

# The Lottery of Legal Licensing: Examiner Leniency in Italy's Bar Exam

APEP\*

@SocialCatalystLab

March 29, 2026

## Abstract

Italy's bar exam uses a public lottery to assign grading commissions across Courts of Appeal, creating quasi-random variation in examiner identity. Using sorteggio pairings and pass rates for 26 courts over four exam sessions (2017–2023), we document that 45% of the cross-court variation in pass rates (24%–71%) is within-court—changing when the randomly assigned grading court changes. Regressions using a residualized leave-one-out leniency measure yield positive but imprecise estimates, consistent with examiner leniency driving part of this variation but lacking statistical power to isolate the effect from four sessions of aggregate data. The sorteggio reveals the scale of outcome instability in a high-stakes licensing exam.

*JEL Codes:* J44, K40, L84

*Keywords:* occupational licensing, examiner leniency, bar exam, lottery, Italy

---

\*Autonomous Policy Evaluation Project. Contributor: DavidYD. Correspondence: scl@econ.uzh.ch

# 1 Introduction

In Italy, the pass rate on the national bar exam ranges from 24% to 71% across Courts of Appeal. A candidate sitting in Caltanissetta faces odds less than half as favorable as one in Milano. Is this because southern Italian law graduates are less prepared, or because some grading commissions are harsher than others?

The answer matters for 20,000 candidates who sit the exam each year, and more broadly for the design of occupational licensing systems that gate entry into professions affecting one-quarter of the European workforce (Kleiner and Krueger, 2013). If geographic disparities reflect genuine differences in candidate quality, the exam is functioning as intended—screening for competence. If they reflect examiner leniency, the licensing gate is partly arbitrary, and careers are determined by a lottery that no candidate chose to enter.

We exploit a unique institutional feature of the Italian bar exam to distinguish these explanations. Since at least 2015, the Ministry of Justice has conducted an annual *sorteggio*—a public lottery—to assign grading commissions. Under this system, anonymous written papers from candidates at one Court of Appeal are shipped to a commission at a *different* Court of Appeal for evaluation. The pairing is determined by lottery within five size tiers (*fasce*). Critically, the assigned grading court changes each year, creating year-to-year variation in examiner identity for the same candidate population.

This design allows a direct test. If pass rates at a given court fluctuate systematically with the identity of the lottery-assigned grading court, examiner standards—not just candidate quality—contribute to the geographic disparity. We construct a panel of 93 court-year observations across four exam sessions (2017, 2019, 2022, 2023), merging *sorteggio* pairings from official Ministry verbales with pass rates from legal news aggregators. We compute both raw and residualized leave-one-out measures of grading court leniency and estimate their association with pass rates controlling for candidate court and year fixed effects.

Two findings emerge. First, the variance decomposition reveals that 45% of the total cross-court variation in pass rates is *within-court*—it changes when the grading court changes. This is a striking descriptive fact: nearly half of Italy’s well-known geographic disparity in bar exam outcomes is unstable, shifting year to year as the lottery rotates grading assignments. Second, we construct a residualized leave-one-out measure of grading court leniency—purged of candidate-court and year effects to avoid contamination from non-random candidate pools—and find a positive but imprecisely estimated coefficient ( $\hat{\beta} = 0.36$ ,  $t = 1.16$  in the written-exam subsample). The direction is consistent with examiner leniency driving within-court variation, but statistical power is limited by four sessions of aggregate court-level data.

These results contribute to the examiner leniency literature, which has documented quasi-random variation in judges (Kling, 2006; Dobbie et al., 2018; Frandsen et al., 2023), patent examiners (Farre-Mensa et al., 2020), and disability evaluators (Maestas et al., 2013), but has never been applied to occupational licensing exams. Bamieh et al. (2024) use an RDD at the pass/fail cutoff of the same Italian exam; we exploit a different and independent source of variation—the *sorteggio*

itself. Pagliero (2019) shows that licensing exam difficulty affects entry salaries in the US; we provide direct evidence on examiner inconsistency in the European context. They also speak to the large literature on occupational licensing (Kleiner and Krueger, 2013; Blair and Chung, 2019), which documents wage premia and employment effects but cannot identify supply-side effects because licensing stringency is typically endogenous. The Italian *sorteggio* provides a rare setting where the stringency of the licensing gate varies quasi-randomly across candidate populations within the same size tier.

The paper is organized as follows. Section 2 describes the institutional setting. Section 3 presents the data. Section 4 outlines the empirical strategy. Section 5 presents results. Section 6 concludes.

## 2 Institutional Background

### 2.1 The Italian Bar Exam

Italy's *Esame di Stato per l'Abilitazione all'Esercizio della Professione di Avvocato* is the mandatory licensing examination for entry into the legal profession. Approximately 20,000–25,000 candidates sit the exam annually at 26 Courts of Appeal (*Corti d'Appello*) distributed across the country. The exam is governed by *decreto legge* 13 March 2021, n. 31 (converted to *legge* 15 April 2021, n. 50), which reformed the format from three written essays to a combined written-oral structure.

Candidates register at the Court of Appeal in whose jurisdiction they completed their legal traineeship (*tirocinio forense*). Written exams are held simultaneously across all 26 courts on the same dates, ensuring common exam content nationwide.

### 2.2 The Sorteggio: A Public Lottery for Grading Assignment

After the written exams, the Ministry of Justice convenes a central commission to conduct the *sorteggio*—a public lottery that determines which court's commission will grade which other court's papers. Courts are grouped into five size tiers (*fascie*) based on candidate volume:

- **Fascia A** (largest): Milano, Roma, Napoli
- **Fascia B**: Bologna, Catania, Catanzaro, Palermo, Venezia
- **Fascia C**: Bari, Firenze, Salerno, Torino
- **Fascia D**: Ancona, Brescia, Cagliari, Genova, L'Aquila, Lecce, Messina, Reggio Calabria
- **Fascia E** (smallest): Caltanissetta, Campobasso, Perugia, Potenza, Trento, Trieste

Lots are drawn within each tier. The result is a one-to-one pairing: papers from Court  $X$  are shipped to Court  $Y$  for correction by Commission  $Y$ . The pairings change each year. Bolzano is excluded due to bilingual examination requirements and always evaluates its own candidates.

The sorteggio records (*verbali*) are published as official Ministry of Justice decrees and are publicly accessible. This institutional feature creates a setting where each court’s candidates face a different grading commission each year, with the assignment determined by lottery.

### 2.3 COVID-Era Format Change

During 2020–2022, the exam switched to an oral-only format (*doppio orale*) due to pandemic restrictions. Under this format, examiners traveled to candidate courts for face-to-face evaluation rather than evaluating shipped written papers. The sorteggio still operated to assign commissions, but the evaluation modality differed. We include these observations in the full sample with appropriate controls and verify robustness in the written-exam subsample.

## 3 Data

We construct a court-year panel combining two data sources.

**Sorteggio pairings.** We extract the lottery-assigned grading court for each candidate court from official Ministry of Justice *verbali* and press releases, covering exam sessions 2017, 2019, 2021, 2022, 2023, 2024, and 2025.

**Pass rates.** We collect candidate court-level pass rates from legal news aggregators (primarily *Formazionegiuridica.org*, supplemented by *Controcampus.it* and *Money.it*), covering sessions 2016–2024. For a subset of court-years, absolute candidate and pass counts are available; for others, only percentages are reported.

Merging these sources yields a panel of 93 court-year observations with both sorteggio pairings and pass rates, spanning four exam sessions (2017, 2019, 2022, 2023) and 26 candidate courts. We drop the 2024 session, which has pass rates for only 3 of 26 courts (results were still being released at time of writing). The panel is unbalanced: 2017 has 25 courts, 2019 has 24, 2022 has 24, and 2023 has 20, reflecting gaps in news aggregator coverage.

Table 1 reports summary statistics. The mean pass rate is 46%, with a standard deviation of 10 percentage points. The range is striking: from 24.4% (Caltanissetta, 2019) to 70.9% (Milano, 2017). The 2022 oral-format session accounts for 24 observations; the remaining 69 are written-format.

Table 1: Summary Statistics

	N	Mean	SD	Min	Median	Max
Pass rate (all sessions)	93	0.460	0.104	0.244	0.450	0.709
Pass rate (written only)	69	0.441	0.109	0.244	0.437	0.709
Raw grading leniency (LOO)	93	0.461	0.071	—	—	—
Residualized leniency (LOO)	93	0.000	0.043	—	—	—

*Notes:* Grading leniency is the leave-one-out mean pass rate when the assigned grading court evaluates other courts’ papers in other years.

## 4 Empirical Strategy

### 4.1 Variance Decomposition

We first decompose the total variance in pass rates into between-court and within-court components. Between-court variance reflects persistent differences in candidate quality across Courts of Appeal. Within-court variance reflects changes over time—coinciding with the sorteggio-determined rotation of grading courts.

If the exam were perfectly meritocratic and all grading commissions applied identical standards, within-court variance would be zero (conditional on year effects). Any within-court variance that correlates with grading court identity provides evidence of examiner leniency.

### 4.2 Grading Court Leniency

We construct a leave-one-out measure of grading court leniency. For each court-year observation  $(c, t)$  where court  $c$  is graded by court  $g$ , leniency is:

$$\text{Leniency}_{g,-ct} = \frac{1}{|\mathcal{S}_g \setminus \{(c, t)\}|} \sum_{(c', t') \in \mathcal{S}_g \setminus \{(c, t)\}} \text{PassRate}_{c', t'} \quad (1)$$

where  $\mathcal{S}_g$  is the set of all court-year observations graded by court  $g$ .

### 4.3 Main Specification

We estimate:

$$\text{PassRate}_{ct} = \alpha_c + \gamma_t + \beta \cdot \text{Leniency}_{g(c,t),-ct} + \varepsilon_{ct} \quad (2)$$

where  $\alpha_c$  are candidate court fixed effects,  $\gamma_t$  are year fixed effects, and  $g(c, t)$  is the sorteggio-assigned grading court. The coefficient  $\beta$  captures the within-court association between grading court leniency and pass rates, identified by the year-to-year rotation of grading assignments under the sorteggio.

Standard errors are heteroskedasticity-robust throughout.

## 5 Results

### 5.1 Variance Decomposition

Table 2 reports the variance decomposition. Of the total cross-court variance in pass rates, 60.5% is between courts (persistent candidate quality differences) and 44.5% is within courts (year-to-year changes). In the written-exam subsample, the within-court share rises to 50.9%. (Shares exceed 100% because the panel is unbalanced; with unequal court-year counts, between and within components need not sum exactly to total variance.) This is a large and policy-relevant finding: between two-fifths and one-half of Italy’s geographic disparity in bar exam outcomes is unstable, shifting when the lottery-assigned grading commission changes.

Table 2: Variance Decomposition of Bar Exam Pass Rates

Component	Variance	Share (%)	Written Only (%)
Total	0.01091	100.0	100.0
Between courts (persistent)	0.00660	60.5	56.5
Within courts (grading court rotation)	0.00486	44.5	50.9

*Notes:* Components do not sum exactly to total due to unbalanced panel. Between-court variance from court means; within-court from deviations. Written Only excludes the 2022 oral-format session.  $N = 93$  (all sessions);  $N = 69$  (written only).

## 5.2 Grading Court Leniency and Pass Rates

A scatter plot of residualized pass rates against residualized grading court leniency (available in the replication materials) shows a positive slope consistent with examiner standards influencing outcomes, though the relationship is noisy.

Table 3 reports regression results using two leniency measures. The *raw* leave-one-out measure (columns 1–2) computes each grading court’s average pass rate when it evaluates other courts. This measure is contaminated: a high value can reflect lenient grading *or* assignment to courts with strong candidates. The raw measure yields positive and significant coefficients ( $\hat{\beta} = 0.26$ ,  $t = 2.03$  with court and year FE;  $\hat{\beta} = 0.35$ ,  $t = 2.23$  for written exams only).

The *residualized* measure (columns 3 and 5) addresses contamination by first purging pass rates of candidate-court and year fixed effects, then computing leave-one-out leniency from the residuals. This isolates the grading court’s contribution from the candidate pool’s quality. The residualized measure yields positive but imprecisely estimated coefficients ( $\hat{\beta} = 0.18$ ,  $t = 0.90$  full sample;  $\hat{\beta} = 0.36$ ,  $t = 1.16$  written only). The direction is consistent with examiner leniency, but the standard errors are large—reflecting both the smaller variation in the purged measure (SD = 0.043 vs. 0.071 for raw) and the limited sample of four court-year sessions.

We interpret these results as follows: the variance decomposition establishes that 45% of pass rate variation is within-court. The leniency regressions provide directional evidence that this within-court variation correlates with grading court identity, but the data cannot precisely quantify the examiner-leniency share.

## 5.3 Within-Court Variation

The within-court ranges are large for most courts—for example, Catanzaro varies from 33% to 52%, while Trieste ranges from 26% to 54%. These within-court swings coincide with changes in the lottery-assigned grading court.

## 6 Discussion and Conclusion

The Italian bar exam sorteggio provides a rare institutional setting for studying regulatory consistency. We document that 45% of the geographic disparity in pass rates—one of the most discussed features

Table 3: Grading Court Leniency and Bar Exam Pass Rates

Dependent Variable: Model:	pass_rate				
	Pooled (1)	Raw+FE (2)	Purged+FE (3)	Raw+FE(W) (4)	Purged+FE(W) (5)
<i>Variables</i>					
Constant	0.3212*** (0.0733)				
raw_leniency	0.2989* (0.1566)	0.2562** (0.1264)		0.3505** (0.1573)	
purged_leniency			0.1837 (0.2044)		0.3614 (0.3117)
<i>Fixed-effects</i>					
candidate_court		Yes	Yes	Yes	Yes
exam_year		Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	93	93	93	69	69
R <sup>2</sup>	0.04164	0.64528	0.62810	0.64090	0.61485
Within R <sup>2</sup>		0.05761	0.01199	0.10566	0.04080

*Heteroskedasticity-robust standard-errors in parentheses*

*Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1*

Dependent variable: candidate court pass rate. Columns 1-2 use raw leave-one-out leniency. Columns 3 and 5 use residualized leave-one-out leniency (purged of candidate-court and year effects). (W) = written exams only (excludes 2022 oral session). Heteroskedasticity-robust SEs in parentheses.

of Italy’s legal profession—is unstable, changing when the lottery rotates grading assignments. The leniency regressions are directionally consistent with examiner standards driving this instability, though the estimates lack precision.

These findings have two implications. First, the scale of within-court variation is itself informative for licensing policy. Italy’s Ministry of Justice introduced the sorteggio to reduce geographic disparities. Our variance decomposition suggests the disparities are not primarily rooted in stable differences in candidate preparation—they shift substantially from year to year. Whether this instability reflects examiner leniency, year-to-year cohort fluctuations, or other factors cannot be definitively resolved with aggregate data.

Second, the sorteggio mechanism itself is a contribution. Professional licensing exams in law, medicine, and accounting affect millions of workers’ careers (Kleiner and Krueger, 2013), yet evaluator consistency is rarely testable because evaluator assignment is typically endogenous. The Italian sorteggio generates the quasi-random variation needed for clean identification. With individual-level data—available through ISTAT’s microdata lab—future work can estimate examiner leniency precisely and trace downstream effects on lawyer supply and earnings.

Three limitations deserve note. First, the sample is modest (93 court-year observations across four sessions), limiting statistical power for the leniency regressions. Second, the panel is unbalanced: pass rates come from news aggregators, not a uniform administrative release, and some courts are missing in some years. Third, the 2022 session used an oral rather than written format, which changes the grading mechanism. We verify robustness in the written-exam subsample.

The sorteggio was designed to ensure fairness. What it reveals is the scale of outcome instability in a high-stakes licensing exam: nearly half of the variation that candidates, policymakers, and the legal profession attribute to geography is not persistent.

## References

- Bamieh, Omar, Andrea Cintolesi, and Mario Pagliero**, “The Value of Professional Licensing: Evidence from the Italian Bar Exam,” *Bank of Italy Working Paper*, 2024, (1440).
- Blair, Peter Q. and Bobby W. Chung**, “How Much of Barrier to Entry Is Occupational Licensing?,” *British Journal of Industrial Relations*, 2019, 57 (4), 919–943.
- Dobbie, Will, Jacob Goldin, and Crystal S. Yang**, “The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges,” *American Economic Review*, 2018, 108 (2), 201–240.
- Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist**, “What Is a Patent Worth? Evidence from the U.S. Patent “Lottery”,” *The Journal of Finance*, 2020, 75 (2), 639–682.
- Frandsen, Brigham R., Lars J. Lefgren, and Emily C. Leslie**, “Judging Judge Fixed Effects,” *American Economic Review*, 2023, 113 (1), 253–277.

**Kleiner, Morris M. and Alan B. Krueger**, “Analyzing the Extent and Influence of Occupational Licensing on the Labor Market,” *Journal of Labor Economics*, 2013, 31 (S1), S173–S202.

**Kling, Jeffrey R.**, “Incarceration Length, Employment, and Earnings,” *American Economic Review*, 2006, 96 (3), 863–876.

**Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand**, “Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt,” *American Economic Review*, 2013, 103 (5), 1797–1829.

**Pagliari, Mario**, “Licensing Exam Difficulty and Entry Salaries in the US Market for Lawyers,” *British Journal of Industrial Relations*, 2019, 57 (4), 767–798.

## A Standardized Effect Direction

Table 4: Standardized Effect Direction (SDE)

Panel A: Pooled					
Outcome	Estimate	SE	<i>p</i> -value	Direction	Classification
Pass rate (purged leniency)	0.184	0.204	0.372	Positive	Imprecise positive
Pass rate (raw leniency)	0.256	0.126	0.047	Positive	Significant positive

  

Panel B: Heterogeneous					
Subsample	Estimate	SE	<i>p</i> -value	Direction	Classification
Written exams only (purged)	0.361	0.312	0.253	Positive	Imprecise positive
Written exams only (raw)	0.351	0.157	0.032	Positive	Significant positive

**Country:** Italy. **Research question:** Does lottery-assigned grading court identity explain geographic disparities in bar exam pass rates? **Policy mechanism:** Annual sorteggio (public lottery) assigns grading commissions across Courts of Appeal within size tiers. **Outcome definition:** Candidate court pass rate on the esame di stato per l’abilitazione all’esercizio della professione di avvocato. **Treatment:** Identity of the lottery-assigned grading court (varies year to year for each candidate court). **Data:** Ministry of Justice sorteggio verbales (grading court assignments) merged with pass rates from legal news aggregators (FormazioneGiuridica, Money.it). Court-year panel, 2017–2023. **Method:** Variance decomposition and OLS with candidate-court and year fixed effects. Leave-one-out grading court leniency (raw and FE-residualized). Heteroskedasticity-robust standard errors. **Sample:** 93 court-year observations across 26 Courts of Appeal and 4 exam sessions. Unbalanced panel (25, 24, 24, 20 courts per session). Classification refers to magnitude, not statistical significance.

## Acknowledgements

This paper was autonomously generated as part of the Autonomous Policy Evaluation Project (APEP).

**Contributors:** @SocialCatalystLab

**First Contributor:** <https://github.com/SocialCatalystLab>

**Project Repository:** <https://github.com/SocialCatalystLab/ape-papers>