect weighted by the share of treated observations within that institutional tier, estimated following Borusyak et al. (2024) with program and graduation-year fixed effects. Institutional quality is defined by the modal CNA accreditation duration of the host institution: Baseline (1–3 years), Enhanced (4–5 years), Top-tier (6–7 years). Control means are in CLP thousands and correspond to never-accredited programs within each institutional tier. Significance levels are denoted as follows: * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Treatment Intensity: Effects by Years Awarded at First Accreditation

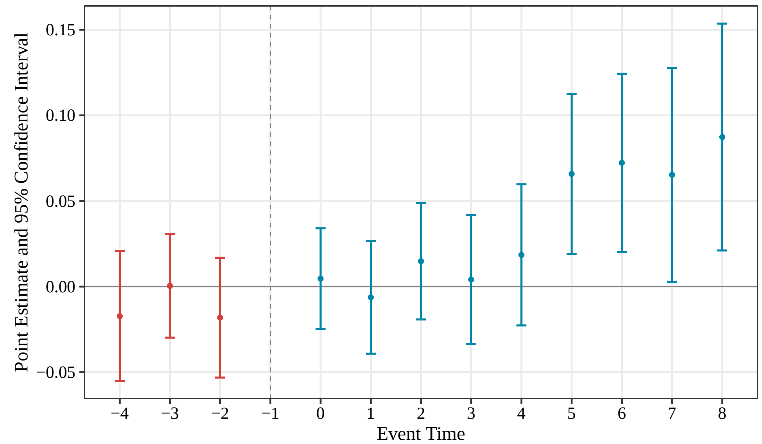
	Log wage yr 1	Log wage yr 2	Log wage yr 3
Treatment Intensity			
Accreditation × Low (1-3 years)	0.107***	0.073***	0.054***
	(0.026)	(0.018)	(0.014)
Accreditation × Mid (4-5 years)	-0.042***	-0.014	-0.008
	(0.012)	(0.011)	(0.009)
Accreditation × High (6-7 years)	-0.015	-0.024**	-0.012
	(0.014)	(0.012)	(0.012)
Num. Obs	18672	18707	18708
Control mean (thousands CLP)	6012.6	8078.4	9806.1

Note: This table displays the effect of first-time accreditation on graduate wages by treatment intensity. Standard errors (clustered at the program level) are in parentheses. The unit of observation is the academic program. Treatment intensity is defined by the number of years awarded at first accreditation: Low (1–3 years), Mid (4–5 years), High (6–7 years). Each row reports the average treatment effect weighted by the share of treated observations within that intensity group, estimated following Borusyak et al. (2024) with program and graduation-year fixed effects. The control mean is the average wage among never-accredited programs, reported in CLP thousands. Significance levels are denoted as follows: * p<0.1, ** p<0.05, *** p<0.01.

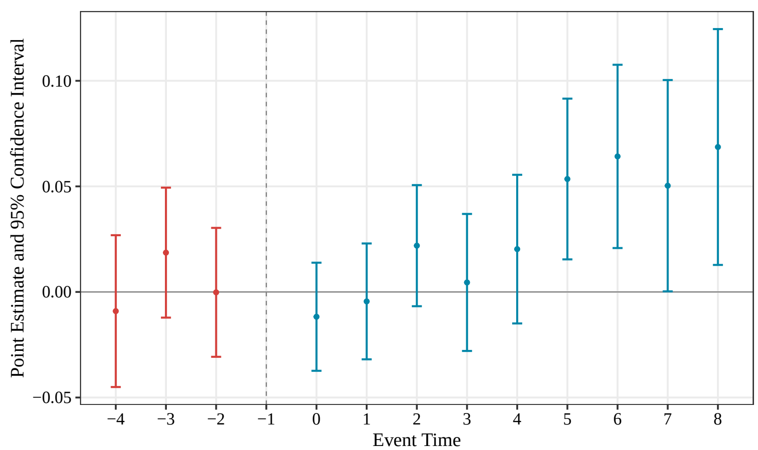
Figure 1: Event study estimates: Wages of employed grads



(a) $1^{st} Year$



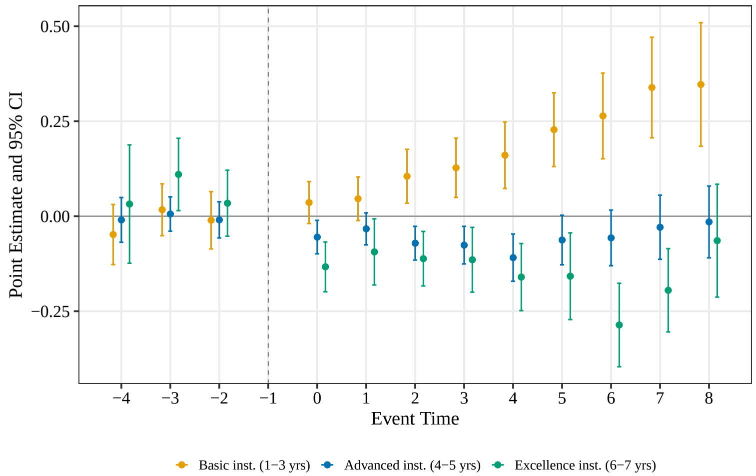
(b) $2^{nd} Year$



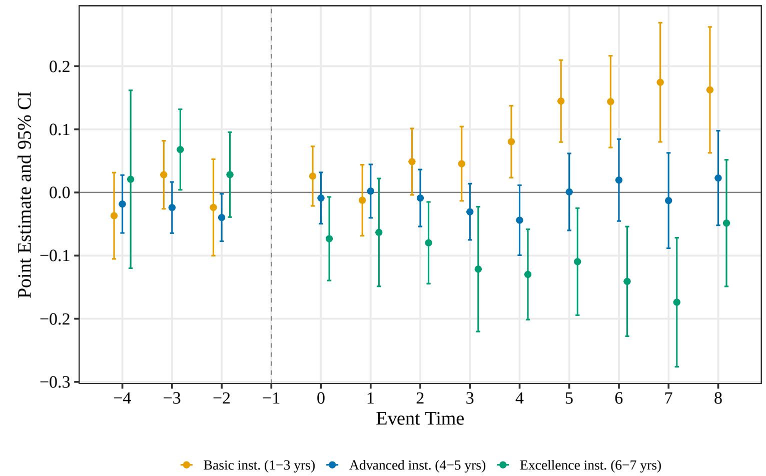
(c) $3^{rd} Year$

Note: This figure plots event study estimates of the effect of first-time accreditation on log wages of graduates (conditional on being employed): (a) 1 year, (b) 2 years, and (c) 3 years after graduation. $\kappa = -1$ is the reference period. Vertical bars represent 95% confidence intervals. Estimates follow Borusyak et al. (2024) with standard errors clustered at the program level.

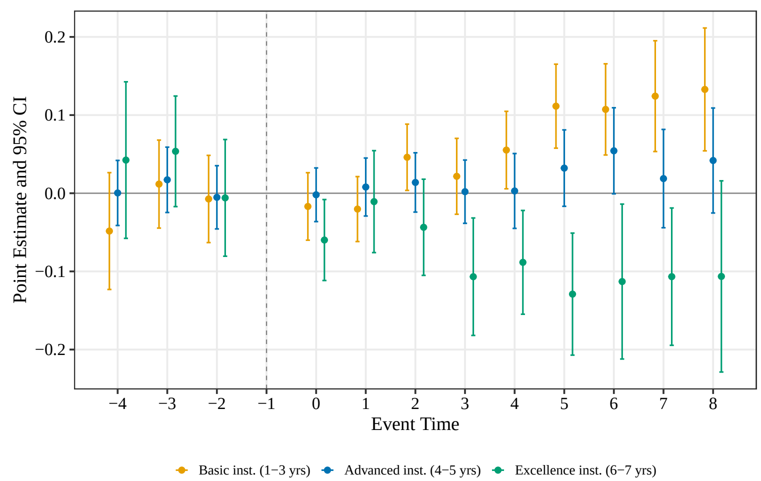
Figure 2: Event study estimates: Wages of employed grads by institutional tier



(a) $1^{st} Year$



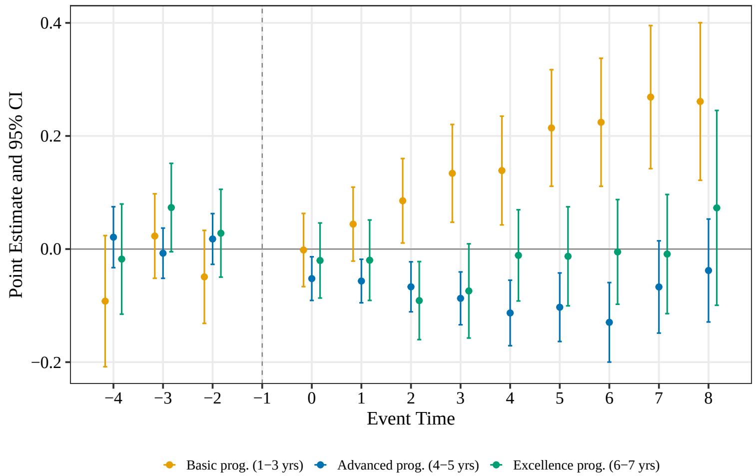
(b) $2^{nd} Year$



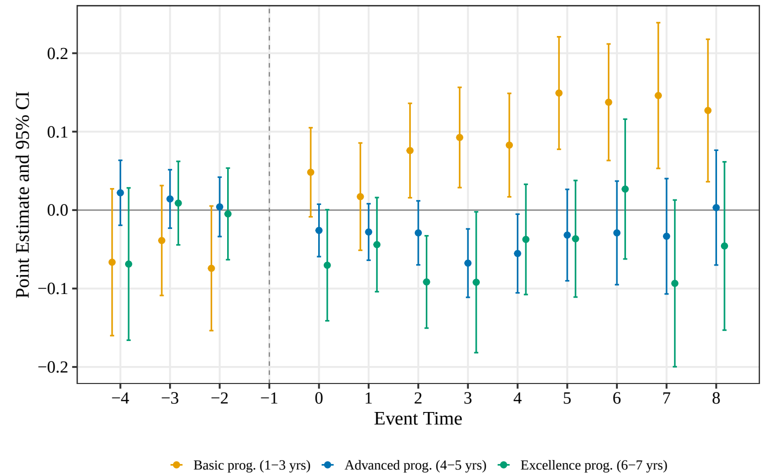
(c) $3^{rd} Year$

Note: This figure plots event study estimates of the effect of first-time accreditation on log wages of graduates by institutional quality tier. Panels (a), (b), and (c) show estimates for wages after 1, 2, and 3 years after graduation. Colors indicate institutional tier: baseline (yellow), enhanced (blue), toptier (green). $\kappa = -1$ is the reference period. Vertical bars represent 95% confidence intervals. Estimates follow Borusyak et al. (2024) with standard errors clustered at the program level.

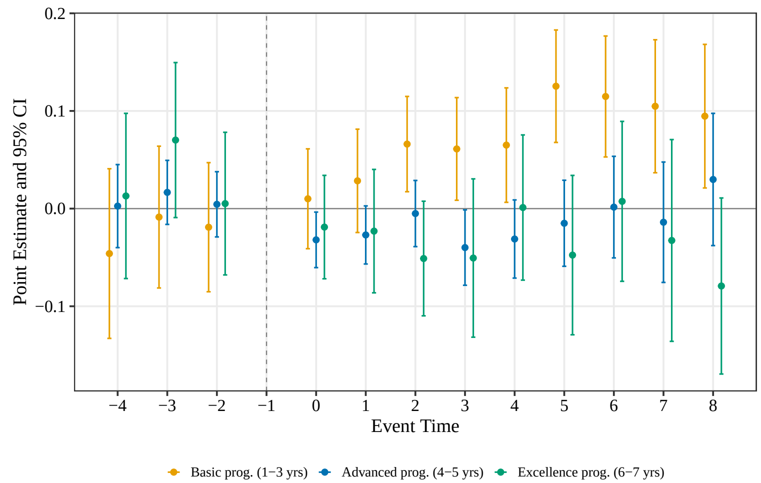
Figure 3: Event study estimates: Wages of employed grads, by treatment intensity



(a) $1^{st} Year$



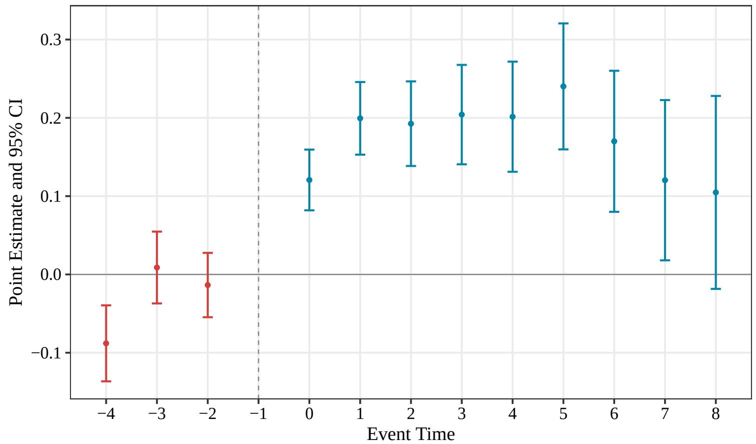
(b) $2^{nd} Year$



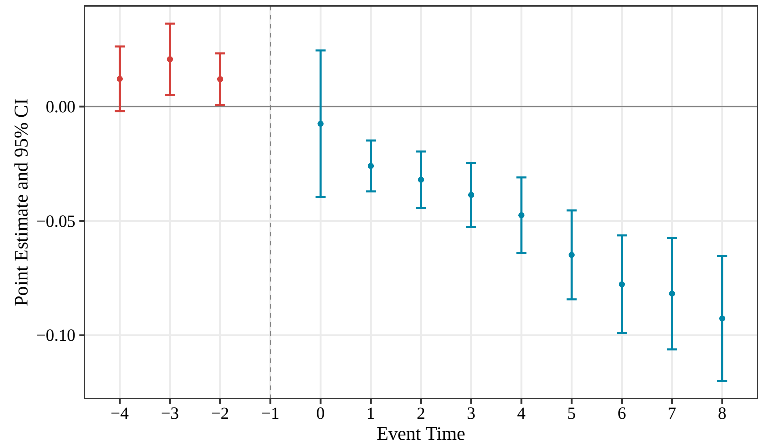
(c) $3^{rd} Year$

Note: This figure plots event study estimates of the effect of first-time accreditation on log wages of graduates by program treatment intensity. Panels (a), (b), and (c) show estimates for wages after 1, 2, and 3 years after graduation. Colors indicate first-accreditation intensity tier: low (yellow), mid (blue), high (green). $\kappa = -1$ is the reference period. Vertical bars represent 95% confidence intervals. Estimates follow Borusyak et al. (2024) with standard errors clustered at the program level.

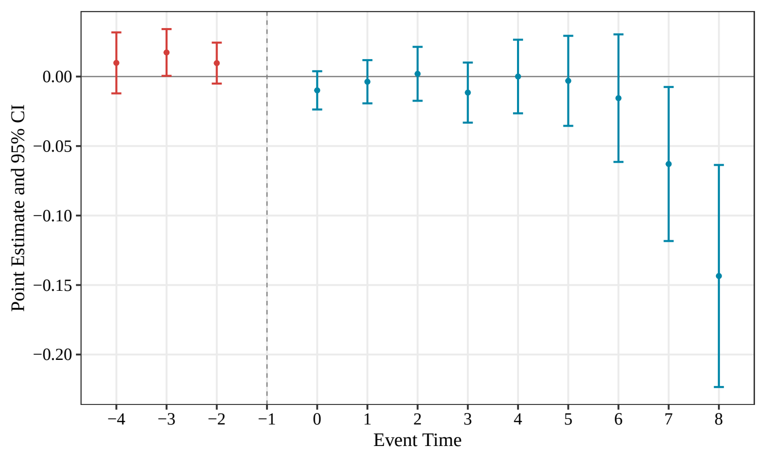
Figure 4: Event study estimates: Mechanism outcomes



(a) Graduation cohort size



(b) Employment rate at 2^{nd} year



(c) On-time graduation rate

Note: This figure plots event study estimates of the effect of first-time accreditation on three mechanism outcomes: (a) log graduation cohort size (total graduates per program–graduation year cell), (b) formal employment rate two years after graduation (employed graduates divided by total SIES graduates), and (c) on-time graduation rate (share of the enrollment cohort completing within the nominal program duration). $\kappa = -1$ is the reference period. Vertical bars represent 95% confidence intervals. Estimates follow Borusyak et al. (2024) with standard errors clustered at the program level.

References

Akerlof, George A., "The market for "lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event study designs: Robust and efficient estimation," *Review of Economic Studies*, January 2024.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Dale, Stacy Berg and Alan B. Krueger, "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 2002, *117* (4), 1491–1527.

Dranove, David and Ginger Zhe Jin, "Quality disclosure and certification: Theory and practice," *Journal of Economic Literature*, December 2010, *48* (4), 935–963.

González-Velosa, Carolina, Graciana Rucci, Miguel Sarzosa, and Sergio Urzúa, "Returns to Higher Education in Chile and Colombia," IDB Working Paper IDB-WP-587, Inter-American Development Bank 2015.

Hastings, Justine S., Christopher A. Neilson, and Seth D. Zimmerman, "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," July 2013.

Hoekstra, Mark, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, 2009, *91* (4), 717–724.

Hörner, Johannes, "Reputation and competition," *American Economic Review*, June 2002, *92* (3), 644–663.

MacLeod, W. Bentley and Miguel Urquiola, "Reputation and school competition," *American Economic Review*, November 2015, *105* (11), 3471–3488.

_ , Evan Riehl, Juan E. Saavedra, and Miguel Urquiola, "The Big Sort: College Reputation and Labor Market Outcomes," *American Economic Journal: Applied Economics*, 2017, *9* (3), 223–261.

Reyes, Loreto, Jorge Rodríguez, and Sergio Urzúa, "Heterogeneous Economic Returns to Postsecondary Degrees: Evidence from Chile," *Journal of Human Resources*, 2016, *51* (2), 416–460.

Spence, Michael, "Job market signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, December 2021, *225* (2), 175–199.

A Wage Data: Record Linkage and Variable Definitions

B Additional Tables and Figures

B.1 Cohort Distribution

B.2 Subgroup Event Studies

Figures **??**–**??** present the full event study estimates for each subgroup separately (Sections 11–12 of the estimation code).

C Robustness Checks

C.1 Minimum Cell Size

Our estimation dataset is a panel of program–graduation year cells, where the outcome in each cell is the *average* log annual wage across all graduates in that cell. Cells with very few graduates produce noisy averages: a single outlier moves the cell mean substantially, inflating residual variance and potentially introducing measurement error into the outcome. We investigate the sensitivity of our estimates to minimum cell-size requirements.

The SIES wage data already applies an implicit threshold: no cell in our sample has fewer than 4 graduates. We assess two additional restrictions:

- **alum** ≥ 5: drops cells with exactly 4 graduates (1,028 observations, 5.5% of the baseline sample of 18,826). The number of treated programs falls from 1,725 to 1,693; never-treated from 847 to 790.
- **alum** ≥ 10: the standard "at least 10 observations per cell" threshold used in aggregated panel studies (see, e.g., **?**). Drops 5,056 observations (26.9%). Treated programs fall to 1,547; never-treated to 579. We note that never-treated control programs are smaller on average (median 11 graduates per cell) than treated programs (median 22), so this trim hits the control group proportionally harder.

Table 5 reports the mean post-period ATT—the simple average of $\widehat{\text{ATT}}(\kappa)$ for $\kappa \geq 0$—for each wage horizon across the three samples.

The ATTs are stable across all three samples. For log wage year 1, the estimates range from 0.038 to 0.046—a spread of less than 1 percentage point around the baseline of 0.042. For year 2 the estimates are nearly identical (0.033–0.034). Year 3 shows slightly more variation (0.034–0.055), driven by noisier estimates at longer horizons where fewer cohorts contribute to the average. The event study shapes are similarly stable: pre-period coefficients remain flat and the gradual post-accreditation growth is present in all three samples. We conclude that the main results are not an artifact of noisy wage averages from small graduation cohorts.

Table 5: Robustness to Minimum Cell Size: Mean Post-Period ATT

	Log annual wage		
Sample	Year 1	Year 2	Year 3
Baseline (alum ≥ 4)	0.042	0.034	0.036
+ alum ≥ 5	0.046	0.034	0.034
+ alum ≥ 10	0.038	0.033	0.055
Observations (baseline)	18,672	18,707	18,708

Notes. Each cell reports the simple average of the BJS event-study estimates over all post-accreditation event times $\kappa \in \{0, 1, \ldots, 8\}$. The baseline sample includes all program–graduation year cells with at least 4 graduates, the minimum present in the SIES wage data. The alum ≥ 5 sample additionally drops the 1,028 cells with exactly 4 graduates (5.5% of baseline). The alum $\geq$ 10 sample retains only cells with at least 10 graduates (73.1% of baseline). Standard errors clustered at the program level; event study plots available in Figures **??**–**??**.