Hunting in Pairs:
Testing for Coworker Effects in Professional Cricket*

Johan Fourie†

Abstract

Sports data make coworkers visible. They do not make coworker assignment random. I study 1,355,340 focal-bowler overs in Tests, domestic multi-day cricket, one-day internationals, T20 internationals and franchise T20. Conventional bowler-by-match estimates imply large, format-dependent associations between a bowler's wicket rate and the pre-match quality of the opposite-end partner. These estimates are highly sensitive to innings phase, workload position and empirical-Bayes support. In pooled long-form cricket the preferred association is −0.43% of the mean wicket rate per standard deviation of partner quality (95% confidence interval: −1.30% to +0.45%). Three prospective pair measures do not predict held-out output. A pair-disrupting injury design has a minimum detectable effect nearly 30 times the largest timing-controlled coefficient. The evidence does not identify a structural zero. It shows that granular sports data and worker fixed effects can produce convincing but misleading evidence of partnership effects unless coworker timing, shrinkage calibration and common experience support are made explicit.

Keywords: peer effects; team production; coworker assignment; empirical Bayes; cricket

JEL codes: J24; J44; L83; M54

*I thank the Cricsheet project for the ball-by-ball match records on which this paper depends, and ESPNcricinfo, *Wisden Cricketers' Almanack* and the Association of Cricket Statisticians and Historians for the archival material quoted in Section 2. This paper was created with the help of Anthropic's Claude Code (Fable 5) and OpenAI's Codex (GPT-5.6). Cite this paper as: Fourie, Johan. 2026. "Hunting in Pairs: Testing for Coworker Effects in Professional Cricket." Working Paper, Department of Economics, Stellenbosch University.

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

1 Introduction

The idea of a productive partnership is central to how teams are selected, coached and discussed. Two workers may be more valuable together than apart. A senior teammate may improve a junior's decisions. Complementary skills may change an opponent's behaviour. Familiarity may reduce coordination costs. If these mechanisms matter, managers should preserve successful pairs. A worker's value should then include the output he creates in others. The difficulty is that managers choose who works together, and when. A further difficulty is less discussed: the quality of a coworker with a short track record is itself an estimate that is shrunk toward a prior chosen by the researcher. A coefficient on 'coworker quality' can therefore mix production, managerial timing and the measurement rule. Whether high-frequency data and strict fixed effects are enough to separate these components is not known.

This paper asks that question in professional cricket, where the coworker is unusually well defined. Cricket also states the belief plainly: fast bowlers, the saying goes, hunt in pairs. Bowlers operate from alternating ends of the ground: for the bowler of over o (an over is a set of six legal deliveries), the bowler of over $o-1$ is the opposite-end partner. A captain can retain or replace that partner between overs. This creates many within-match changes in a worker's closest teammate. Ball-by-ball records register output, workload and player identity. The setting appears almost designed for a coworker regression: compare the same bowler in the same match when paired with stronger and weaker partners. It is not a natural experiment. A team's best bowlers tend to open the bowling together. Later spells (a spell is one bowler's unbroken run of overs) come with changes in the condition of the ball, the batters' objectives and fatigue. And the captain chooses the timing of every change.

I assemble 1,355,340 eligible focal overs from five settings: Test matches, domestic multi-day cricket, one-day internationals (ODIs), twenty-over internationals (T20Is) and franchise T20. Partner quality is an empirical-Bayes wicket rate built only from matches dated before the focal match. All main regressions include match-innings and focal-bowler-by-match fixed effects. The comparison is therefore among overs bowled by the same player in the same match. Legal-ball weights and two-way clustering by focal bowler and match apply throughout. Section 4.3 states two criteria that govern the interpretation of these estimates.

The first set of estimates follows the conventional approach. With fixed effects alone, a one-standard-deviation increase in partner quality is associated with a wicket rate about two to three per cent of the setting mean lower in the two long-form settings (Tests and domestic multi-day cricket). In the three short formats it is associated with a wicket rate three to six per cent higher. The intuitive interpretation is that partnerships are unproductive or ceremonial in long-form cricket and complementary in faster, more constrained formats.

The second set of estimates does not support that interpretation. The controls for innings position and the focal bowler's own over number are determined before the focal over. Adding them eliminates the long-form and ODI associations and cuts the two remaining short-format associations to well under half their size. In pooled long-form cricket the estimate falls to -0.43%

of the mean, with a 95% confidence interval of −1.30% to +0.45%. The negative pattern appeared in the format where partnerships have the greatest opportunity to develop. It reflects the timing of innings and spells. It is not stable evidence that strong partners reduce one another's output.

The remaining pattern in the white-ball formats (ODIs and the two T20 settings) is less reliable than its full-sample precision suggests. Empirical-Bayes shrinkage matters most for inexperienced partners. Short histories are especially common in T20Is. The five setting slopes remain different in the full sample and among partners with at least 1,000 prior balls. But the equality test fails to reject once the least-experienced quartile is excluded. The format pattern also lacks a coherent within-setting experience gradient. The positive T20I estimate is concentrated among the least experienced partners. The franchise estimate appears only among the most experienced. This is insufficient for an institutional claim about short-format bowling constraints.

A third set of estimates tests for a durable pair-specific component in the long-form overs retained by the dyad specifications. Three measures are dated strictly before the focal match. Prior pair familiarity has a small association with focal wickets that is statistically indistinguishable from zero. A partner's ability to suppress runs does not predict the focal bowler's wickets. The most direct test uses prior pair-specific wicket excess, the wickets a pair took beyond what the two bowlers' individual pre-match quality predicts. This measure does not predict held-out focal output. The estimate is as close to zero as the design can measure. These conclusions do not change when standard errors are clustered on the focal bowler, the partner or the pair, in each case together with the match. A complementary design follows established bowlers across pair-disrupting injuries. It yields only 93 clean events. Its minimum detectable effect is nearly 30 times the largest timing-controlled coefficient, so injuries cannot supply causal identification in this archive either.

This paper contributes to three literatures. The first is the economics of coworker and peer effects. The designs that identify these effects change exposure to peers through something neither the worker nor the manager fully controls: shift overlap among supermarket cashiers (Mas and Moretti, 2009), quasi-random peer composition in administrative data (Cornelissen, Dustmann and Schönberg, 2017), randomised encouragement to seek advice from colleagues (Sandvik et al., 2020), office closures that separate junior workers from seniors (Emanuel, Harrington and Pallais, 2026) and unexpected coworker deaths (Jäger and Heining, 2022). A parallel literature infers learning from coworkers through wage dynamics and equilibrium structure (Jarosch, Oberfield and Rossi-Hansberg, 2021; Herkenhoff et al., 2024). This paper supplies the complementary cautionary result: observational granularity is not a substitute for an exposure shock. Managers create value by allocating workers to jobs and moments (Lazear, Shaw and Stanton, 2015; Minni, 2026). In cricket, rotations occur within bowler-by-match cells. The captain's timing of each pairing therefore survives even worker-by-workplace fixed effects. The confounder is not only who works with whom but when the pair is activated within the production cycle. The design concerns current output. It does not test the slower accumulation of skill from coworkers, the margin at the centre of Emanuel, Harrington and Pallais (2026).

The second contribution is to the econometrics of worker-quality measurement. Empirical-Bayes shrinkage is standard when estimated quality is the object of study: teacher value added (Chetty, Friedman and Rockoff, 2014), firm and worker effects (Kline, Saggio and Sølvsten, 2020; Bonhomme et al., 2023) and employer-level discrimination (Kline, Rose and Walters, 2022). Walters (2024) surveys these methods. Chen (2026) shows that standard shrinkage misleads when estimation precision predicts the parameter. This paper documents the same problem on the right-hand side of a regression, where the posterior is the treatment rather than the estimand. Inexperienced partners are shrunk hardest, so the prior's location in each environment's quality distribution becomes part of the treatment definition. Under fixed common priors, a debutant counts as an above-average partner in Tests. In franchise T20 the same debutant counts as a partner nearly two standard deviations below average. Recentring the prior on each setting's own rate and requiring results to hold on common experience support removes the basis for an institutional reading of the cross-format pattern. Studies that regress outcomes on shrunk coworker, teacher or manager quality should therefore report the prior's position in each estimation sample, not only the prior's strength.

The third contribution is to sports economics, which has long served as a labour-market laboratory (Kahn, 2000). Palacios-Huerta (2025) makes the general case: sports settings pair the clean measurement, known rules and observable incentives of the laboratory with the stakes, expertise and sample sizes of the field. Some of the best evidence on workplace peers comes from sports. Random groupings in professional golf, a setting with little task interdependence, show scant evidence of peer effects (Guryan, Kroft and Notowidigdo, 2009). Hickman and Metz (2018) decompose that null into learning and motivation components that offset. When a research design isolates a teammate's mere presence, for example when a swimmer barely qualifies for the same final, performance can improve (Jiang, 2020). Where task interdependence is real, granular data yield productivity spillovers, teammate facilitation and effort spillovers (Gould and Winter, 2009; Arcidiacono, Kinsler and Price, 2017; Cohen-Zada, Dayag and Gershoni, 2024). Cricket offers the sharpest teammate definition of all: the partner changes over by over and is recorded at every ball. The belief that bowlers hunt in pairs has, to my knowledge, no quantitative literature of its own. Performance-evaluation research measures bowlers one at a time, and reviews of that work call for deeper analysis of bowling tactics (Chathurangi et al., 2025). This paper identifies a limit of the laboratory. The features that make sports data attractive do not by themselves protect a coworker-quality regression. The manager schedules every pairing. The treatment is an estimated quantity. As a result, conventional partnership coefficients move with innings and workload timing rather than with any stable pair signal. Pair statistics and teammate-quality regressions should be validated on held-out matches, purged of innings and workload timing, and evaluated on common experience support before they enter selection or player valuation.

2 Cricket as a coworker-assignment setting

2.1 Alternating production

A cricket innings is divided into overs, normally of six legal deliveries. One bowler delivers an over from one end of the ground. In standard cricket, the next over is delivered from the other end by a different bowler. If bowler i operates in over o, bowler j in over $o-1$ is the opposite-end partner: together they constitute the active bowling pair surrounding i's work.

The Hundred is the exception. Its five-ball units come in ten-ball end blocks. Both units within a block are therefore delivered from the same end. For its 167 matches, I define the partner from the bowler who completed the previous ten-ball end rather than mechanically using over $o-1$.

The belief that bowlers operate in pairs is old, general and still current.[1] The 1963 edition of *Wisden Cricketers' Almanack* reviewed the Test season under the heading 'Bowlers Hunt in Pairs' and named Turner and Ferris, Australia's new-ball pair of the 1880s, as the bowlers who 'set the trend long ago'. By 1989 the claim was common enough for a reviewer in *The Cricket Statistician* to dismiss the statement that 'all great quick bowlers' hunt in pairs as 'an oft-repeated assertion which does not stand up to close examination'. Players state the belief as a method. Stuart Broad described his pairing with James Anderson in exactly these terms: when one bowler is taking wickets, 'the other one can calm the other end down and make it impossible for the batsman to go anywhere'. The same language appears in live coverage. In a corpus of about eight million balls of ESPNcricinfo text commentary, bowlers 'hunt in pairs', 'bowl in partnership' and apply 'pressure from both ends'. The belief this paper tests is therefore not a construction of the research design. It is held by players, journalists and commentators.

This definition captures what cricketers mean by a bowling partnership. A restrictive partner may create pressure that induces a batter to attack the focal bowler. A wicket-taking partner may expose the focal bowler to a new batter. Different styles may limit adaptation. Communication and familiarity may improve tactical coordination. The pair operates sequentially rather than simultaneously. But each over changes the match situation that the bowler at the other end faces next. The partner is deliberately defined as a single bowler. Longer tactical units may involve several bowlers. The preceding-over definition has the advantage that every focal over gets a precisely timed coworker.

The captain chooses both bowlers and their timing. The choice is constrained. The same bowler cannot bowl consecutive overs. Limited-overs formats cap each bowler's total overs. Yet those constraints do not make realised partner identity exogenous. Captains decide when a bowler

[1] Sources: *Wisden Cricketers' Almanack* 1963, 'Stars of the Tests', accessed through ESPNcricinfo's Almanack archive; *The Cricket Statistician* No. 66 (Summer 1989), book-review pages, reviewing David Lemmon's *The History of Worcestershire County Cricket Club*; Stuart Broad quoted in Oliver Yew, 'Are James Anderson and Stuart Broad England's greatest bowling pair?', Sky Sports, 9 June 2016; Arun Gopalakrishnan, 'Hunting in Pairs: 5 Bowling Duos Who Cemented a Place in History', The Quint, 12 August 2020, which calls the saying 'a famous phrase used in cricket commentary' that 'has stood the test of time'; ESPNcricinfo ball-by-ball commentary for Ireland v Scotland, ODI, 12 January 2015 ('hunting in pairs') and Pakistan v South Africa, ODI, 19 December 2024 ('sustained pressure from both ends'). Phrase counts by corpus are in the replication materials.

begins, how long a spell lasts, who operates at the other end and how to respond to batters and match state. A partner change can coincide with an old ball, a new batting objective, fatigue or an intentional change of attack.

2.2 Why five settings are informative

Tests and domestic multi-day cricket allow long spells and repeated pair exposure. They are the settings in which sustained partnerships are most feasible and most prominent in cricket strategy. ODIs cap a bowler at ten overs in a fifty-over innings. T20 cricket generally caps a bowler at four overs in a twenty-over innings. T20Is and franchise T20 also contain more player turnover and shorter observed histories.

These differences make cross-setting comparison useful. They do not make it causal. Format changes many things at once: the wicket base rate, scoring objectives, roster composition, assignment frequency, over caps and the amount of prior data available for player-quality estimation. I therefore use the five settings as replications of a measurement exercise. A formal gradient must survive common measurement and experience support before it can motivate an institutional interpretation.

3 Data and construction

3.1 Match coverage

The source files are structured ball-by-ball match records. International and franchise coverage contains 10,774 matches across the four international and franchise categories. The domestic build adds 2,162 multi-day matches from the County Championship, Sheffield Shield, New Zealand's Plunket Shield, the Bob Willis Trophy and Sri Lanka's Major League Tournament. Table 1 reports descriptive statistics for the five settings. Franchise T20 includes 167 Hundred matches, about 3.6% of the franchise sample. Domestic coverage starts at different dates for different competitions. Long-horizon exposure histories are therefore noisier in domestic cricket than in Tests.

Every match in the domestic sample is read successfully, with no duplicate match-innings-over keys, no ambiguous player identities and no bowler appearing in two domestic matches on the same date. A small number of overs have more than one bowler. The international and domestic samples apply the same rule for these overs.

3.2 Focal overs and partners

An observation is a focal bowler-over. It enters when:

1. the focal over has at least one legal delivery;
2. the immediately preceding over exists in the same innings;
3. focal and partner bowlers differ;

4. neither the partner-defining over nor the focal over is a mixed-bowler over;

5. the observation is not a super over; and

6. pre-match partner-quality information can be constructed.

Mixed-bowler overs are rare. Excluding both the mixed over and the over after it avoids two errors: crediting an entire over to the bowler of its first delivery, and defining a partner from that attribution. For The Hundred, the same exclusion is applied to the focal unit and the over closing the prior ten-ball end. The resulting eligible corpus contains 1,355,340 overs. Table 1 reports the slightly smaller samples that remain after fixed-effect singleton removal.

The primary outcome is bowler-credited wickets divided by legal balls in the focal over. All bowler-credited dismissals recorded on a delivery are counted, including a stumping on a wide. Retired hurt does not count as a dismissal and does not reduce wickets in hand. Wickets are rare and reward only one dimension of bowling. Runs scored at the focal end per legal ball therefore provides a second output measure. This total includes byes, leg-byes and penalty runs. It is therefore not a measure of the bowler's economy alone. Outcome and eligibility rules are identical for the international and domestic samples.

3.3 Predetermined partner quality

Let W_{jt} and B_{jt} denote partner j's wickets and legal balls in matches dated strictly before match t. Histories are setting-specific for Tests, ODIs and domestic multi-day cricket. T20I and franchise observations instead use a pooled all-T20 history. A player's record in one T20 competition then also informs his measured quality in the other. Partner wicket quality in setting s is

$$q_{jt}^{s} = \frac{W_{jt} + 300\,\bar{p}_s}{B_{jt} + 300},$$

where 300 is the prior strength and $\bar{p}_s$ is the wicket rate in setting s's full-period eligible corpus. The prior mean is a fixed calibration constant estimated once from the full eligible corpus. The player-specific wicket and ball counts exclude the focal date. The treatment is standardised within setting after construction. The setting-specific prior applies even when the T20 counts pool T20I and franchise histories.

The setting-specific prior is important. A natural alternative fixes the same small set of prior rates across settings. Relative to the realised quality distributions, such fixed priors sit above the mean in the two long-form settings and well below it in the white-ball settings. In franchise T20 the gap is nearly two standard deviations. The same 'no history' condition would therefore imply an above-average partner in long-form cricket and a far-below-average partner in short cricket. The measure used here instead recentres the prior on each setting's eligible-corpus rate. Table 3 reports both priors and the share of observations whose partner has fewer than 1,000 prior balls.

This recentring does not solve the assignment problem. It removes a mechanical reason that sparse partners are ranked differently across settings. The analysis therefore also reports a $\geq$1,000-

ball sample and four within-setting quartiles of partner prior balls, without restandardising quality within those selected samples.

4 Empirical framework

4.1 Conditional association

For focal bowler i in match m, innings n and over o, I estimate

$$y_{imno} = \beta_s \, q^{s}_{jt(m)} + g_s(o) + \rho_s \, h_{imo} + \alpha_{im} + \delta_{mn} + \varepsilon_{imno},$$

where q is pre-match partner quality, $g_s(o)$ is a flexible five-over innings-position profile, h is the number of the focal bowler's own over in the innings, α_{im} is a focal-bowler-by-match fixed effect and δ_{mn} is a match-innings fixed effect. The model is weighted by legal balls. Standard errors are clustered by focal bowler and match.

The coefficient compares the same focal bowler within the same match when the opposite-end partner's pre-match quality changes. Table 1 shows how much variation this leaves. A focal bowler works with about three and a half different partners in an average long-form match and with about two in T20. Match-innings fixed effects remove common conditions. Focal-bowler-by-match effects remove player-match form or selection. The timing controls address two remaining sources of covariance visible before the focal over: innings phase and the focal bowler's workload position.

I call this the timing specification, not a causal specification. Captains retain private information. Observable timing cannot absorb batter-specific matchups, tactical intentions or unrecorded ball condition. The estimand is a conditional association. It is useful for testing whether a partnership interpretation is stable under predetermined design controls.

4.2 Nested specifications

Table 2 reports three specifications. The first contains only the two sets of fixed effects. It represents the conventional comparison. The second adds innings position and focal-over number and is primary. The third adds wickets in hand and the score rate at the focal-over start.

The third column is a bracket rather than a superior adjustment. Current state can confound partner assignment. But earlier partners may themselves have changed wickets and scoring. Controlling for cumulative state can therefore block part of a genuine sequential channel. Separating timing from cumulative state avoids treating every additional control as automatically more causal.

4.3 Criteria for interpretation

The data permit many plausible partner definitions, formats, states and subgroups. Two criteria therefore govern the interpretation of the estimates.

The first applies to the long-form estimate. If the pooled long-form timing estimate lies within $\pm 1\%$ of the base wicket rate and its 95% interval lies inside $\pm 2\%$, that interval is reported as a bound on the conditional association, and the search for a current-production effect ends there.

The second applies to the cross-format pattern. An institutional gradient across formats is claimed only if equality of the five slopes is rejected at 5% in three samples: the full eligible sample, the sample of partners with at least 1,000 prior balls, and the sample that excludes the least-experienced quartile. A pattern that holds in some of these samples but not others is reported as descriptive.

These criteria are not a registered analysis plan. Their value lies in the completeness of what is reported against them. The tables contain every estimate to which the criteria refer: five settings, two outcomes, three specifications and four experience quartiles. A reader who applies a different criterion can do so with the same tables.

5 Results

5.1 The conventional pattern

The fixed-effects-only estimates in Figure 1 look like strong evidence of format heterogeneity. In Tests and domestic multi-day cricket, a one-standard-deviation increase in partner quality predicts a wicket rate two to three per cent of the setting mean lower. The signs reverse in white-ball cricket. The association also grows as the format shortens. It reaches more than five per cent of the mean in franchise T20. Each estimate except the Test-to-domestic difference is precise enough to suggest a format-specific interpretation.

The same warning appears in the run results. With fixed effects only, partner quality also predicts noticeably more runs per ball in ODIs and franchise T20, alongside smaller long-form relationships. A high-quality partner may create pressure that shifts batters' risk-taking toward the focal end. Wickets and runs can then rise together. But this mechanism is only one possible interpretation of a common phase covariance.

5.2 The association under timing controls

The open markers in Figure 1 add only variables known before the focal over: innings-position bins and the focal bowler's over number. The two long-form estimates fall to a fraction of one per cent of the mean and are no longer distinguishable from zero. The ODI association disappears. The T20I and franchise associations remain. Both are now well under half their former size.

Pooling Tests and domestic multi-day cricket while allowing setting-specific timing profiles yields -0.43% of the pooled mean wicket rate, with a 95% interval from -1.30% to $+0.45\%$. Adding cumulative state moves the point estimate just above zero, with a similarly narrow interval. Both estimates are centred close to zero. By the first criterion in Section 4.3, this interval is reported as a bound on the conditional association.

The result is best understood as a bound on the conditional association remaining after the timing controls. It rejects the empirical importance of the large negative long-form coefficient, not every possible pair complementarity. A benefit specific to particular bowling styles could average to zero. Endogenous matching could attenuate or reverse a causal effect. Quality measured by wicket rate may miss teaching or tactical value. Those are limits of the estimand, not reasons to restore the discarded coefficient.

The run results support the timing interpretation. The preferred long-form setting estimates are essentially zero. The ODI and franchise run associations remain positive but shrink by roughly two-thirds to three-quarters. The evidence is consistent with phase and tactical assignment affecting both dimensions of output.

5.3 Measurement support and the cross-format gradient

The full-sample timing estimates remain jointly different across settings, and the $\geq$1,000-prior-ball sample also rejects equality. Those tests alone would support continued study of an institutional gradient. They are not the complete criterion.

Figure 2 reports the same timing specification by within-setting quartile of partner prior balls. Tests, domestic multi-day cricket and ODIs show no systematic relationship in any quartile. In T20Is, the positive estimate is concentrated in quartile 1: +10.31% of the mean, based on a median of only 24 prior T20-family balls. At that median, the partner's own record supplies just 7.4% of the quality measure. The prior supplies the rest. Quartiles 2–4 are individually indistinguishable from zero and do not form a monotone sequence. Franchise T20 instead has a positive estimate only in the most-experienced quartile. When quartile 1 is excluded across settings, the formal equality test has $p = 0.0525$. Figure 3 places every setting-by-quartile cell on the shrinkage weight curve and plots each cell's estimate against the data share behind its treatment. The T20I quartile-1 estimate stands alone at the uninformative end of the curve.

This number is not interpreted as 'almost significant'. It fails the second criterion in Section 4.3. More importantly, the underlying cells do not describe one stable experience mechanism. The short-format pattern changes location between T20Is and franchise cricket. The paper therefore drops the institutional gradient and does not claim that over caps or reassignment frequency create the remaining association.

6 Durable pair capital

The main regressions study partner quality. A productive partnership could instead be pair-specific: two ordinary bowlers may be unusually effective together, or familiarity may create value not captured by individual quality. The long-form panel is the strongest place to look because pairs have time to recur. I construct all dyad measures at the date level and lag cumulative stocks, excluding the entire focal date.

6.1 Familiarity

The first test adds $\log(1 + \text{prior pair balls}/300)$. Familiar pairs are selected. Successful pairs may be kept together. Stable teams may differ from unstable teams. The estimate is therefore not a treatment effect of familiarity. It asks only whether accumulated pair exposure adds predictive information beyond partner quality and the state controls.

In 882,053 pooled long-form overs, the estimate is small and statistically indistinguishable from zero. It stays so under partner-and-match and pair-and-match clustering. The conclusion therefore does not depend on treating repeated pair observations as independent. Setting-level estimates are unstable and opposite in sign. Test and domestic coefficients sometimes differ in the models without state controls. Neither setting has a detectable state-adjusted pressure or pair-excess association.

6.2 A cross-dimensional pressure channel

Cricket's most plausible immediate mechanism is pressure. A partner who concedes few runs may change batters' risk-taking against the focal bowler, potentially raising both wickets and runs at the focal end. I construct the partner's pre-match run-suppression quality and ask whether it predicts focal wickets conditional on the partner's wicket quality.

The estimate has the opposite sign to the simple pressure prediction. It is too small and imprecise to interpret in either direction. Alternative clustering leaves the conclusion unchanged.

6.3 Held-out pair excess

The strongest validation test asks whether a pair's own past excess performance recurs. For each pair, I aggregate wickets in matches before the focal date and subtract the performance predicted by the two bowlers' individual pre-match quality. The resulting pair-excess stock is standardised and used to predict held-out focal output.

The estimate is as close to zero as the design can measure. Its 95% interval spans less than one per cent of the pooled base rate in either direction. The interval concerns persistence in this pair-excess measure. It is not a universal causal bound on all partnership mechanisms. Still, if durable pair capital existed, it should show up in held-out output. It does not.

Together, the three tests yield no evidence that a stable pair component explains the conditional association. Familiarity may be endogenous. Cross-dimensional quality may be incomplete. Pair excess is noisy. Their value is triangulation: three measures with different errors all fail to restore the strong partnership interpretation suggested by the fixed-effects-only model.

7 Injury as a power diagnostic

An injury that stops a bowler mid-over appears to offer an exogenous partner disruption. The immediate substitute, however, is often a temporary bowler who merely completes the over. He is

not the new sustained partner. The relevant design therefore follows the established bowler at the opposite end until the captain selects a planned replacement partner.

This sequence is rare. Tests and domestic multi-day cricket supply only 93 clean long-form events in which the established opposite-end bowler returns and later operates with the planned replacement. The estimated change is small and imprecise. The design's 80%-power minimum detectable effect is nearly 30 times the largest timing-controlled setting coefficient and larger than the mean long-form wicket rate itself.

The design also conditions on the focal bowler returning after the disruption, which is a post-injury selection margin. Injury timing can also correlate with fatigue or workload. The exercise is therefore presented only as a power and feasibility result. Its null cannot validate the main estimate. Its positive point estimate cannot be read as an injury effect.

8 Conclusion

Does the quality of the bowler at the other end change what a bowler produces? The evidence from five professional cricket settings supports a plain answer: not detectably, once the timing of the pairing is held fixed. Coworker quality is not itself a coworker effect.

The preferred pooled long-form association is -0.43% of the mean wicket rate per standard deviation of partner quality, with a 95% interval from -1.30% to $+0.45\%$. That interval is narrow relative to the estimates it replaces: every fixed-effects-only setting coefficient lies outside it. The remaining short-format pattern fails the study's common-experience support rule. Familiarity, a pressure channel and prior pair-specific excess performance do not predict held-out long-form output.

For coworker research the first lesson is about timing. The literature emphasises reflection, sorting and common shocks. This setting shows that manager-chosen timing can be the principal confounder even after worker-by-workplace fixed effects. Rotations occur within those cells. Who works with whom matters. So does when the pair is activated within a production cycle.

The second lesson is about measurement. Shrinkage is useful because raw quality is noisy. But the prior has economic content for workers with little history. If inexperienced workers enter different tasks, teams or market segments, the prior's placement can create a treatment gradient that resembles institutional heterogeneity. Reporting prior strength without reporting the prior's location in each estimation sample is therefore insufficient. Together with common experience support, these are not robustness checks to run after a headline coefficient is found. Each can produce a coefficient that looks like a partnership effect.

For team analysts the implication is not that coaches should ignore partnerships. It is that retrospective pair statistics and teammate-quality regressions can value a player for when the captain deploys him rather than for what he produces in others. That bias remains even in comparisons within player and match. Pair metrics should be validated on future matches, separated from innings and workload timing, and evaluated on common experience support.

The same distinction affects recruitment and selection. A player who appears to improve teammates may simply be assigned during favourable phases or beside the first-choice attack. Conversely, a bowler used as a defensive partner may create value through runs prevented that a wicket-only measure misses. Multidimensional player valuation remains warranted. But a teammate spillover should predict held-out teammate performance before it is priced as portable pair capital.

The caution extends to format comparisons. Short formats have more frequent partner changes, binding over limits and sparse international histories. They also have larger changes in batting intent across innings phases. A larger T20 coefficient is therefore not evidence of stronger complementarity without a design that separates all three. The failed common-support rule prevents the paper from converting descriptive differences into a causal institutional claim.

The result holds for current output in men's professional cricket across the five settings studied. The estimand is a conditional average association. Average nulls can hide complementarity between particular bowling styles, handedness matchups or match states. The design also does not test the slower accumulation of skill from coworkers. These conclusions complement research with exogenous changes in interaction. Emanuel, Harrington and Pallais (2026) use office closures and return-to-office changes to show that proximity alters feedback and junior work quality while costing senior output. Sandvik et al. (2020) experimentally encourage advice-seeking and find persistent gains. The present study has richer task timing but no comparable exposure shock. Where exposure is set by a shock rather than by a manager's minute-to-minute choices, the confounder documented here is absent. That is why granularity cannot substitute for that source of identification.

What would settle the remaining question is variation the captain does not control. Two designs would identify what this archive cannot. The first is a pre-specified test of style complementarity built on external bowling classifications fixed before estimation. The second is an institutional shock to pair availability, such as a rule change that removes one member of established pairs. Until then, a partnership claim should be required to predict held-out output. This one did not.

Data and replication statement

The paper uses publicly available ball-by-ball match records. The data and the code that reproduce all tables and figures will be made available on the author's GitHub page.

References

Arcidiacono, Peter, Josh Kinsler, and Joseph Price. 2017. "Productivity Spillovers in Team Production: Evidence from Professional Basketball." *Journal of Labor Economics*, 35(1): 191–225.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2023. "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*, 41(2): 291–322.

Chathurangi, A. K. D. K., R. M. Silva, N. Withanage, and C. L. Jayasinghe. 2025. "Impact Ranking Methodologies in Limited-Overs Cricket: A Systematic Review of Performance Metrics." *International Journal of Sports Science & Coaching*, 20(3): 1307–1319.

Chen, Jiafeng. 2026. "Empirical Bayes When Estimation Precision Predicts Parameters." *Econometrica*, 94(2): 305–340.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104(9): 2593–2632.

Cohen-Zada, Danny, Itay Dayag, and Naomi Gershoni. 2024. "Effort Peer Effects in Team Production: Evidence from Professional Football." *Management Science*, 70(4): 2355–2381.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg. 2017. "Peer Effects in the Workplace." *American Economic Review*, 107(2): 425–456.

Emanuel, Natalia, Emma Harrington, and Amanda Pallais. 2026. "The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?" *Quarterly Journal of Economics*, 141(3): 1825–1870.

Gould, Eric D., and Eyal Winter. 2009. "Interactions between Workers and the Technology of Production: Evidence from Professional Baseball." *Review of Economics and Statistics*, 91(1): 188–200.

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1(4): 34–68.

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips. 2024. "Production and Learning in Teams." *Econometrica*, 92(2): 467–504.

Hickman, Daniel C., and Neil E. Metz. 2018. "Peer Effects in a Competitive Environment: Evidence from the PGA Tour." *Economic Inquiry*, 56(1): 208–225.

Jäger, Simon, and Jörg Heining. 2022. "How Substitutable Are Workers? Evidence from Worker Deaths." National Bureau of Economic Research Working Paper 30629.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2021. "Learning from Coworkers." *Econometrica*, 89(2): 647–676.

Jiang, Lingqing. 2020. "Splash with a Teammate: Peer Effects in High-Stakes Tournaments." *Journal of Economic Behavior and Organization*, 171: 165–188.

Kahn, Lawrence M. 2000. "The Sports Business as a Labor Market Laboratory." *Journal of Economic Perspectives*, 14(3): 75–94.

Kline, Patrick, Evan K. Rose, and Christopher R. Walters. 2022. "Systemic Discrimination Among Large U.S. Employers." *Quarterly Journal of Economics*, 137(4): 1963–2036.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-Out Estimation of Variance Components." *Econometrica*, 88(5): 1859–1898.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton. 2015. "The Value of Bosses." *Journal of Labor Economics*, 33(4): 823–861.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–145.

Minni, Virginia. 2026. "Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs." *Quarterly Journal of Economics*, 141(3).

Palacios-Huerta, Ignacio. 2025. "The Beautiful Dataset." *Journal of Economic Literature*, 63(4): 1363–1423.

Sandvik, Jason, Richard Saouma, Nathan Seegert, and Christopher Stanton. 2020. "Workplace Knowledge Flows." *Quarterly Journal of Economics*, 135(3): 1635–1680.

Walters, Christopher R. 2024. "Empirical Bayes Methods in Labor Economics." In *Handbook of Labor Economics*. Vol. 5 of *Handbooks in Economics*, ed. Christian Dustmann and Thomas Lemieux, 183–260. Amsterdam:Elsevier.

Tables

Table 1. Descriptive statistics

	Test	Domestic multi-day	ODI	T20I	Franchise T20
Panel A. Sample					
Matches	885	2,162	2,547	3,442	3,900
Eligible overs	279,535	603,854	215,093	117,946	138,912
Regression N	279,155	602,898	214,042	112,490	133,447
Focal bowlers	781	1,541	1,463	3,200	1,626
Bowling pairs	4,662	12,334	9,354	14,812	19,643
Panel B. Output					
Wickets per legal ball	0.0163	0.0175	0.0261	0.0556	0.0515
Wickets per over	0.098	0.105	0.156	0.330	0.305
standard deviation	(0.315)	(0.323)	(0.391)	(0.554)	(0.530)
Runs per legal ball	0.549	0.560	0.874	1.272	1.396
standard deviation	(0.504)	(0.518)	(0.621)	(0.793)	(0.810)
Legal balls per over	5.97	5.97	5.96	5.94	5.91
Panel C. Partner quality (the treatment)					
Mean	0.0168	0.0179	0.0267	0.0553	0.0517
Standard deviation	0.0035	0.0034	0.0049	0.0062	0.0063
10th percentile	0.0126	0.0136	0.0205	0.0476	0.0437
90th percentile	0.0210	0.0222	0.0330	0.0632	0.0595
Partner's prior balls (median)	2,677	2,784	998	269	828
Partner under 1,000 balls (%)	28.4	26.4	50.1	81.5	55.4
Panel D. Timing and match state					
Focal bowler's over number	12.26	10.23	4.73	2.28	2.33
standard deviation	(9.33)	(7.46)	(2.64)	(1.05)	(1.06)
Wickets in hand	6.52	6.38	7.02	7.21	7.42
Runs per over before the over	3.16	3.15	4.69	7.09	7.57
Overs per bowler-match	27.2	22.2	7.05	2.91	3.02
Partners per bowler-match	3.49	3.51	2.78	2.08	2.29

Note. The unit of observation is the focal bowler-over, and all statistics are computed over the eligible overs of each setting. Eligible overs precede fixed-effect singleton removal; regression N is the number retained by the timing specification. Wickets are bowler-credited dismissals. Runs are those scored at the focal end and include byes, leg-byes and penalty runs. Partner quality is the empirical-Bayes wicket rate of the opposite-end bowler, built only from matches dated before the focal match and reported here in levels rather than standardised. The last two rows describe the variation the fixed effects use: the number of overs a focal bowler delivers in a match, and the number of distinct opposite-end partners he works with in it. The domestic setting pools five first-class competitions, listed in Section 3.

Table 2. Partner quality and focal-bowler output under nested specifications

Setting	Outcome	Specification	Estimate (SE)	% of mean	p value	N
Test	Wickets/ball	Fixed effects only	−0.000385 (0.000146)	−2.36	0.008	279,155
Test	Wickets/ball	Timing controls	−0.000127 (0.000154)	−0.78	0.409	279,155
Test	Wickets/ball	Timing plus cumulative state	−0.000099 (0.000154)	−0.61	0.519	279,155
Domestic multi-day	Wickets/ball	Fixed effects only	−0.000490 (0.000088)	−2.80	0.000	602,898
Domestic multi-day	Wickets/ball	Timing controls	−0.000049 (0.000106)	−0.28	0.643	602,898
Domestic multi-day	Wickets/ball	Timing plus cumulative state	+0.000120 (0.000109)	+0.69	0.270	602,898
ODI	Wickets/ball	Fixed effects only	+0.000978 (0.000208)	+3.75	0.000	214,042
ODI	Wickets/ball	Timing controls	−0.000122 (0.000201)	−0.47	0.544	214,042
ODI	Wickets/ball	Timing plus cumulative state	+0.000097 (0.000203)	+0.37	0.634	214,042
T20I	Wickets/ball	Fixed effects only	+0.002306 (0.000409)	+4.15	0.000	112,490
T20I	Wickets/ball	Timing controls	+0.000959 (0.000399)	+1.73	0.016	112,490
T20I	Wickets/ball	Timing plus cumulative state	+0.001042 (0.000393)	+1.87	0.008	112,490
Franchise T20	Wickets/ball	Fixed effects only	+0.002904 (0.000368)	+5.64	0.000	133,447
Franchise T20	Wickets/ball	Timing controls	+0.000804 (0.000347)	+1.56	0.021	133,447
Franchise T20	Wickets/ball	Timing plus cumulative state	+0.001016 (0.000345)	+1.97	0.003	133,447
Test	Runs/ball	Fixed effects only	−0.001501 (0.001338)	−0.27	0.262	279,155
Test	Runs/ball	Timing controls	−0.000027 (0.001364)	−0.00	0.984	279,155
Test	Runs/ball	Timing plus cumulative state	−0.000364 (0.001352)	−0.07	0.788	279,155
Domestic multi-day	Runs/ball	Fixed effects only	−0.003370 (0.000928)	−0.60	0.000	602,898
Domestic multi-day	Runs/ball	Timing controls	+0.000473 (0.001089)	+0.08	0.664	602,898
Domestic multi-day	Runs/ball	Timing plus cumulative state	+0.000274 (0.001098)	+0.05	0.803	602,898
ODI	Runs/ball	Fixed effects only	+0.016513 (0.001959)	+1.89	0.000	214,042
ODI	Runs/ball	Timing controls	+0.006158 (0.001655)	+0.70	0.000	214,042
ODI	Runs/ball	Timing plus cumulative state	+0.008348 (0.001640)	+0.96	0.000	214,042
T20I	Runs/ball	Fixed effects only	+0.009768 (0.003282)	+0.77	0.003	112,490
T20I	Runs/ball	Timing controls	−0.001238 (0.003126)	−0.10	0.692	112,490
T20I	Runs/ball	Timing plus cumulative state	−0.000248 (0.003169)	−0.02	0.938	112,490
Franchise T20	Runs/ball	Fixed effects only	+0.031937 (0.003064)	+2.29	0.000	133,447
Franchise T20	Runs/ball	Timing controls	+0.008246 (0.002727)	+0.59	0.002	133,447
Franchise T20	Runs/ball	Timing plus cumulative state	+0.010761 (0.002740)	+0.77	0.000	133,447

Note. The unit of observation is the focal bowler-over. Quality is the setting-calibrated empirical-Bayes wicket rate, standardised within setting. T20I and franchise observations use pooled prior T20-family sufficient statistics; other histories are setting-specific. All models contain match-innings and focal-bowler-by-match fixed effects, use legal-ball weights and cluster by focal bowler and match. The timing specification is primary; cumulative state is a bracket.

Table 3. Quality measurement and common experience support

Panel A. Prior calibration

Setting	Base wicket rate	Fixed prior	Fixed prior position (SD)	Calibrated prior position (SD)	Partner under 1,000 balls (%)
Test	0.0163	0.018	+0.25	−0.12	28.4
Domestic multi-day	0.0175	0.020	+0.47	−0.10	26.4
ODI	0.0261	0.020	−0.88	−0.12	50.1
T20I	0.0556	0.025	−1.32	+0.04	81.5
Franchise T20	0.0515	0.025	−1.78	−0.03	55.4

Panel B. Timing estimates by partner-experience quartile

Setting	Quartile	Median prior balls	Data weight at median (%)	Estimate (% of mean)	95% CI	p value	N
Test	Q1	258	46.2	−0.92	[−5.70, 3.86]	0.705	69,047
Test	Q2	1,610	84.3	−0.91	[−5.95, 4.13]	0.723	69,336
Test	Q3	4,209	93.3	−0.05	[−5.87, 5.78]	0.988	69,445
Test	Q4	11,330	97.4	+0.74	[−6.14, 7.62]	0.833	69,549
Domestic multi-day	Q1	336	52.8	−1.57	[−4.01, 0.86]	0.206	148,930
Domestic multi-day	Q2	1,737	85.3	−0.45	[−3.08, 2.19]	0.739	149,374
Domestic multi-day	Q3	4,222	93.4	+0.34	[−3.80, 4.47]	0.874	149,758
Domestic multi-day	Q4	9,908	97.1	−1.53	[−5.79, 2.73]	0.482	150,160
ODI	Q1	116	27.9	−2.85	[−8.26, 2.56]	0.302	50,529
ODI	Q2	607	66.9	+0.80	[−3.83, 5.43]	0.735	50,599
ODI	Q3	1,536	83.7	−1.28	[−5.42, 2.86]	0.545	50,727
ODI	Q4	3,569	92.2	−0.74	[−4.68, 3.19]	0.711	51,291
T20I	Q1	24	7.4	+10.31	[2.38, 18.24]	0.011	23,041
T20I	Q2	156	34.2	+0.75	[−3.37, 4.87]	0.721	21,838
T20I	Q3	438	59.3	+2.32	[−1.78, 6.42]	0.267	21,726
T20I	Q4	1,419	82.5	−1.67	[−4.72, 1.38]	0.283	24,640
Franchise T20	Q1	106	26.1	+0.50	[−4.79, 5.79]	0.853	23,239
Franchise T20	Q2	507	62.8	−0.67	[−5.17, 3.82]	0.769	22,961
Franchise T20	Q3	1,263	80.8	+0.52	[−3.14, 4.18]	0.781	23,594
Franchise T20	Q4	2,992	90.9	+3.79	[0.20, 7.37]	0.038	24,798

Note. The unit of observation is the focal bowler-over; standard errors cluster by focal bowler and match. Quartiles are defined from the within-setting rank of prior balls. Quality is not restandardised within a quartile.

Table 3 (continued)

Panel C. Formal equality tests across settings

Treatment	Sample	χ^2	df	p value	N
Fixed prior	Full eligible sample	10.98	4	0.0268	1,342,032
Calibrated prior	Full eligible sample	12.57	4	0.0136	1,342,032
Calibrated prior	Partner $\geq$1,000 prior balls	12.69	4	0.0129	815,668
Calibrated prior	Experience quartiles 2–4	9.37	4	0.0525	996,668

Note. Equality tests are joint tests that the five setting slopes are equal. An institutional gradient is claimed only if equality is rejected in all three calibrated support samples.

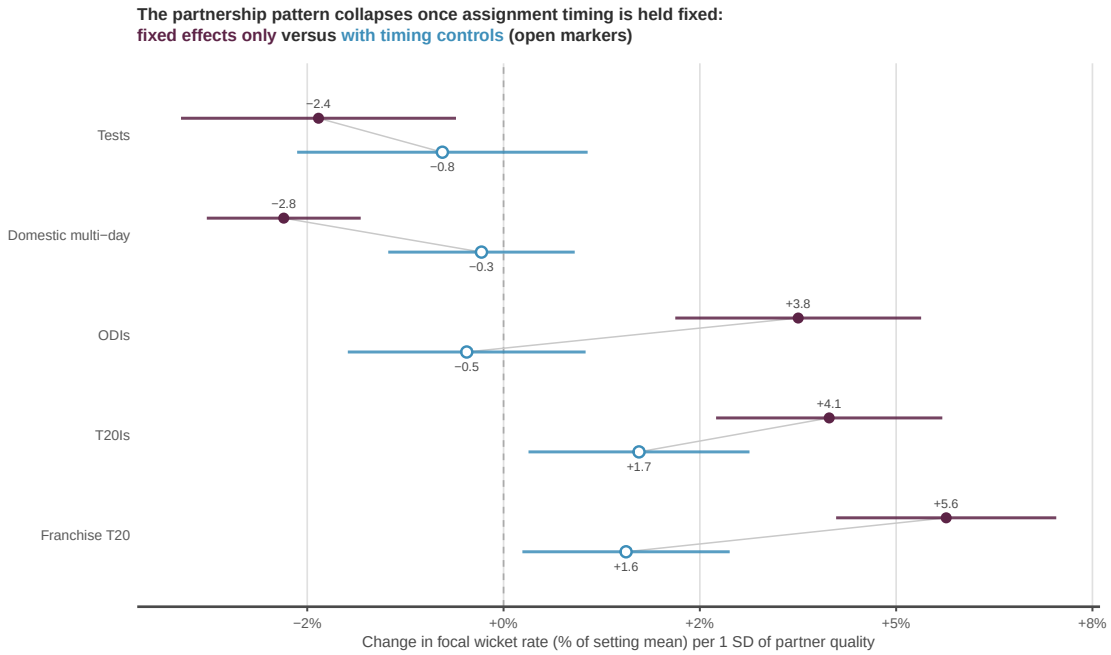
Table 4. Prospective long-form dyad validation

Test	Estimate (SE)	95% CI	p value	N
Prior-match pair familiarity	+0.000153 (0.000131)	$[-0.000103, +0.000409]$	0.243	882,053
Pressure cross-effect	−0.000033 (0.000078)	$[-0.000186, +0.000120]$	0.674	882,053
Held-out prior pair excess	+0.000001 (0.000073)	$[-0.000141, +0.000144]$	0.984	882,053

Note. The sample is pooled long-form overs (Tests and domestic multi-day cricket). Measures use dates strictly before the focal match. Models include partner quality, start-state controls, match-innings fixed effects and focal-bowler-by-match fixed effects. Standard errors shown cluster by focal bowler and match. Clustering instead by partner and match, or by pair and match, leaves every conclusion in the table unchanged.

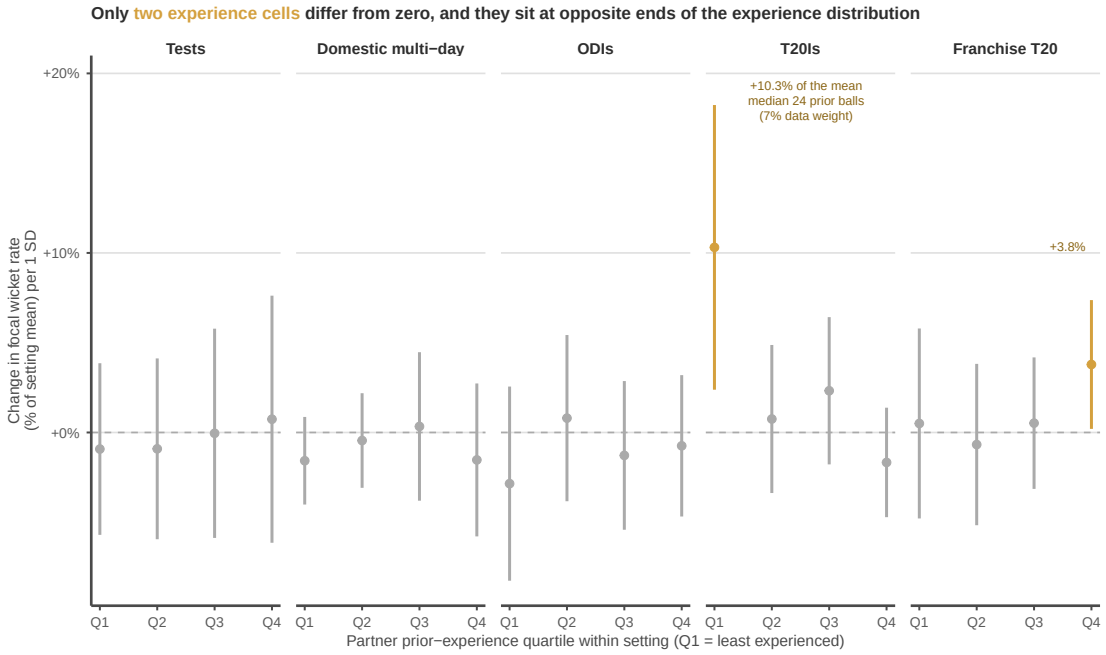
Figures

Figure 1. Timing sensitivity of the coworker-quality association



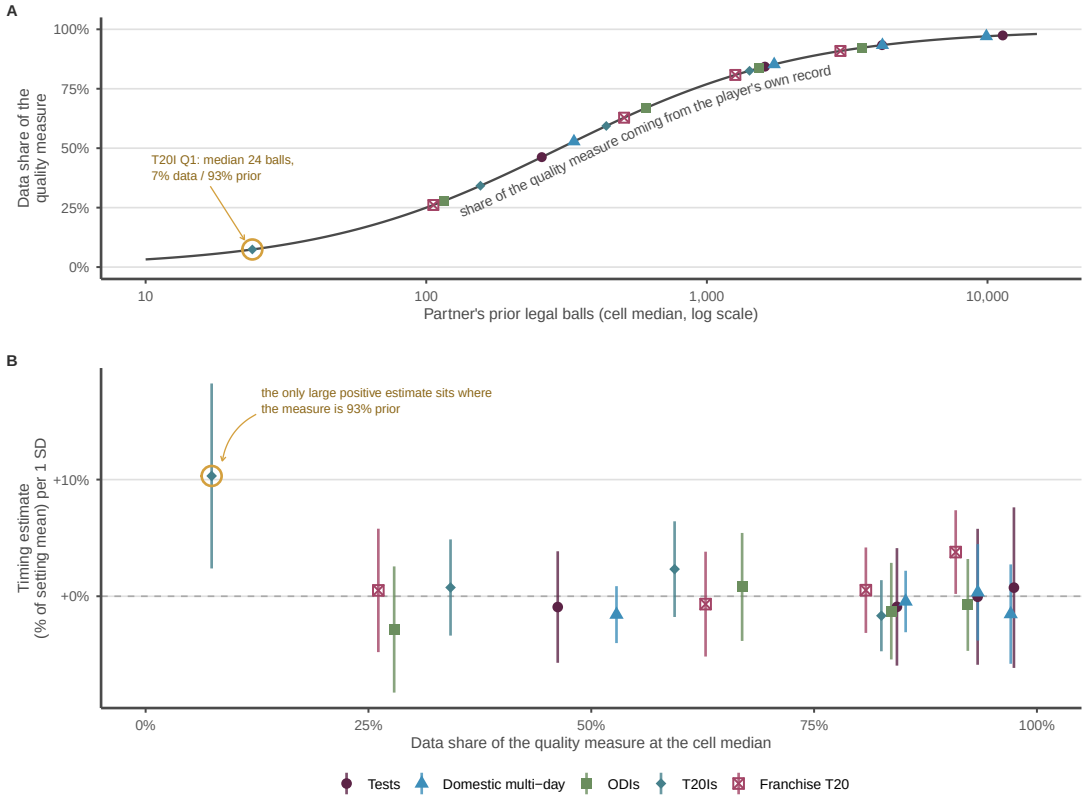
Note. The unit of observation is the focal bowler-over. Filled markers report the fixed-effects-only specification; open markers add innings-position bins and the focal bowler's own over number. Estimates are expressed as a percentage of the setting mean wicket rate, with 95% confidence intervals. All models contain match-innings and focal-bowler-by-match fixed effects, use legal-ball weights and cluster by focal bowler and match. Sample sizes are reported in Table 1.

Figure 2. Common experience support



Note. Timing-specification estimates by within-setting quartile of the partner's prior legal balls, expressed as a percentage of the setting mean wicket rate, with 95% confidence intervals. The unit of observation is the focal bowler-over; standard errors cluster by focal bowler and match. Cells whose interval excludes zero are highlighted; all others are grey. Cell medians, data weights and sample sizes are reported in Table 3, Panel B.

Figure 3. The shrinkage window: how much data sits behind the treatment



Note. Panel A shows the empirical-Bayes weight $B/(B+300)$ implied by the prior strength of 300, the share of the quality measure coming from the partner's own record at B prior legal balls. The markers place the median prior balls of each setting-by-quartile cell from Table 3, Panel B on this curve. Panel B plots each cell's timing-specification estimate, expressed as a percentage of the setting mean wicket rate with 95% confidence intervals, against the data share at the cell median. The unit of observation is the focal bowler-over; standard errors cluster by focal bowler and match. The highlighted cell is T20I quartile 1.