

Pulling the Power Play: A Win Probability Strategy for Mixed Doubles Curling

Trey Elder (MSE '27), Samen Hossain (CAS '27), & Jonathan Pipping

January 28, 2026

What is Mixed Doubles Curling?

Ends. Mixed doubles games are eight ends long—like innings, but on ice. In each end, both teams alternate throwing five stones apiece.

Scoring an end. Only one team can score points in each end: the team with the stone closest to the button. That team scores one point for each of their stones in the house that is closer than the opponent’s nearest stone. If nobody has a stone in the house, the end is said to be “blank,” and no points are scored.

Hammer. The team that throws last in an end has a big advantage: they control the final shot. Under mixed doubles rules, the hammer passes to the team that did *not* score in the previous end; if the end is blank, the hammer switches to the team that did not throw last.

Pre-positioned stones. Every end starts with two stones already in play: one as a guard and one behind the button. Figure 1 shows the standard mixed doubles setup at the start of an end.

Winning the game. After eight ends, the team with more points wins; if tied, play continues with sudden-death extra ends until one team scores.

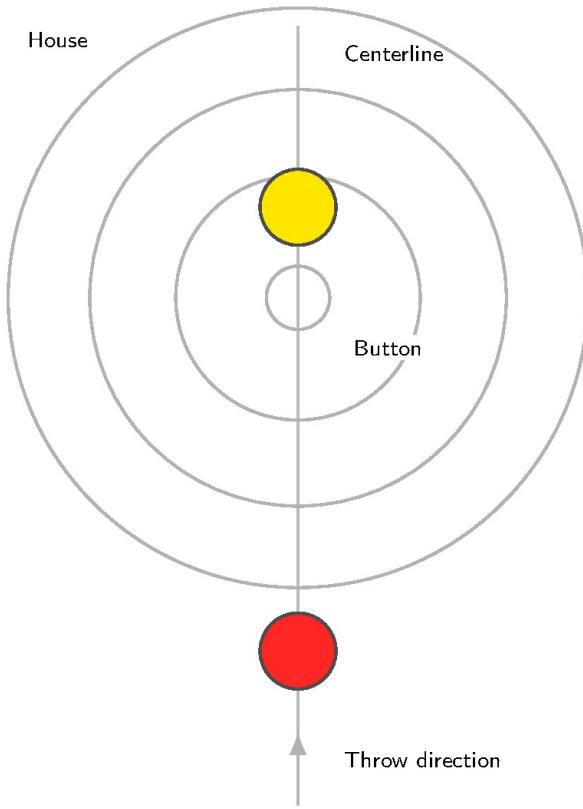


Figure 1: Stone orientation at the start of an end.

What is the Power Play?

At the start of an end, the team with the hammer may call a Power Play. When they do, the two pre-positioned stones shift to one side, opening a cleaner scoring lane. Figure 2 contrasts the regular setup with the Power Play configuration.

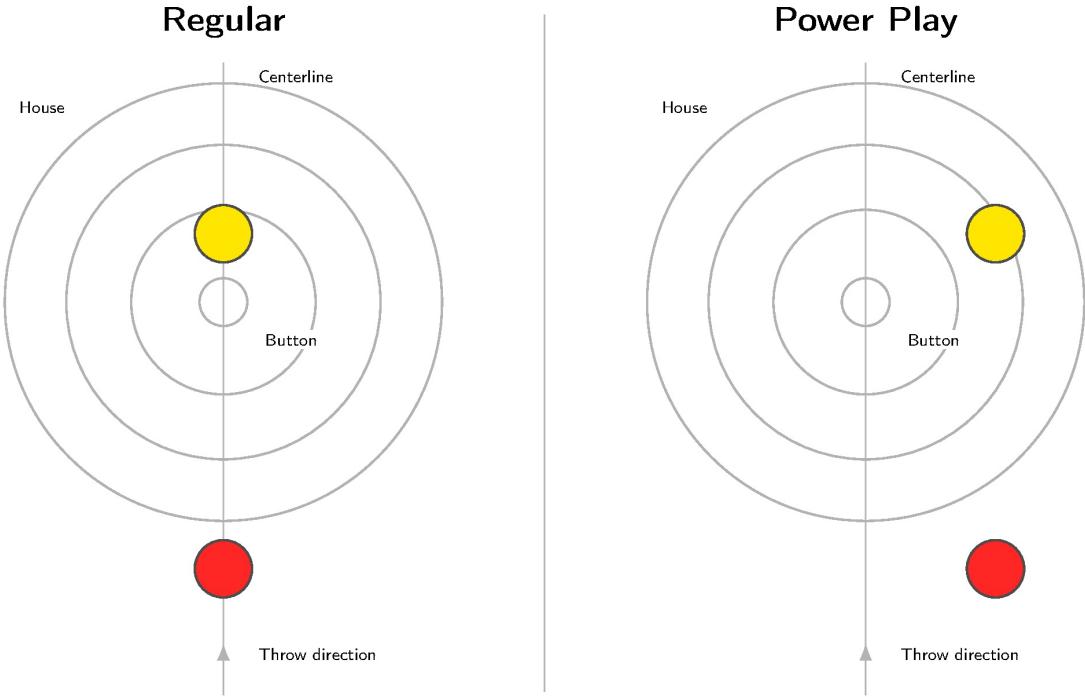


Figure 2: Prepositioned stone setup for regular ends (left) and Power Plays (right).

The ground rules for using the Power Play are:

- Each team can use Power Play at most once.
- It can be used only in ends 1–8 (not in extra ends).
- Only the team with the hammer can call it.

On average, using the Power Play increases the hammer team’s expected score by about 0.29 points in the end. It’s clearly valuable, but the strategic question remains: *when is the right time to deploy it?*

Our Approach

The Power Play’s once-per-game restriction gives it real *option value* for teams: spending it now may provide some immediate benefit, but it also forfeits the opportunity to use it later in a higher-leverage situation. Furthermore, a point gained in the second or third end may be “fixable” over the remaining ends, while a point gained (or saved) in the seventh or eighth end can be decisive. That’s why we evaluate Power Play timing in terms of *winning the game*, not just the expected points gained in the current end.

Our approach is to compute, at the start of each end, the team’s win probability under two choices—use Power Play now or save it—and take the difference as the “Power Play value” in

that situation. This produces a practical decision rule: use the Power Play when it increases win probability more than the value of keeping it for later.

We implement this in two steps:

- Predict end outcomes.** Using historical curling data from the CSAS 2026 Data Challenge ([Connecticut Sports Analytics Symposium \(CSAS\), 2026](#); [CSAS Data Challenge, 2026](#)), we use a multiclass XGBoost model to predict the likelihood of each possible end result—from blank ends to multi-point ends—based on what we know at the start of the end: the end number, the current score, who has hammer, whether Power Play is used, and a simple measure of relative team strength.
- Solve “use now vs. save.”** For each start-of-end situation, we estimate each team’s chance of winning under two choices: using the Power Play now or saving it for later. We compute these by combining the predicted odds of each end outcome with the rules for how the game state updates, then work backward from the final end. The difference in win probabilities is the “Power Play value” in that situation:

$$\Delta WP = WP(\text{use now}) - WP(\text{save})$$

Positive values mean “use now,” and negative values mean “save.”

Our Findings

Figure 3 summarizes the optimal timing policy as heatmaps over end number and score differential: ΔWP is the win-probability gain from using Power Play now rather than saving it.

Our optimal policy boils down to three takeaways:

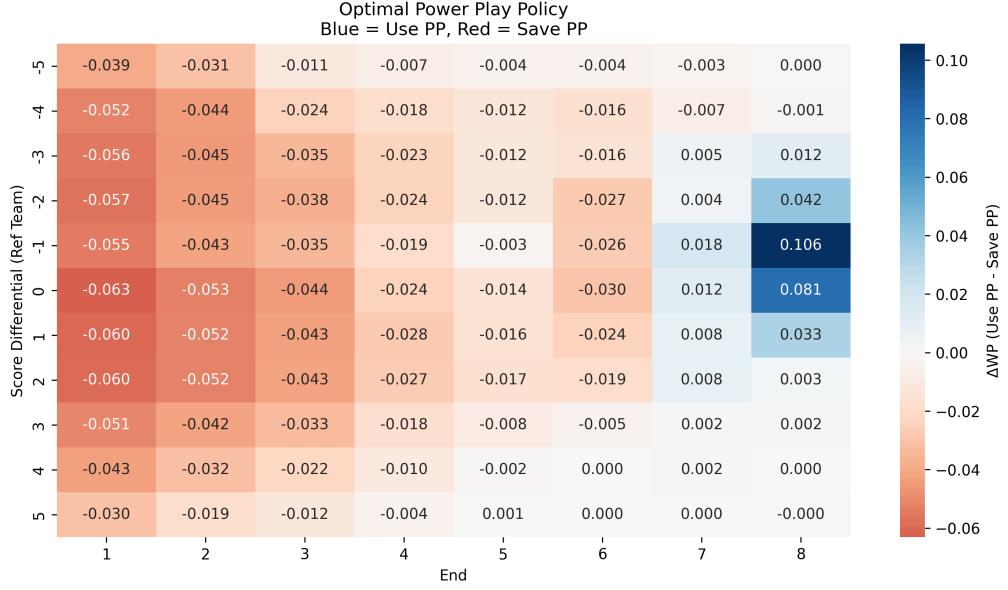
- Power Play is most valuable late in close games.** With the hammer in the final end, using the Power Play meaningfully swings win probability in tight contests: for example:

- End 8, tied: +8–10 percentage points of win probability.
- End 8, down 1: +10–11 percentage points of win probability.
- End 8, down 2: +4–5 percentage points of win probability.

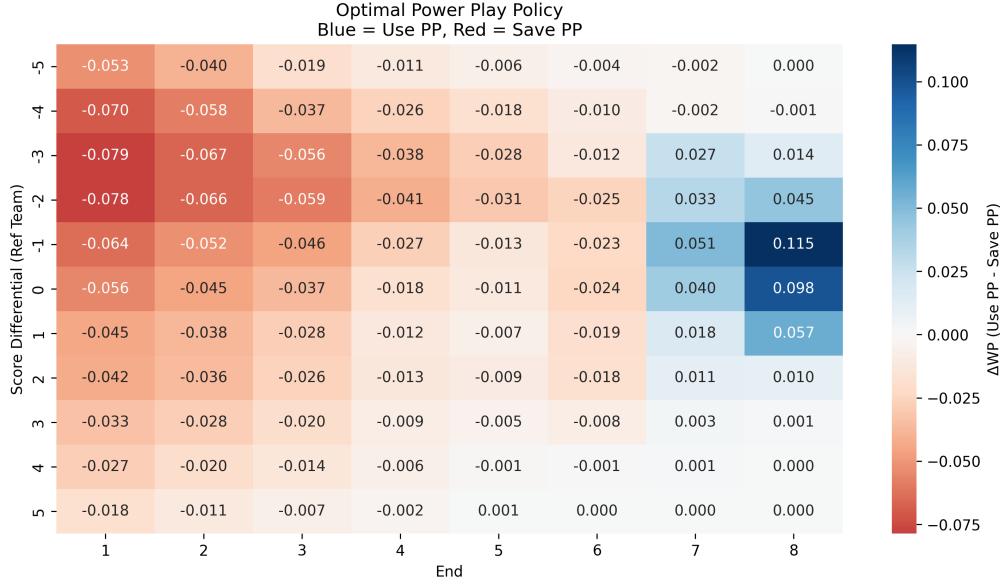
Why? Late ends are high leverage, and they leave the opponent little or no time to recover from falling behind. Using a Power Play also increases the chance of a *multi-point* end with the hammer, which is exactly what teams need to come from behind or score the winning point.

- Early in the game, option value dominates.** A Power Play tends to add expected points in any end, but spending it early can reduce win probability because it sacrifices the ability to deploy it later when a one-end swing is more likely to decide the match. In other words, the key tradeoff is not “points now” versus “points later,” but “win probability now” versus the value of keeping the option.

3. Power Play is worth more when the opponent has already used theirs. If the opponent has spent their Power Play, yours becomes the only remaining lever to open the sheet. That asymmetry increases the value of saving it for late, close situations: the opponent can no longer respond symmetrically.



(a) Opponent PP available



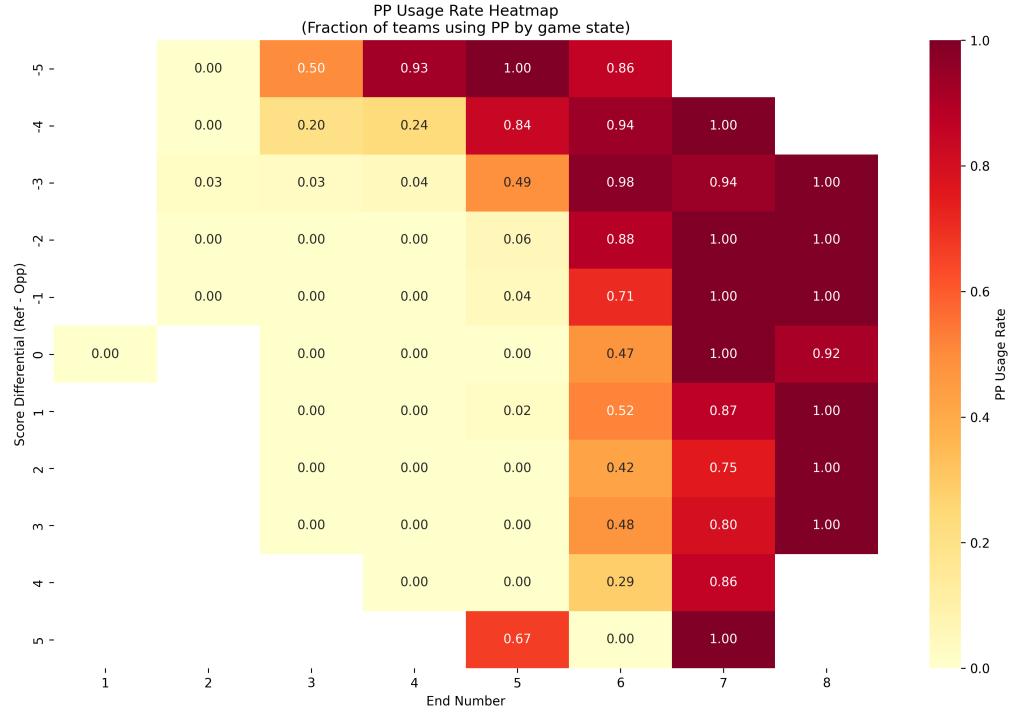
(b) Opponent PP used

Figure 3: Optimal PP policy heatmaps: $\Delta WP = WP(\text{use now}) - WP(\text{save})$.

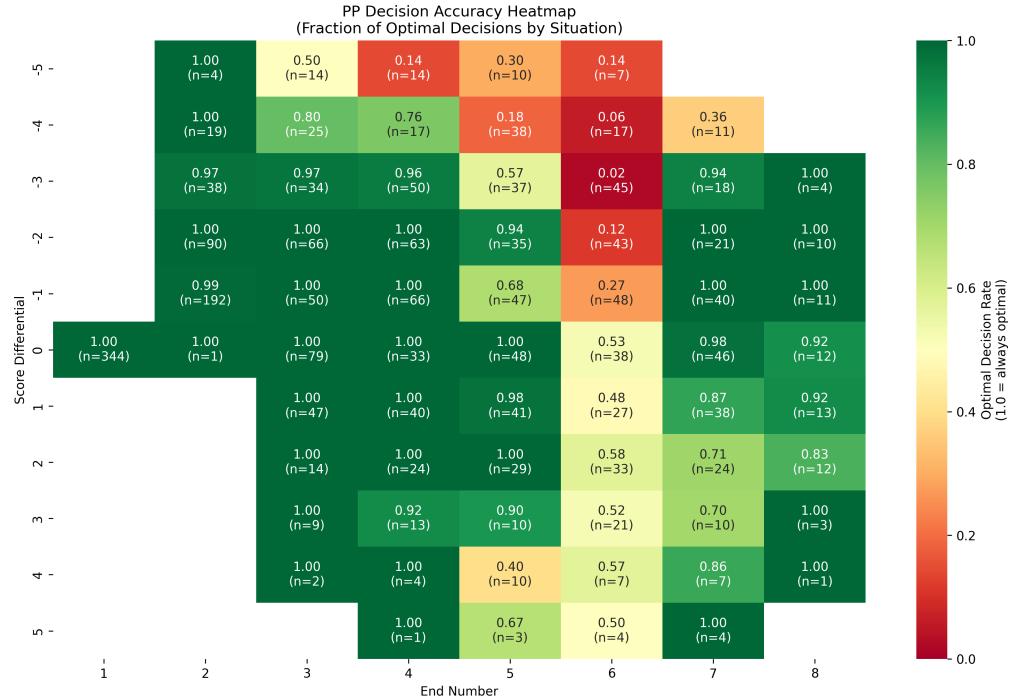
What Teams Actually Do

Figure 4a shows *when* teams deploy the Power Play across ends and score situations. Figure 4b shows how often those choices align with the optimal recommendation (accuracy by situation). Overall, teams rarely call a Power Play in the earliest ends (aside from a few lopsided cases), with most usage concentrated in ends 6–8. Consistent with that pattern, decision accuracy is relatively high early in the game. The main departures from the optimal policy occur in the mid-game: in *ends 5–6*, many teams deploy their Power Play even when the model recommends saving it for later. In those states, using it “too early” carries a measurable cost in win probability on average (roughly 2–3 percentage points in the most pronounced cells). The takeaway is simple: although a Power Play in end 5 or 6 can increase expected points in that end, it often comes at the expense of preserving the Power Play for the highest-leverage moments—late, close ends that can swing the final result.

Figure 5 summarizes teams’ timing performance: how often their Power Play decisions match the optimal recommendation, and the net win probability they gained or lost from those decisions (relative to the optimal policy). In our data, Canada aligns most closely with the optimal policy, while Finland deviates the most. That said, accuracy and impact are not identical: Austria loses the least win probability from its timing choices, whereas Norway loses the most.

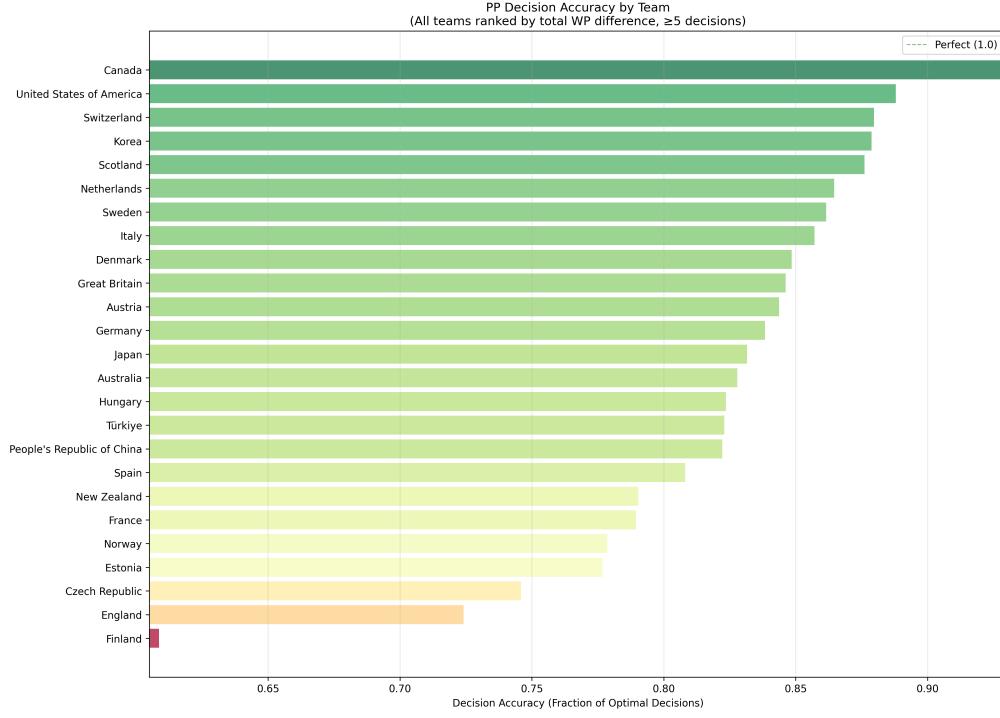


(a) Power Play usage rate by end and score differential among decision-eligible states (hammer + PP available). Cells with no observations are masked.

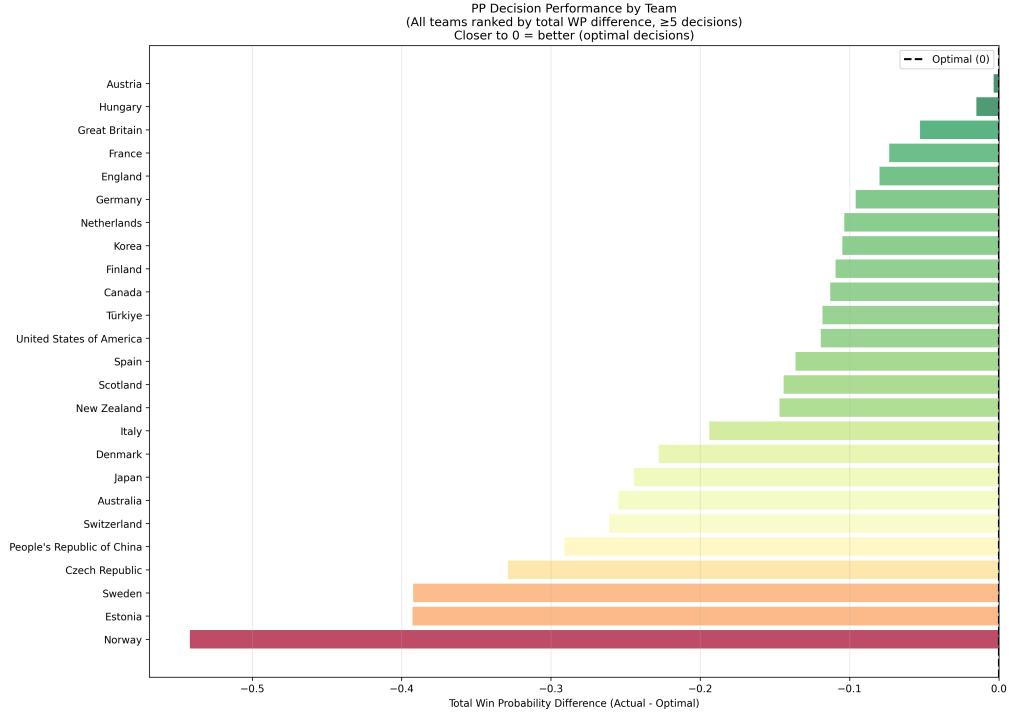


(b) Decision accuracy: fraction of situations where the observed choice matches the optimal policy (labels show sample size).

Figure 4: Observed Power Play timing and how often it matches the optimal policy.



(a) Decision accuracy by team (≥ 5 decisions).



(b) Total WP difference by team (≥ 5 decisions).

Figure 5: Team-level Power Play decision-making: accuracy and net win-probability impact.

Discussion

Practical guidance. The optimal Power Play policy can be summarized with a simple heuristic: in **ends 1–6**, teams should generally **save** their Power Play. Even in a close game, there is still enough time to absorb small swings, and the option value of holding until ends 7–8 is massive. In **ends 7–8**, the recommendation flips: teams should **use** the Power Play almost always. Late ends are high leverage, so deploying it then can create large win-probability swings with little time for the opponent to respond, and the option value is essentially zero. For matchup-specific guidance, we provide an interactive app and all reproducible code on [GitHub](#).

Caveat. Power Play usage is not randomized: teams may choose when to deploy it based on outside factors (ice conditions, matchups, or tactical preferences) that we cannot fully observe in data. With that in mind, teams should treat our policy as a data-informed benchmark: useful for guiding decisions, but best combined with situational judgment. Ultimately, the right framing remains not “does this help right now?” but rather “is this the moment when spending the Power Play is most likely to swing the match outcome in our favor?”.

References

- Connecticut Sports Analytics Symposium (CSAS) (2026). Csas 2026 data challenge: Mixed doubles curling power play. <https://statds.org/events/csas2026/challenge.html>. Accessed: 2025-12-15.
- CSAS Data Challenge (2026). Csas 2026 data challenge data repository. <https://github.com/CSAS-Data-Challenge/2026>. Accessed: 2025-12-15.