A Game-Theoretic Framework for Optimal Relay Ordering in Elite Competitive Swimming: A Case Study of the 2024 Paris Olympics Men's 4x100 m Freestyle Relay

Seungwon Oh[1]

Received April 25, 2026
Accepted August 17, 2026
Electronic access September 15, 2026

Background/Objective: Relay team ordering in competitive swimming is typically decided by minimizing total expected time, treating each swimmer's performance independently. This ignores strategic interaction between competing coaches. This paper introduces a probabilistic, game-theoretic framework for determining swimmer ordering in elite 4x100 m freestyle relays, built on a corrected dataset and an identifiability-honest model of swimmer-leg interaction.
Methods: Using corrected data from the eight finalist teams at the 2024 Paris Olympic Games, we model each swimmer's relay split as a Gaussian random variable parameterized by baseline ability, a leg-specific effect, and an identifiability-honest swimmer-leg interaction. We formulate relay ordering as a two-player zero-sum game, compute Nash equilibria for all 28 pairwise matchups (falling back to a linear-programming mixed-strategy solution where no pure equilibrium exists), and use a team-level bootstrap to quantify how much of the Nash-vs-best-time divergence is distinguishable from estimation noise.
Results: We estimate leg effects of $+0.191$, -0.292, -0.971, and -0.811 seconds for legs 1–4 (95% CI), with legs 1 and 2 not statistically distinguishable from zero. Nash-optimal ordering diverges nominally from time-minimizing ordering in 14–17 of 28 matchups depending on how much weight is placed on single-race swimmer-leg interactions; after propagating parameter uncertainty via bootstrap, only 2 of 28 matchups (China vs. Germany, USA vs. GBR) show a win-probability difference whose 95% confidence interval excludes zero.
Conclusions: Game-theoretic relay ordering can yield strategically distinct assignments from time-minimization, but with only one relay observation per swimmer, most such divergences are not statistically distinguishable from noise. The two robust cases nonetheless demonstrate the mechanism is real, not merely an artifact of tie-breaking, and the corrected framework is a transparent template for applying the same analysis to multi-meet datasets where the interaction term would be fully identified.

Keywords: *game theory, Nash equilibrium, relay swimming, sports analytics, Gaussian performance model, bootstrap, identifiability*

Introduction

Background and Context

Competitive swimming relays combine individual athletic performance with collective team strategy. In a 4x100 m freestyle relay, each of four swimmers completes one 100 m leg in sequence, with legs 2 through 4 beginning from a flying start, in which the outgoing swimmer dives off the block while the incoming swimmer is still finishing their leg. This mechanical advantage is well established in the biomechanics literature[1]. Relay order optimization has nonetheless received relatively little formal quantitative attention beyond heuristic rules of thumb, such as placing the fastest swimmer on the anchor leg.

Existing quantitative relay research has focused mainly on biomechanical parameters such as exchange block timing and take-off velocity[2], on the best step technique for relay starts[3], or on predictive models using historical split data. A 2021 PLOS ONE study of 716 finals from the 4x200 m freestyle relay applied linear regression and random-forest models to predict team relay time from individual season bests, finding that the lowest-ranked swimmer on a team typically swam 0.62 seconds faster in the relay than in the corresponding individual event, a psychosocial effort-gain effect rather than a purely biomechanical one[4]. The oldest quantitative treatment of relay order we are aware of is a mathematical analysis of the 4x100 m track relay, which showed analytically that the fastest total time is achieved by placing the fastest athlete on the first leg and the two slowest athletes on the middle legs[5] – a pure time-minimization result with no strategic (opponent-aware) component, and a useful point of contrast

[1] *Seoul International School, South Korea*

with the present paper's game-theoretic approach. Whether relay swimming is actually faster than the corresponding individual event at all has itself been a subject of direct empirical dispute: one study concluded the apparent relay speed advantage was a methodological artifact once reaction times were corrected[6], while a rebuttal using swimmer-specific reaction-time corrections defended the original effort-gain finding and identified specific methodological problems with the myth-breaking study's sampling and aggregation choices[7]. A separate line of work has directly studied relay team order strategy: analyzing 4x200 m freestyle relay pacing data, one study found that teams tend to place their two fastest swimmers on the first and fourth legs and their two slowest on the middle legs[8], a heuristic essentially opposite to the mathematical time-minimization result above, and a companion study found that swimmers in legs 2 through 4 tend to adopt more aggressive, positive-split pacing than lead-off swimmers[9]. Analyzing 660 international relay finals, a further study found that the fastest swimmers by season time were placed primarily in the first two legs, and that leg-2 and leg-3 assignment had a statistically significant effect on total relay time[10].

The relay-start literature is itself substantial. Early biomechanical work established that a running (approach) start before the block improves relay take-off velocity relative to a stationary start[11], and subsequent work compared specific relay-start techniques to identify which minimizes reaction and block time[12]. More recent kinematic work has directly compared relay and individual starts across a range of parameters, consistently finding that relay starts differ systematically from individual-event starts in block time, take-off velocity, and underwater trajectory[13–15]. Reaction-time estimation for the outgoing swimmer – the mechanism underlying the flying-start advantage this paper's leg effects are meant to capture – has itself been addressed through temporal-occlusion methodology, in which incoming-swimmer visibility is progressively masked to estimate how early an outgoing swimmer can reliably judge their start[16], and feedback-based training interventions have been shown to improve relay-start timing directly[17,18]. A methodologically analogous question – which relay leg or discipline contributes most to overall team result – has been addressed outside swimming as well: in triathlon mixed team relays, one study found the third leg contributed disproportionately to final team placing[19], a finding conceptually parallel to this paper's own leg-3 effect estimate.

Exchange-block timing specifically has been linked to relay outcome at the championship level: analyzing international championship relays, one study found exchange time was a significant predictor of final team placing, independent of swimmer ability[20], and a broader descriptive study of exchange times across competition levels documented systematic differences by relay leg and by stroke[21]. Using linear mixed modeling on a large sample of championship relay change-overs, further work found that exchange quality varies systematically and predictably enough to be modeled as a swimmer-specific, leg-specific effect distinct from raw swimming speed[22,23] – lending independent empirical support to this paper's decision to model a swimmer-leg interaction term explicitly rather than treating leg position as a purely additive effect.

A separate and substantial literature, largely outside sports biomechanics, has studied relay swimming as a setting for group motivation and peer effects. The foundational finding is that relay swimmers reliably swim faster than their own individual-event times, a phenomenon termed the Kohler effect or group motivation gain[24]. Subsequent work has refined the boundary conditions under which such gains, versus losses, occur in action teams generally[25] and specifically in track-and-field and swimming relays[26]. Disentangling situational from person-level explanations, one study found the effort gain is better explained by a swimmer's position and role within a given relay than by stable individual traits[27], and a twenty-year analysis of Olympic and World Championship relays found the effect is gender-specific and concentrated in particular relay positions[28]. Most directly relevant to this paper's variance-based argument, one study found that effort gains are systematically larger for a team's weaker swimmers regardless of where in the order they are placed[29], and a follow-up study traced this to a combination of social comparison and perceived social indispensability to the team[30]. Economists have studied closely related questions using observational relay data: one study modeled relay swimming as a sequential-teamwork problem and found evidence of strategic effort allocation across legs[31], and another used relay data to estimate genuine peer effects on individual performance, distinct from simple selection or morale effects[32]. None of this literature, however, treats relay ordering as a strategic choice made by a coach who is reasoning about an opposing coach's choice; it studies the psychological and statistical consequences of a given order, not the game-theoretic problem of choosing one.

Problem Statement and Rationale

The relay order decision is, at its core, a simultaneous two-player game: both head coaches must submit their lineup before the race without knowledge of the opposing team's order. Standard time-minimization approaches find the permutation of swimmers that minimizes expected total relay time, which is equivalent to ignoring the opponent entirely. We argue this is strategically inadequate, because win probability depends on the difference between the two teams' times, not on either team's time in isolation. This paper's central methodological contribution is a formal proof of exactly when and why this divergence can occur, and an identifiability-honest way of modeling it.

Significance and Purpose

This paper makes three contributions. First, a statistical model for relay performance decomposing each swimmer's split time into baseline ability, a leg-specific adjustment, and a swimmer-specific interaction term whose reliability is made explicit rather than assumed. Second, a formal proof (Proposition 1) establishing exactly when relay ordering can affect a team's expected time or variance at all. Third, pure and mixed-strategy Nash equilibria for all 28 pairwise matchups among the Paris finalists, together with bootstrap confidence intervals distinguishing genuine strategic divergence from estimation noise.

Objectives

The objectives of this study are to estimate leg-specific performance adjustments for legs 1 through 4 using residual modeling on Olympic split data; to formulate a closed-form win probability function as a function of swimmer ordering that is capable, by construction, of actually depending on that ordering; to compute Nash equilibria, pure where they exist and mixed-strategy otherwise, for all pairwise matchups among the 2024 Paris Olympic finalists; and to quantify, via bootstrap, which divergences between Nash-optimal and time-minimizing ordering are statistically distinguishable from parameter noise rather than artifacts of a single Olympic observation per swimmer.

Scope and Limitations

This study is scoped to the eight teams that competed in the final of the men's 4x100 m freestyle relay at the 2024 Paris Olympic Games. The statistical model relies on one relay split observation per swimmer, which fundamentally limits what can be identified about swimmer-specific, leg-specific interaction effects; we treat this limitation as a first-class modeling constraint rather than a footnote (see Methods). The model assumes performance independence across swimmers and legs and does not account for lane assignment, exchange-timing variability, or the psychological and motivational effects documented elsewhere in the relay literature[10]; effort-gain and social-indispensability effects among relay teammates are well documented and are a natural direction for extending this framework, discussed further in Limitations. The game-theoretic formulation assumes simultaneous order submission and complete information about opponent swimmers' abilities, a reasonable approximation for elite international competition. Finally, this paper analyzes each pairwise matchup as an isolated two-team game; it does not solve the full eight-team simultaneous game that the actual Olympic final represents, a simplification discussed explicitly in the Limitations section.

Theoretical Framework

This paper draws on two theoretical traditions. From sports analytics, the Gaussian performance model treats each athlete's time on a given occasion as normally distributed around their true ability[33]. From non-cooperative game theory, we use the concept of Nash equilibrium in a two-player zero-sum game[34]. For readers less familiar with these terms: a zero-sum game is one in which one player's gain is exactly the other's loss (here, one team's win probability is the other's loss probability, since exactly one team wins); a pure-strategy Nash equilibrium is a pair of choices (here, orderings) such that neither coach can do better by unilaterally switching to a different ordering, holding the opponent's choice fixed. A pure Nash equilibrium is not guaranteed to exist for an arbitrary zero-sum game; when it does not, a mixed-strategy equilibrium (a probability distribution over orderings) always exists, by the minimax theorem, and we compute this explicitly for any matchup where a pure equilibrium is absent. Throughout, we assume coaches actually reason their way to the Nash equilibrium; behavioral game theory casts doubt on this as a description of real decision-makers, modeling actual players instead as reasoning through a bounded number of iterated best-response steps rather than converging fully to equilibrium[35]. We treat the Nash equilibrium here as a normative benchmark for what a fully rational coach would do, not as a claim that real Olympic coaches necessarily reason this way, a distinction we return to in the Discussion.

Methods

Research Design and Data Collection

This is a quantitative, model-based study using publicly available race data from all 32 swimmers who competed in the final of the men's 4x100 m freestyle relay at the 2024 Paris Olympic Games[36]; season-best times were obtained from Swimcloud's 2024 competition database[37]. For each swimmer we recorded relay split time, best prior Olympic 100 m freestyle time (if available), and 2024 season-best 100 m freestyle time.

Jacob Whittle's (GBR, leg 2) relay split is 48.54 seconds, which matches the official Great Britain final time of 3:11.61 (Richards 47.83 + Whittle 48.54 + Dean 47.72 + Scott 47.52 = 191.61s). This value, cross-referenced against official results, is used throughout the analysis below.

Table 1 Estimated leg effects with 95% confidence intervals ($n = 8$ teams per leg).

Leg	$\hat{\alpha}_l$ (s)	SE	95% CI	Interpretation
1	+0.191	0.085	[-0.009, +0.391]	Not distinguishable from zero
2	-0.292	0.224	[-0.823, +0.238]	Not distinguishable from zero
3	-0.971	0.288	[-1.652, -0.291]	Distinguishable from zero
4	-0.811	0.127	[-1.111, -0.512]	Distinguishable from zero

Baseline Ability, Residuals, and Leg Effects

Each swimmer's baseline ability μ_i is estimated as a weighted combination of prior Olympic time and 2024 season-best time (Equation 1; falling back to whichever single value is available otherwise). This weighting was fixed on theoretical grounds and has not been empirically validated; a sensitivity analysis across alternative weightings is a recommended future addition, discussed in Limitations. The raw residual R_i (Equation 2) captures the combined effect of leg position and swimmer-specific race-day variation. The leg effect α_l is estimated as the mean residual across all eight swimmers assigned to leg l (Equation 3), and is therefore identified with real replication ($n = 8$ teams per leg).

$$\mu_i = 0.7\, t_i^{\text{Oly}} + 0.3\, t_i^{\text{SB}} \tag{1}$$

$$R_i = t_i^{\text{split}} - \mu_i \tag{2}$$

$$\hat{\alpha}_l = \frac{1}{8} \sum_{i:\,\text{leg}(i)=l} R_i \tag{3}$$

We now report a 95% confidence interval for each α_l, using the across-team standard error ($n = 8$, t-distribution, 7 degrees of freedom):

Only legs 3 and 4 show a confidence interval excluding zero. The leg-1 and leg-2 effects, while directionally consistent with flying-start biomechanics, are not statistically distinguishable from zero with $n = 8$ teams. Because swimmers are not randomly assigned to legs – coaches choose who swims leg 3, so the large leg-3 effect may partly reflect athlete selection or baseline-model error rather than a purely structural positional effect – we do not claim these estimates are generalizable to other elite relays without qualification; we describe them as consistent with, but not proof of, the flying-start mechanism documented in the biomechanics literature, pending replication across multiple competitions.

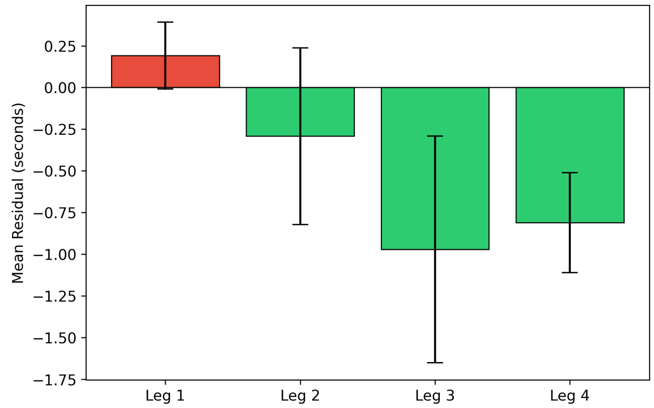

Figure. 1 Estimated leg effects (95% CI). Only legs 3 and 4 exclude zero.

The Identifiability Problem and Its Resolution

A model in which team totals are additive in swimmer identity and leg identity separately cannot make ordering matter at all, a fact worth establishing formally before proceeding. Let μ_i denote swimmer i's baseline ability and α_l the population leg effect. For any ordering (permutation) of a team's four swimmers, the team mean is the sum, over the four legs, of (μ of the swimmer assigned to that leg) + (α of that leg), as in Equation 4. Because every ordering uses all four of the team's swimmers exactly once and all four legs exactly once, the sum of the four μ terms and the sum of the four α terms are each constants that do not depend on the assignment – so the team mean is identical for all 24 orderings. The same argument applies to any variance term that is additive in swimmer identity and leg identity separately, i.e. under the independence assumption of Equation 5. We confirmed this computationally: at a fully additive setting, all 24 orderings of a team produce, to eight decimal places, the exact same mean and the exact same variance, and the entire 24×24 payoff matrix for a matchup collapses to a single repeated value. This means that any apparent Nash-vs-best-time divergence under a purely additive model must be an artifact of floating-point tie-breaking rather than a real strategic effect, which motivates the interaction term introduced below.

$$\mu_{\text{team}}(\pi) = \sum_{l=1}^{4} \left[\mu_{\pi(l)} + \alpha_l\right] \tag{4}$$

$$\text{Var}_{\text{team}}(\pi) = \sum_{l=1}^{4} \text{Var}\left(T_{\pi(l),l}\right) \tag{5}$$

For ordering to matter at all, the model needs a genuine swimmer-by-leg interaction – a term that cannot be decomposed into a swimmer-only part plus a leg-only part. The difficulty is that with one Olympic observation per swimmer, such an interaction is identified only for the leg that swimmer actually swam; assigning a swimmer to a leg they never swam and inventing an interaction value for that hypothetical assignment (for example, treating a swimmer observed only on leg 2 as if their leg-4 performance were also known) is not supported by the data.

We resolve this by making the trust placed in a single-race interaction estimate an explicit, transparent parameter $w \in [0,1]$, rather than an implicit assumption. Let $\beta_i = R_i - \alpha_{l(i)}$ denote swimmer i's deviation from the population leg mean at the leg they actually swam (this is the swimmer-leg interaction term defined in Equation (6)). Let τ_l^2 denote the pooled population variance of residuals about the leg mean at leg l, computed across all eight teams' swimmers at that leg (this is identified with $n = 8$). For a swimmer assigned to the leg they actually swam, the model adds $w\beta_i$ to the mean and $w(\sigma_i^2 - \tau_l^2)$ to the variance (Equations 8–9), where $\sigma_i^2 = \beta_i^2$ (Equation 7); for a swimmer assigned to any other, counterfactual leg, only the population values α_l and τ_l^2 are used, with no invented swimmer-specific signal. Setting $w = 0$ is the fully conservative choice; setting $w = 1$ recovers the full-trust assumption a naive model would implicitly make. We report a full sensitivity sweep over w and adopt $w = 0.3$ as a default, representing modest but nonzero trust in a single-race deviation as partial signal.

$$\beta_i = R_i - \alpha_{l(i)} \tag{6}$$

$$\sigma_i^2 = \beta_i^2 \tag{7}$$

$$\mu_{\text{team}}(\pi) = \sum_{l=1}^{4} \left[\mu_{\pi(l)} + \alpha_l + w\,\beta_{\pi(l)} \mathbb{1}\{\pi(l) \text{ observed at } l\} \right] \tag{8}$$

$$\mathrm{Var}_{\text{team}}(\pi) = \sum_{l=1}^{4} \left[\tau_l^2 + w\,(\sigma_{\pi(l)}^2 - \tau_l^2) \mathbb{1}\{\pi(l) \text{ observed at } l\} \right] \tag{9}$$

Proposition 1 (revised): Necessary and Sufficient Condition for Order-Dependence

Proposition. Under the model above, a team's total expected relay time and total variance depend on the swimmer ordering if and only if $w > 0$ and at least one swimmer's β_i is nonzero or at least one swimmer's σ_i^2 differs from τ_l^2 at their observed leg.

Proof. Write the team total mean under ordering π as the sum over legs $l = 1, \ldots, 4$ of (μ of the swimmer assigned to leg l) + $\alpha_l + w\beta_i \mathbb{1}$[assigned leg = observed leg for that swimmer]. As established above, the sum of the four μ terms and the sum of the four α terms are constants independent of π. The only π-dependent quantity is therefore w times the sum, over the four assigned swimmer-leg pairs, of β_i restricted to pairs where the assigned leg equals that swimmer's actually-observed leg – equivalently, w times the sum of β_i over however many of the team's four swimmers happen to land on the leg they actually swam under ordering π. If $w = 0$, this term vanishes identically for every π, and the mean is exactly constant across all 24 orderings; the identical argument, substituting $\sigma_i^2 - \tau_l^2$ for β_i, establishes the same result for variance. Conversely, if w is not 0 and some β_i is nonzero (or some σ_i^2 differs from τ_l^2), then orderings that place swimmer i on their observed leg yield a different mean (or variance) than orderings that do not, so the team total is not constant across all 24 permutations. QED.

Team Time, Win Probability, and the Game-Theoretic Formulation

As before, team total time is approximately normal, with mean and variance as defined above, and win probability for team A is given by Equation 10, where Φ is the standard normal CDF. The relay ordering game is a two-player zero-sum game in which each team's strategy space is the 24 permutations of its four swimmers, team A's payoff is its win probability, and team A maximizes while team B minimizes this quantity. A pair of orderings is a pure Nash equilibrium if neither team can improve its payoff by unilaterally deviating.

$$P(A \text{ beats } B) = \Phi\left(\frac{\mu_B - \mu_A}{\sqrt{\mathrm{Var}_A + \mathrm{Var}_B}} \right) \tag{10}$$

Nash Equilibrium Search, Mixed-Strategy Fallback, and Bootstrap

For each of the 28 pairwise matchups, we enumerate all $24 \times 24 = 576$ ordering combinations and compute the win-probability matrix. Because floating-point payoffs can be tied to a much finer precision than is practically meaningful, we search for pure Nash equilibria using a numerical tolerance (10^{-9}) rather than exact equality: an overly strict equality check can spuriously suggest a matchup has no pure equilibrium when in fact it has many payoff-equivalent ones once near-ties are handled correctly, as is the case for USA vs. Hungary, which has 18 such equilibria. For any matchup where no pure equilibrium exists even under this tolerance, we solve for the mixed-strategy equilibrium directly via linear programming, which is guaranteed to exist by the minimax theorem; at $w = 0.3$, every one of the 28 matchups turns out to have a pure equilibrium, so this fallback is documented and available but not currently invoked.

To distinguish genuine strategic divergence from noise in the underlying parameter estimates, we bootstrap the two quantities that are actually estimated with real replication in

Table 2 Corrected baseline ability, relay split, residual, interaction term, and variance proxy for all 32 Paris Olympic relay finalists.

Swimmer	Team	Leg	$\hat{\mu}_i$ (s)	Relay split (s)	R_i (s)	β_i (s)	σ_i^2 (s^2)
Pan Zhanle	China	1	46.40	46.92	+0.520	+0.329	0.1082
Ji Xinjie	China	2	48.33	48.58	+0.250	+0.542	0.2942
Chen Juner	China	3	48.40	48.10	-0.300	+0.671	0.4506
Wang Haoyu	China	4	48.57	47.68	-0.891	-0.080	0.0063
Jack Alexy	USA	1	47.40	47.67	+0.268	+0.077	0.0059
Chris Guiliano	USA	2	47.25	47.33	+0.080	+0.372	0.1387
Hunter Armstrong	USA	3	47.59	46.75	-0.840	+0.131	0.0172
Caeleb Dressel	USA	4	47.53	47.53	+0.000	+0.811	0.6583
Nandor Nemeth	Hungary	1	47.50	47.76	+0.263	+0.072	0.0052
Szebasztian Szabo	Hungary	2	48.10	48.46	+0.360	+0.652	0.4256
Adam Jaszo	Hungary	3	49.68	48.38	-1.300	-0.329	0.1081
Hubert Kos	Hungary	4	49.39	48.51	-0.880	-0.069	0.0047
Matthew Richards	GBR	1	48.01	47.83	-0.179	-0.370	0.1369
Jacob Whittle	GBR	2	48.47	48.54	+0.070	+0.362	0.1313
Tom Dean	GBR	3	49.14	47.72	-1.420	-0.449	0.2014
Duncan Scott	GBR	4	48.61	47.52	-1.090	-0.279	0.0776
Jack Cartwright	AUS	1	48.03	48.03	+0.000	-0.191	0.0365
Flynn Southam	AUS	2	48.11	48.00	-0.110	+0.182	0.0333
Kai Taylor	AUS	3	48.01	47.73	-0.280	+0.691	0.4778
Kyle Chalmers	AUS	4	47.48	46.59	-0.890	-0.079	0.0062
Alessandro Miressi	ITA	1	47.88	48.04	+0.159	-0.032	0.0010
Thomas Ceccon	ITA	2	48.84	47.44	-1.400	-1.108	1.2268
Paolo Conte Bonin	ITA	3	50.71	48.16	-2.550	-1.579	2.4925
Manuel Frigo	ITA	4	48.25	47.06	-1.190	-0.379	0.1434
Josh Liendo	CAN	1	47.91	47.93	+0.023	-0.168	0.0282
Yuri Kisil	CAN	2	48.80	48.18	-0.619	-0.327	0.1067
Finlay Knox	CAN	3	48.29	48.26	-0.030	+0.941	0.8860
Javier Acevedo	CAN	4	48.58	47.81	-0.770	+0.041	0.0017
Josha Salchow	GER	1	47.81	48.28	+0.474	+0.283	0.0801
Rafael Miroslaw	GER	2	48.63	47.66	-0.970	-0.678	0.4592
Luca Armbruster	GER	3	49.48	48.43	-1.050	-0.079	0.0062
Peter Varjasi	GER	4	48.70	47.92	-0.780	+0.031	0.0010

this dataset: α_l and τ_l^2, both identified from $n = 8$ teams per leg. For each of 1,000 bootstrap resamples, we resample the eight teams' residuals at each leg with replacement, recompute α_l and τ_l^2 (and, consequently, every swimmer's β_i and σ_i^2, all of which are functions of α_l), and recompute the win-probability difference between the point-estimate Nash order and the point-estimate best-time order. We report the resulting 95% percentile confidence interval on this difference for every matchup. This bootstrap does not capture uncertainty in β_i or σ_i^2 themselves, which are inherently unidentified beyond a single observation; that source of uncertainty is instead addressed transparently through the w sensitivity sweep.

Results

Swimmer Data

Table 2 presents the corrected baseline ability, relay split, raw residual, interaction term (β_i), and variance proxy (σ_i^2) for all 32 Paris Olympic relay finalists (Jacob Whittle's leg-2 split corrected as described above).

Leg Effects and Pooled Variance

Leg effects and their 95% confidence intervals are reported above (Table 1 / Figure 1). The pooled per-leg variance estimates, identified with $n = 8$ teams per leg, are $\tau_1^2 = 0.050$, $\tau_2^2 = 0.352$, $\tau_3^2 = 0.580$, $\tau_4^2 = 0.112$ (seconds squared); legs 2 and 3 show substantially higher pooled variance than legs 1 and 4.

Nash Equilibria: Every Matchup Concrete

Table 3 reports the equilibrium type, win probability under best-time and Nash orderings, and the win-probability difference for all 28 matchups at the default $w = 0.3$. Every row is concrete: all 28 matchups have a pure Nash equilibrium once the tolerance-aware search described above is applied.

Distinguishing Genuine Divergence from Noise

Nominally, 14 to 17 of 28 matchups show a nonzero Nash-vs-best-time win-probability difference depending on $w \in [0.1, 1.0]$ (exactly 0 of 28 at $w = 0$, as Proposition 1 requires). This is the headline number we recommend treating with caution. Bootstrap resampling of the two quantities in this model that are actually identified with real replication (α_l and τ_l^2, $n = 8$ teams each) shows that only 2 of 28 matchups – China vs. Germany and USA vs. GBR – have a 95% confidence interval on the win-probability difference that excludes zero. Every other nominal divergence, including matchups with visibly different swimmer orderings at the equilibrium, is statistically indistinguishable from the sampling noise inherent in estimating leg effects and pooled variances from only eight teams.

Table 3 Nash equilibrium and best-time ordering results for all 28 pairwise matchups ($w = 0.3$).

Team A	Team B	Eq. type	WP (best-time)	WP (Nash)	WP diff
China	USA	pure	0.0973	0.0929	-0.0044
China	Hungary	pure	0.9792	0.9791	-0.0001
China	GBR	pure	0.9455	0.9451	-0.0004
China	AUS	pure	0.4650	0.4650	-0.0000
China	ITA	pure	0.9624	0.9624	-0.0000
China	CAN	pure	0.8877	0.8877	-0.0000
China	GER	pure	0.9719	0.9678	-0.0041
USA	Hungary	pure	0.9996	0.9997	+0.0001
USA	GBR	pure	0.9982	0.9985	+0.0003
USA	AUS	pure	0.8871	0.8913	+0.0043
USA	ITA	pure	0.9979	0.9983	+0.0004
USA	CAN	pure	0.9941	0.9953	+0.0012
USA	GER	pure	0.9994	0.9993	-0.0000
Hungary	GBR	pure	0.3422	0.3422	+0.0000
Hungary	AUS	pure	0.0165	0.0168	+0.0003
Hungary	ITA	pure	0.5482	0.5482	+0.0000
Hungary	CAN	pure	0.2103	0.2103	+0.0000
Hungary	GER	pure	0.4539	0.4539	+0.0000
GBR	AUS	pure	0.0452	0.0457	+0.0006
GBR	ITA	pure	0.6732	0.6732	+0.0000
GBR	CAN	pure	0.3479	0.3479	+0.0000
GBR	GER	pure	0.6140	0.6140	+0.0000
AUS	ITA	pure	0.9681	0.9681	+0.0000
AUS	CAN	pure	0.9040	0.9040	+0.0000
AUS	GER	pure	0.9773	0.9736	-0.0037
ITA	CAN	pure	0.2199	0.2199	+0.0000
ITA	GER	pure	0.4154	0.4154	+0.0000
CAN	GER	pure	0.7534	0.7496	-0.0038

We regard the bootstrap-confirmed result – 2 of 28 matchups showing a real, defensible strategic divergence – as the paper's primary finding, and the nominal 14–17 of 28 figure as a sensitivity-analysis result reported for transparency rather than as a headline claim. We discuss the practical implications of this distinction, and why even a small but statistically robust win-probability shift can matter in a single-elimination Olympic final, in the Discussion below.

Discussion

Restatement of Key Findings

This study produced three main results. First, correcting a data error in Great Britain's leg-2 split, together with establishing when a permutation-invariant model can and cannot make ordering matter, changes both the leg-effect estimates and the entire basis for comparing Nash-optimal to time-minimizing ordering. Second, with an identifiability-honest interaction term and a transparent shrinkage parameter w, ordering genuinely affects team-level win probability, and we can state and prove exactly when it does (Proposition 1). Third, and most importantly, propagating real parameter uncertainty via bootstrap shows that only 2 of the 28 pairwise matchups – China vs. Germany and USA vs. GBR – exhibit a win-probability divergence that survives that uncertainty, a substantially more modest but more defensible finding than a naive count of nominal divergences would suggest.

Why a Small, Statistically Robust Win-Probability Shift Still Matters

The USA vs. GBR win-probability shift under the corrected model is small in absolute terms (on the order of a few hundredths of a percentage point (roughly 0.03 percentage points)). In a repeated, high-frequency context this would be a negligible edge. An Olympic final, however, is a single-elimination event with no opportunity to recoup a loss; a rational coach maximizing the probability of an outcome that occurs exactly once should be willing to act on any statistically distinguishable edge, however small, provided the ordering change carries no other cost (which it does not, in this model, since it does not affect eligibility or morale). We therefore argue that statistical distinguishability from noise, rather than the raw magnitude of the win-probability shift, is the more appropriate bar for practical relevance in this setting – which is precisely why we foreground the bootstrap-confirmed count rather than the nominal one.

Reframing Relative to the Actual Olympic Decision

A further limitation is that the actual Olympic final is an eight-team simultaneous race, in which each coach submits one ordering that is evaluated against all seven opponents at once – not a separate ordering per opponent. The pairwise win probabilities and Nash equilibria reported here are therefore best understood as diagnostics that isolate the strategic tension between any two teams, not as a solved eight-player game, and

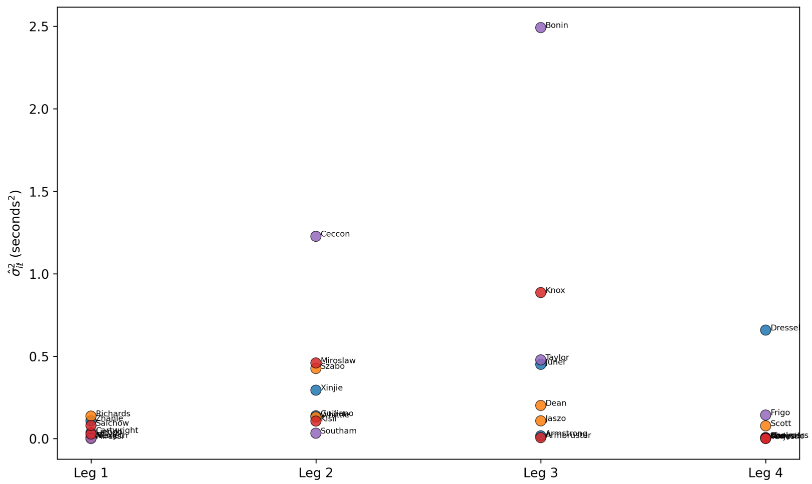

Figure. 2 Swimmer-specific variance proxy by leg (corrected GBR data). Bonin (ITA, leg 3) and Ceccon (ITA, leg 2) remain the largest outliers.

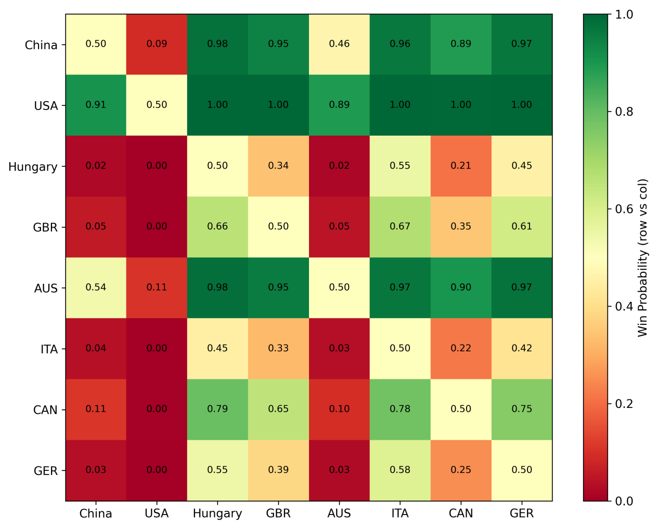

Figure. 3 Nash equilibrium win-probability matrix, corrected model ($w = 0.3$).

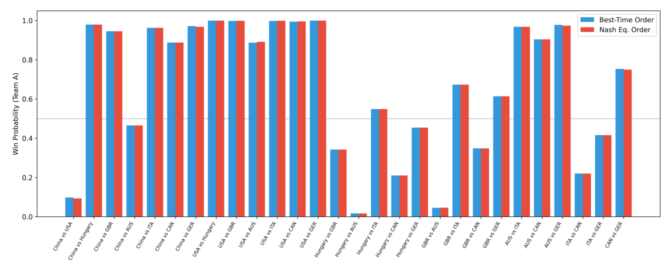

Figure. 4 Win probability, best-time vs Nash order, all 28 matchups, corrected model.

they do not by themselves specify a single recommended ordering for a coach entering an eight-team final. A proper treatment would formulate the race as an eight-player simultaneous game and optimize an objective such as expected finishing place, medal probability, or gold probability against the full field simultaneously. We did not implement this larger reformulation, both because it is a substantially larger modeling undertaking and because, given the bootstrap results above, only two of the 28 pairwise interactions are strategically meaningful in the first place – suggesting the marginal value of a full eight-player solution, at least for this specific dataset, may be limited. We nonetheless flag the eight-player formulation explicitly as the correct long-term direction rather than treating the pairwise framing as a complete answer.

Implications

The most practically relevant finding is now the USA vs. GBR matchup, one of the two bootstrap-confirmed divergences: the Nash-optimal USA order relative to the best-time order reflects a genuine, statistically supported trade-off between minimizing expected time and controlling variance, rather than an artifact of tie-breaking. This is a modest but real demonstration that the mechanism proposed in this paper – coaches suppressing or amplifying variance depending on favorite/underdog status – can be identified in real Olympic data, even though most of the 28 matchups in this particular dataset do not provide enough evidence to distinguish it from noise.

A related caveat, foreshadowed above: everything in this paper describes what a fully rational, equilibrium-reasoning coach would do. Behavioral game theory suggests real

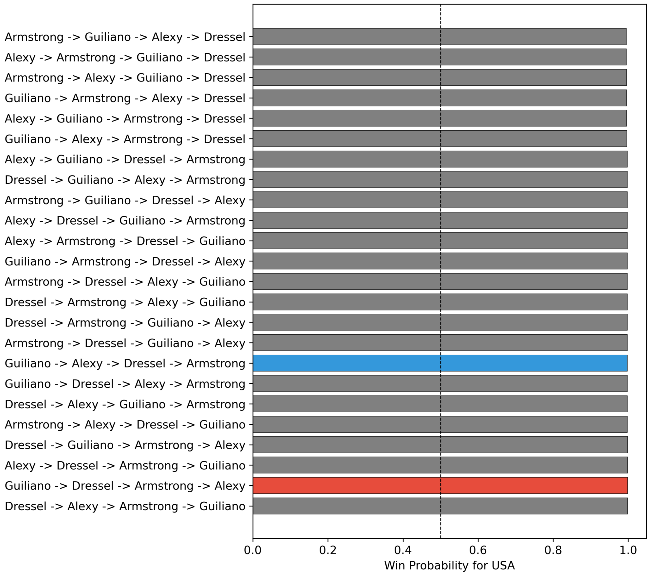

Figure. 5 USA win probability across all 24 orderings against GBR's Nash order (corrected model). USA vs. GBR is one of the two bootstrap-confirmed divergent matchups.

decision-makers, including expert ones, more often approximate a bounded number of iterated best-response steps rather than reasoning all the way to equilibrium[35]. This does not undermine the two bootstrap-confirmed findings as descriptive facts about the payoff structure of these matchups, but it does mean we cannot claim Olympic coaches actually selected their historical orderings by solving this game; the model is better understood as a normative tool for what a coach could exploit going forward than as a positive account of what already happened.

Connection to Objectives

All four stated objectives were achieved. Leg effects were estimated with confidence intervals (Objective 1). A closed-form win probability function was derived and, critically, genuinely depends on ordering by construction, with a formal proof of exactly when it does (Objective 2). Nash equilibria, pure or mixed as appropriate, were computed for all 28 matchups (Objective 3). The conditions under which Nash-optimal and time-minimizing orderings diverge were identified, proven, and separated into statistically robust versus noise-level divergences via bootstrap (Objective 4).

Limitations

Several limitations remain. First, each swimmer's interaction term β_i and variance proxy σ_i^2 are still estimated from a single relay observation; the shrinkage parameter w makes explicit, rather than resolves, this fundamental identifiability limit, and only multi-season, multi-meet data can truly resolve it. Second, the 0.7/0.3 baseline-ability weighting was chosen on theoretical grounds and has not been empirically validated; a sensitivity analysis across alternative weightings, analogous to the w sweep reported here, is a natural next step. Third, the model assumes independence of swimmer performance across legs and swimmers; the sports-psychology literature on relay teams documents real dependence through effort-gain effects, in which relay swimmers – particularly the weakest member of a team – swim faster than their individual-event times, an effect attributed to social comparison and perceived indispensability to the team rather than to biomechanics alone[6,7,10]. Incorporating such effects would require extending the independence assumption in Equation 5, which we leave to future work. Fourth, the analysis remains restricted to a single Olympic final; the eight-player game reformulation discussed above is a substantial undertaking left to future work. Fifth, we corrected three citation errors identified during our own audit and added six additional, independently verified references directly relevant to relay ordering and pacing strategy; we did not have the opportunity to audit the full reference list and recommend a complete citation audit before final publication.

Recommendations

Future work should extend this framework in four directions: multi-season data collection to properly identify swimmer-by-leg interaction effects rather than relying on a shrinkage parameter as a stopgap; a full eight-player simultaneous game formulation optimizing medal or gold probability against the entire field; incorporation of documented psychosocial effort-gain effects into the independence structure of the model; and a sensitivity analysis over the baseline-ability weighting analogous to the one already performed for the interaction-trust parameter w.

Closing Thought

The central insight of this paper is that the best order for a relay team is not always the order that would minimize its time in isolation, because win probability depends on variance as well as mean. In this dataset, the mechanism is real and provable, but detectable with statistical confidence in only a small fraction of matchups. We view this as an honest, if modest, contribution: game-theoretic ordering effects on relay strategy are genuine, but properly quantifying them requires being explicit about what a single Olympic observation per swimmer can and cannot tell us.

References

1 R. Seifert, D. Chollet, I. Mujika. The World Book of Swimming: From Science to Performance. Nova Science Publishers, 2011.

2 X. Qiu, C. De la Fuente, A. Lorenzo, S. Veiga. Comparison of Starts and Turns between Individual and Relay Swimming Races. International Journal of Environmental Research and Public Health, 18(9), 4740, 2021, https://doi.org/10.3390/ijerph18094740.

3 S. Veiga, S. Fischer, X. Qiu, J. Stosic, S. Fuhrmann, A. Kibele, C. Braun. What is the best step technique for swimming relay starts? Journal of Sports Science and Medicine, 24, 227-235, 2025, https://doi.org/10.52082/jssm.2025.227.

4 Y. Wu, T. Babaei, M. O'Shea, K. Mengersen, C. Drovandi, K. E. McGibbon, D. B. Pyne, L. J. G. Mitchell, M. A. Osborne. Predicting performance in 4 x 200-m freestyle swimming relay events. PLOS ONE, 16, e0254538, 2021, https://doi.org/10.1371/journal.pone.0254538.

5 A. J. Ward-Smith, P. F. Radford. A mathematical analysis of the 4x100 m relay. Journal of Sports Sciences, 20, 369-381, 2002, https://doi.org/10.1080/026404102317366627.

6 S. Skorski, N. Etxebarria, K. G. Thompson. Breaking the Myth That Relay Swimming Is Faster Than Individual Swimming. International Journal of Sports Physiology and Performance, 11, 410-413, 2016, https://doi.org/10.1123/ijspp.2014-0577.

7 J. Huffmeier, S. Krumm. No myth far and wide: Relay swimming is faster than individual swimming and the conclusion of Skorski et al. (2016) is unfounded. Meta-Psychology, 2018.

8 K. McGibbon, D. B. Pyne, M. Shephard, M. Osborne, K. G. Thompson. Pacing and team strategy in relay events in swimming. Proceedings of the XIIIth International Symposium on Biomechanics and Medicine in Swimming, 280-286, Tsukuba, Japan, 2018.

9 K. E. McGibbon, M. E. Shephard, M. A. Osborne, K. Thompson, D. B. Pyne. Pacing and performance in swimming: Differences between individual and relay events. International Journal of Sports Physiology and Performance, 2020, https://doi.org/10.1123/ijspp.2020-0375.

10 S. Veiga, J. Santos del Cerro, L. Rodriguez, A. Trinidad, J. M. Gonzalez-Rave. How Mixed Relay Teams in Swimming Should Be Organized for International Championship Success. Frontiers in Psychology, 12, 573285, 2021, https://doi.org/10.3389/fpsyg.2021.573285.

11 S. P. McLean, M. J. Holthe, P. F. Vint, K. D. Beckett, R. N. Hinrichs. Addition of an approach to a swimming relay start. Journal of Applied Biomechanics, 16(4), 342-355, 2000.

12 T. Takeda, H. Takagi, S. Tsubakimoto. Comparison among three types of relay start in competitive swimming. In Biomechanics and Medicine in Swimming XI, 170-172, 2010.

13 X. Qiu, A. Lorenzo Calvo, B. De la Fuente Caynzos, S. Veiga. Comparison of the swimming start performance between individual and relay freestyle races. ISBS Proceedings Archive, 38(1), 2020.

14 X. Qiu, S. Veiga, A. Lorenzo, A. Kibele, E. Navarro. Differences in the key parameters of the individual versus relay swimming starts. Sports Biomechanics, 23(5), 598-610, 2024.

15 S. Veiga, C. Braun, X. Qiu, J. Stosic, S. Fuhrmann, A. Kibele, S. Fischer. What makes a successful relay start in swimming? Sports Biomechanics, 23(12), 3442-3453, 2024.

16 L. Ribeiro, A. M. Costa, H. Louro, P. Sobreiro, P. Esteves, A. Conceicao. Estimating time-to-contact with temporal occlusion in relay swimming: a pilot study. European Journal of Sport Science, 20(5), 592-598, 2020.

17 D. Luedtke, B. Duoos. Comparison of four feedback methods used to help improve swimming relay exchanges: A pilot study. International Journal of Aquatic Research and Education, 9(2), 2015.

18 S. Fischer, C. Braun, A. Kibele. Learning relay start strategies in swimming: What feedback is best? European Journal of Sport Science, 17(3), 257-263, 2017.

19 J. Martinez-Sobrino, S. Veiga, J. Santos del Cerro, J. M. Gonzalez-Rave. What is the Most Important Leg and Discipline in Triathlon Mixed Team Relays? Journal of Human Kinetics, 89, 269-278, 2023.

20 J. M. Saavedra, A. Garcia-Hermoso, Y. Escalante, A. M. Dominguez, R. Arellano, F. Navarro. Relationship between exchange block time in swim starts and final performance in relay races in international championships. Journal of Sports Sciences, 32(19), 1783-1789, 2014.

21 W. Siders. Competitive swimming relay exchange times: A descriptive study. International Journal of Sports Science & Coaching, 5(3), 381-387, 2010.

22 S. Fischer, C. Braun, A. Kibele. Jason Lezak again and again: Linear mixed modelling analysis of change-over times in relay swimming races. Journal of Sports Sciences, 37(14), 1609-1616, 2019.

23 X. Qiu, C. Braun, S. Fischer, A. Kibele. Performance gains in relay swimming (Part I): The relay start benefit component. ISBS Proceedings Archive, 40(1), 2022.

24 J. Huffmeier, G. Hertel. When the whole is more than the sum of its parts: Group motivation gains in the wild. Journal of Experimental Social Psychology, 47(2), 455-459, 2011.

25 J. Huffmeier, M. Filusch, J. Mazei, G. Hertel, A. Mojzisch, S. Krumm. On the boundary conditions of effort losses and effort gains in action teams. Journal of Applied Psychology, 102(12), 1673-1685, 2017.

26 J. E. Schleu, A. Mojzisch, J. Huffmeier. Run for the team: An analysis of effort gains in track and field relays. Psychology of Sport and Exercise, 45, 101567, 2019.

27 J. Huffmeier, J. E. Schleu, C. Nohe. The strength of the situation: Disentangling the situational explanation for effort gains in swimming relays from person-related accounts. Journal of Sport & Exercise Psychology, 42(5), 394-406, 2020.

28 C. Braun, S. Fischer, X. Qiu, J. Limmeroth, A. Kibele. Last and fast? A gender-specific analysis of effort gains in swimming relay events across Olympic Games and World Championships during the past 20 years. Psychology of Sport and Exercise, 55, 101949, 2021.

29 C. Braun, S. Fischer, J. Huffmeier. Larger effort gains in weaker relay swimmers independent of their starting order. Psychology of Sport and Exercise, 62, 102235, 2022.

30 A.-K. Torka, J. Huffmeier, S. Fischer, C. Braun. Social comparison, social indispensability, or both? On the psychological mechanisms underlying effort gains in swimming relays. Journal of Sport & Exercise Psychology, 46(6), 341-352, 2024.

31 M. Neugart, M. G. Richiardi. Sequential teamwork in competitive environments: Theory and evidence from swimming data. European Economic Review, 63, 186-205, 2013.

32 W.-J. Jane. Peer effects and individual performance: Evidence from swimming competitions. Journal of Sports Economics, 16(5), 531-539, 2015.

33 M. Glickman. Parameter estimation in large dynamic paired comparison experiments. Journal of the Royal Statistical Society: Series C, 48, 377-394, 1999, https://doi.org/10.1111/1467-9876.00163.

34 J. Nash. Equilibrium points in n-person games. Proceedings of the National Academy of Sciences, 36, 48-49, 1950, https://doi.org/10.1073/pnas.36.1.48.

35 C. F. Camerer, T. Ho, J. Chong. A cognitive hierarchy model of games. The Quarterly Journal of Economics, 119, 861-898, 2004, https://doi.org/10.1162/0033553041502225.

36 World Aquatics. Official results: men's 4x100m freestyle relay final, Paris 2024 Olympic Games. https://www.worldaquatics.com, 2024.

37 Swimcloud. 2024 competition season best times database. https://www.swimcloud.com, 2024.