

The Decline of the Follow-On in 125 Years of First-Class Cricket*

Johan Fourie[†]

Abstract

Whether experienced professionals choose the strategy that maximises their chance of success is hard to test, because decisions and their outcomes are rarely recorded together. We study the follow-on in cricket, a choice governed by a written rule: a side may make its opponent bat again immediately, but only if its first-innings lead reaches a statutory margin. The decision and its outcome have been recorded for more than 125 years. Using 22,113 first-class matches from 1900 to 2025, we exploit the margin in a fuzzy regression discontinuity design. Conventional estimates imply that enforcing raises the probability of winning and lowers the probability of losing. Captains nevertheless enforce far less often than earlier generations did. Because opponents bat to stay below the margin, bounds that allow for this sorting do not identify the sign of the effect. A matched event study finds that the response to the rare defeat after enforcing fades within a year, so personal experience does not explain the decline. The findings suggest that professionals overweight rare and conspicuous failures, and the setting illustrates a general point about natural experiments: rules that create discontinuities also create incentives to avoid them.

Keywords: follow-on; regression discontinuity; experience effects; sports economics; managerial decision-making

JEL codes: D81; D91; L83; Z20

*I thank Ric Finlay and Jim Palfreyman, compilers of the DAFT database. The research received no external funding and I report no conflict of interest. This paper was created with the help of Anthropic’s Claude Code and OpenAI’s Codex, which assisted with statistical code development, reproducibility and methods audits, and language editing; I reviewed and verified all analyses, results and text and take full responsibility for the content. Cite this paper as: Fourie, Johan. 2026. “The Decline of the Follow-On in 125 Years of First-Class Cricket.” Working Paper, Department of Economics, Stellenbosch University.

[†]Department of Economics, Stellenbosch University. Email: johanf@sun.ac.za.

1 Introduction

Economists have long asked whether experienced professionals choose the strategy that maximises their chance of success. The question is hard to test inside firms, where the decision, the information behind it and the outcome that follows are rarely recorded together. Professional sport records all three. Strategic choices in sport have therefore become a standard test of optimising behaviour, and the recurring finding is excess caution: decision-makers avoid an aggressive option with a higher expected payoff when its failure would be rare and conspicuous (Romer, 2006; Walker et al., 2018).

The follow-on in cricket, one of the world’s most popular sports, is such a choice. After both sides have batted once, the side that batted first may require its opponent to bat again immediately, but only if its lead reaches a margin fixed in the Laws of Cricket. The rule makes the sport an unusual laboratory. Whether the option exists depends on a written threshold rather than on judgement, and the decision and its outcome have been recorded for more than 125 years. Captains treat enforcing as dangerous because the bowlers are tired and the opponent may recover, yet only three Test matches since 1900 have ended in defeat for the enforcing side. Enforcement among eligible captains has nevertheless fallen from 77 per cent in 1950–1979 to 55 per cent since 2010.

We make two contributions. First, we use the statutory margin in a fuzzy regression discontinuity design to estimate the effect of enforcing. The previous study fits an outcome equation to 391 Tests and finds no measurable effect (Scarf and Akhtar, 2011). Our design separates the decision from the lead that preceded it. It also shows where strategic sorting limits that comparison. Second, we ask whether the decline reflects experience effects. Personal experience of rare adverse events changes economic behaviour for years afterwards (Malmendier and Nagel, 2011; Malmendier, 2021). A captain whose enforcement ends in defeat experiences the feared outcome directly, so the follow-on provides a field test of that mechanism among professionals. We also correct a classification error: only the side batting first may enforce. Identifying the potential enforcer from its score overstates eligible matches by two thirds.

2 Data

The DAFT database covers men’s Test cricket and the domestic first-class competitions of England, Australia, New Zealand, South Africa, the West Indies, Pakistan and Ireland. We use 22,113 competitive matches from 1900 to 2025. Of these, 4,227 were eligible and 2,743 enforced. We exclude tour and university fixtures and the 1961–1962 County seasons, when the follow-on was suspended.

3 Empirical strategy

The statutory margin was 150 runs for matches of at least three days from 1900 to 1979. From 1980 it is 200 for matches of at least five days and 150 for three or four days. Let x_m be the first-innings lead minus the applicable margin. Eligibility is $\mathbf{1}[x_m \geq 0]$. Crossing zero sharply raises the

probability of enforcing (Table 1). We estimate local linear fuzzy discontinuities using a triangular kernel and robust bias-corrected inference (Calonico, Cattaneo and Titiunik, 2014).

The side batting second can bat to stay below the margin. We test for sorting with the local polynomial density test of Cattaneo, Jansson and Ma (2020), report donut estimates and apply the partial-identification method of Gerard, Rokkanen and Rothe (2020). These bounds allow manipulation and infer the manipulated share from the density. The assumptions are weaker than standard continuity. They identify a set rather than a point.

For the behavioural analysis, treated events are the 34 matches in which a team enforced and lost. Controls are other enforcements matched without replacement within competition, two seasons and 60 runs of the lead. The design follows Fadlon and Nielsen (2021), who compare units that experience a shock with matched units at similar risk of it. We follow each team for five years either side and censor observations at its next such defeat. The model includes matched-set by team and matched-set by event-year fixed effects, a flexible lead control and scheduled match length. We cluster on team, but with 34 treated events cluster-asymptotic inference is unreliable (MacKinnon, Nielsen and Webb, 2023). We therefore also reassign the defeat within each matched set 2,000 times, using a studentised statistic for the first post-event year and the five-year average. The five-year average pools all post-event observations within five years, including the first year. We invert the test for the average. The defeat is not randomly assigned, so this remains a matched observational design.

4 Results

Figure 1(a) shows that enforcement rises at the legal margin. The Test discontinuity moves from 150 to 200 in 1980 while the domestic discontinuity stays at 150. Figure 2 shows the corresponding outcome changes. Table 1 reports the estimates. Enforcing raises the probability that the side batting first wins by 24.0 percentage points (standard error 5.3), lowers the probability that it loses and makes a decisive result more likely. These are local effects for matches near the statutory margin and for captains whose decision changes with eligibility. The effect on reaching a result is consistent with enforcing saving time. Estimates for Test matches alone, pooled or at the separate pre- and post-1980 cutoffs, are too imprecise to interpret.

The feared outcome is rare. Enforcing sides lost 1.3 per cent of domestic matches and 0.9 per cent of Tests. Eligible sides that declined lost 9.1 and 1.1 per cent respectively. These are descriptive comparisons.

Sorting is visible. There are more matches just below the cutoff than just above it, and the density test rejects the absence of manipulation. The estimate changes little when matches within five runs of the cutoff are excluded. Formal bounds are less precise. The excess density implies that about a third of matches just below the cutoff may have avoided eligibility, and the resulting identified set includes zero (supplemental appendix). The conventional estimates suggest that enforcing helps, but the weak-assumption bounds do not identify its sign.

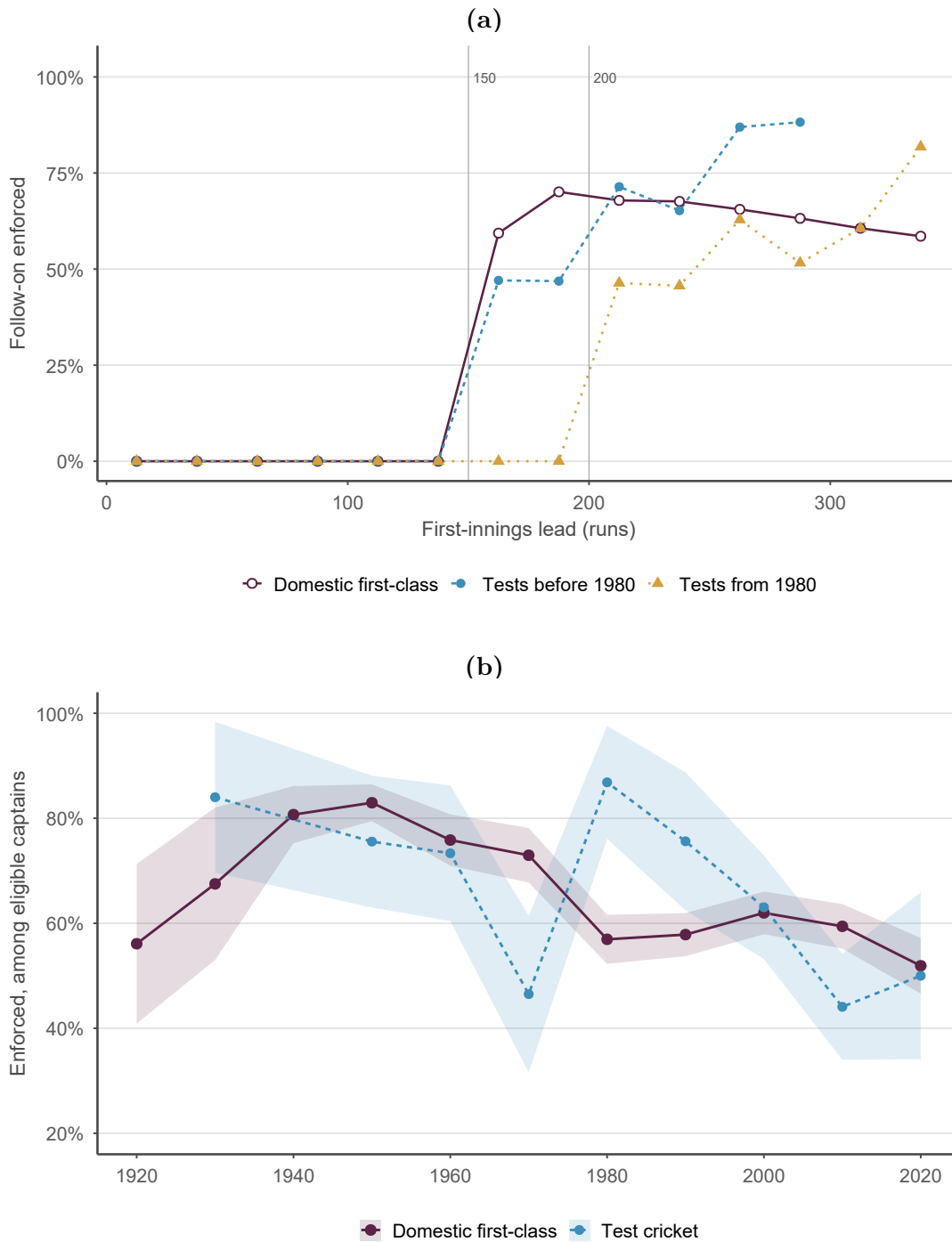


Figure 1: The rule and the decline in its use. Panel (a) gives the share of matches with the follow-on enforced by first-innings lead. Panel (b) gives the share of eligible captains who enforced by decade, with 95 per cent bands. Unit of observation is the match. Sample: 22,113 competitive first-class matches, 1900–2025.

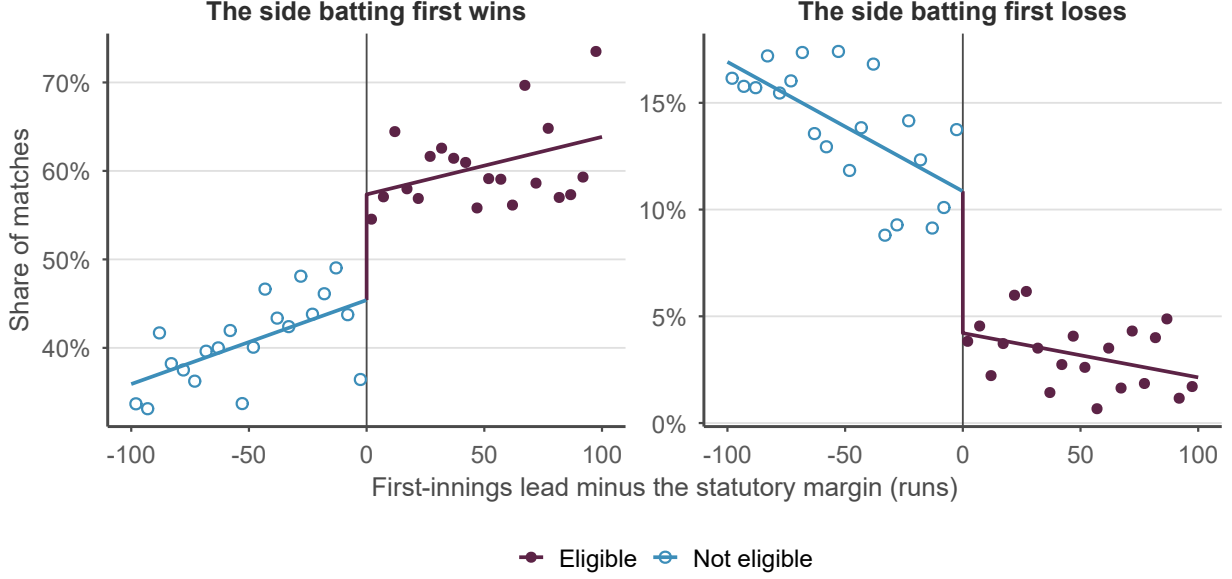


Figure 2: Outcomes by distance from the statutory margin. Points are averages in five-run bins; lines are local linear fits either side. The discontinuities give the conventional reduced-form estimates. Unit of observation is the match.

Table 1: Fuzzy regression discontinuity estimates; robust bias-corrected standard errors in parentheses. Unit of observation is the match. Effective samples for the all-match fuzzy estimates are 7,026 (win), 6,522 (loss) and 9,137 (result); the Test win sample is 1,070. *** $p < 0.01$, ** $p < 0.05$.

Outcome	All matches	Test matches
<i>Panel A. First stage</i>		
Enforced the follow-on	0.547*** (0.026)	0.438*** (0.066)
<i>Panel B. Reduced form: effect of eligibility</i>		
Side batting first wins	0.136*** (0.030)	0.068 (0.074)
Side batting first loses	-0.068*** (0.016)	
Match reaches a result	0.062** (0.027)	
<i>Panel C. Fuzzy estimate</i>		
Side batting first wins	0.240*** (0.053)	0.161 (0.177)
Side batting first loses	-0.122*** (0.029)	
Match reaches a result	0.107** (0.047)	

Figure 3 and Table 2 qualify the behavioural result. Teams that enforced and lost enforce markedly less during the first year after the defeat, then return towards their earlier level. Averaged over five years, the change is close to zero. The pre-event coefficients are jointly indistinguishable from zero. The evidence is consistent with a short-lived response. The individual coefficients are imprecise, so the first-year estimate should be read as suggestive.

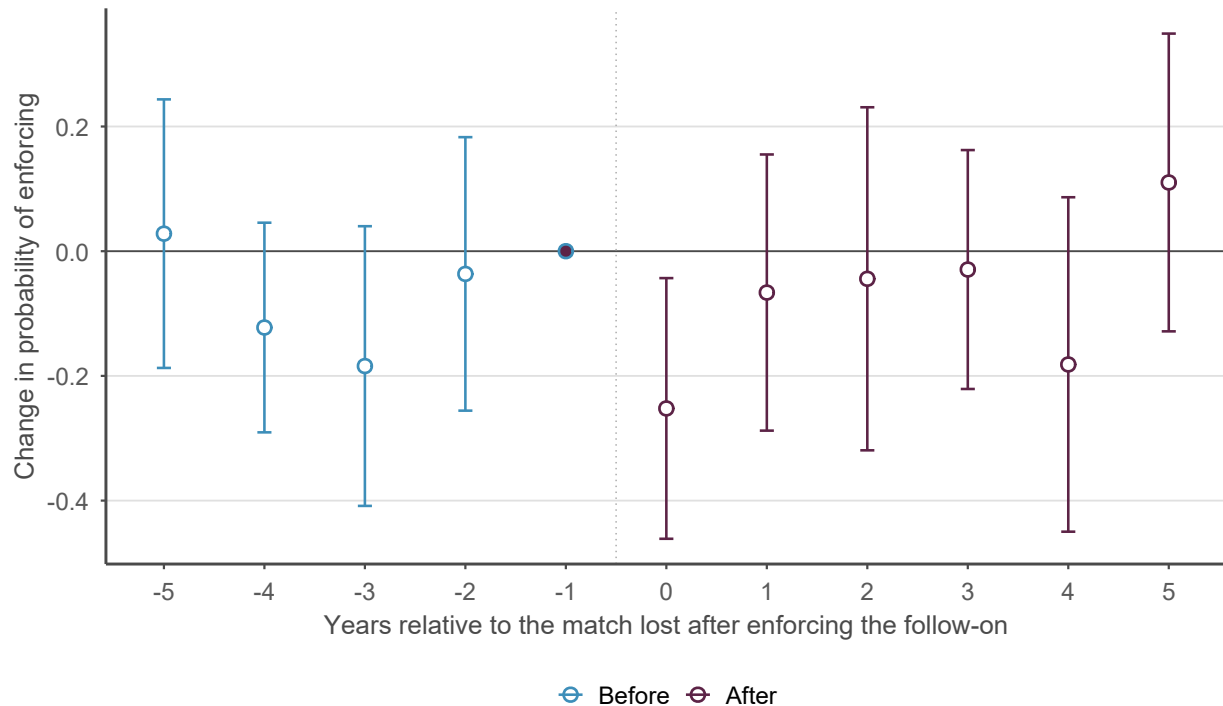


Figure 3: Enforcement before and after a team enforced and lost. Coefficients compare that team with matched teams that enforced and did not lose. Intervals are clustered on team. Year zero is the first 365 days after the event. Unit of observation is the team-year; 34 treated events.

Table 2: Change in the probability of enforcing after a team enforced and lost. Unit of observation is the team-year; 34 treated events. Standard errors clustered on team; randomisation p-values from 2,000 within-matched-set reassignments of the defeat.

Window	Estimate	Std. error	Rand. p	95% interval
First post-event year	-0.252	0.107	0.053	$[-0.461, -0.043]$
Five-year average	-0.035	0.046	0.498	$[-0.140, 0.070]$

5 Conclusion

The conventional regression discontinuity estimates favour enforcing. Wins rise, losses fall and more matches reach a result. This conclusion is conditional on continuity around the margin.

The opposing side can save the follow-on, and bounds that admit this sorting include zero. The data therefore provide qualified evidence for enforcing. They do not provide a manipulation-robust causal answer.

The estimates nevertheless have a practical reading. Modern captains decline the follow-on more often than any generation before them. Nothing in our results supports the added caution. Taken at face value, the estimates imply that a captain who declines near the margin gives up a substantial probability of winning. Even the descriptive rates give the cautious option no advantage: eligible sides that declined lost more often than sides that enforced. The defeat that captains fear has occurred three times in 125 years. Professional captains appear to overweight a rare and conspicuous failure, as American football coaches do on fourth down (Romer, 2006).

A defeat after enforcing appears to reduce enforcement during the next year. The reduction is not detectable across the five-year window, so we find no evidence that personal experience accounts for a decline lasting several decades. This contrasts with household finance, where personal experience of rare events changes behaviour for decades (Malmendier and Nagel, 2011). Captains decide often and receive immediate feedback, which may shorten the effect of any single match. The interval in Table 2 nevertheless admits a moderate lasting fall. The event study has only 34 treated events and conditions on an observed defeat. Its short-run estimate should be read as suggestive. Supplemental appendix checks separate who responds. Post-event estimates are similar whether or not the event captain remains, and rival teams that observed a disaster show no detectable change. The durable change may reflect a shared professional norm. A norm formed by observing the disasters of others would rest on indirect rather than personal experience (Alfonsi, Dossi and Monk, 2026), which this design cannot identify.

The setting also carries a methodological lesson. A statutory threshold, a recorded decision and thousands of repetitions are the conditions economists seek in a natural experiment. The same strategic play that makes the decision worth studying weakens the design, because the side batting second bats against the threshold and sorts the running variable. Rules that create discontinuities also create incentives to avoid them.

Data availability

The derived match-level data and replication code are available at <https://github.com/johanfourieza/research/tree/main/2026/follow-on>. The licensed DAFT source files cannot be redistributed.

References

Alfonsi, Livia, Gaia Dossi, and Thomas Monk. 2026. “Indirect experience.” Harvard Business School Working paper.

- Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik.** 2014. “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica*, 82(6): 2295–2326.
- Cattaneo, Matias D., Michael Jansson, and Xinwei Ma.** 2020. “Simple local polynomial density estimators.” *Journal of the American Statistical Association*, 115(531): 1449–1455.
- Fadlon, Itzik, and Torben Heien Nielsen.** 2021. “Family labor supply responses to severe health shocks: evidence from Danish administrative records.” *American Economic Journal: Applied Economics*, 13(3): 1–30.
- Gerard, François, Miikka Rokkanen, and Christoph Rothe.** 2020. “Bounds on treatment effects in regression discontinuity designs with a manipulated running variable.” *Quantitative Economics*, 11(3): 839–870.
- MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb.** 2023. “Cluster-robust inference: a guide to empirical practice.” *Journal of Econometrics*, 232(2): 272–299.
- Malmendier, Ulrike.** 2021. “Experience effects in finance: foundations, applications, and future directions.” *Review of Finance*, 25(6): 1339–1363.
- Malmendier, Ulrike, and Stefan Nagel.** 2011. “Depression babies: do macroeconomic experiences affect risk taking?” *Quarterly Journal of Economics*, 126(1): 373–416.
- Romer, David.** 2006. “Do firms maximize? Evidence from professional football.” *Journal of Political Economy*, 114(2): 340–365.
- Scarf, Philip, and Sohail Akhtar.** 2011. “An analysis of strategy in the first three innings in test cricket: declaration and the follow-on.” *Journal of the Operational Research Society*, 62(11): 1931–1940.
- Walker, Jesse, Jane L. Risen, Thomas Gilovich, and Richard Thaler.** 2018. “Sudden-death aversion: avoiding superior options because they feel riskier.” *Journal of Personality and Social Psychology*, 115(3): 363–378.