Discussion Paper Series

IZA DP No. 18625

May 2026

Process Utility in High-Stakes Competition

Arnaud Dupuy
University of Luxembourg
and IZA@LISER

IZA DP No. 18625

May 2026

Process Utility in High-Stakes Competition[*]

Abstract

We study how individuals trade off outcome (what) and process (how) utility in high-stakes strategic decisions. We exploit optimality conditions and high-frequency choices in professional tennis to derive nonparametric bounds on process utility and implement a structural approach to estimate player-specific preferences. Under mild shape restrictions, these bounds imply that a large majority of players place positive weight on process utility. Our structural estimates further show that most players systematically sacrifice success probabilities to increase process utility, generating economically meaningful effects on match outcomes and expected earnings.

JEL classification
D91, D81, D01, C57

Keywords
process utility, intrinsic motivation, outcome utility, salience weight, strategic behavior, nonparametric, structural estimation

Corresponding author
Arnaud Dupuy
arnaud.dupuy@uni.lu

---

[*] We are very grateful to Jeff Sackmann for providing us access to the data from the Match Charted Project. We thank seminar participants at Maastricht University and, in particular, Lex Borghans, Arian Schwidder, and Christian Seel for interesting discussions and very useful comments on an earlier draft of the paper.

“When we act because we enjoy the activity itself, not because of what it leads to, the experience is autotelic — literally, having its goal within itself.” — Mihaly Csikszentmihalyi, Csikszentmihalyi (1990).

“Success is a journey, not a destination. The doing is often more important than the outcome.” — Arthur Ashe.
https://www.edwardssports.co.uk/news/post/practice-tennis-like-a-pro

1 Introduction

A central question in economics is the extent to which individuals are willing to trade off extrinsic rewards against the intrinsic enjoyment of an activity (e.g., Frey and Stutzer (2002); Bénabou and Tirole (2003)).[1] In many instances, this trade-off can be framed as one between an outcome (what) and the process (how) by which it is obtained. For instance, Adam Smith’s discussion of compensating wage differentials and its formalization in modern labor economics (e.g., Rosen (1986)) acknowledges that job attributes such as effort, risk, or enjoyment enter workers’ utility alongside wages. Empirically, however, quantifying this trade-off remains challenging. Observational data typically raises two important issues: preferences are confounded with constraints, selection, and unobserved heterogeneity, and identification relies on between-person comparisons, as indeed the frequency of such decisions is generally limited.[2] Experimental evidence circumvents these issues and has provided important insights (e.g., Gneezy and Rustichini (2000b,a); Ariely et al. (2008)), but often relies on low-stakes laboratory environments with limited external validity. As a result, there is relatively little direct field evidence at the individual level on whether people are willing to sacrifice measurable economic outcomes in exchange for intrinsically rewarding processes in high-stakes, competitive settings.

Professional tennis provides a uniquely well-suited setting to study this trade-off. Players operate in a high-stakes environment with strong extrinsic motivation (monetary and reputational incentives) to maximize performance. Yet, they face repeated, strategically rich decisions—most notably in serving with the second serve rule—where they can choose between higher-risk strategies that increase the probability of winning the point immediately in one shot (unreturned serve) but also of making a mistake, and more conservative strategies that lead to longer rallies. Crucially, because points are won either in a single shot or through multi-shot exchanges, tennis offers a setting to separately identify outcome and process motivations. Finally, being one of the most broadcast sports on the planet, a wealth of point-level data is recorded for many matches, allowing a rich analysis of tennis rallies for a large number of players. This combination of high incentives, clear strategic margins,

---

[1] See also Deci et al. (1999), Bandiera et al. (2005), Falk and Kosfeld (2006), Bryan et al. (2011), Larkin et al. (2012), Falk et al. (2013), Kube et al. (2013), Corgnet et al. (2016) among others.

[2] For instance, in our labor market example, workers who accept lower wages for more enjoyable jobs may differ systematically in ability or outside options (e.g., Mas (2025); Lavetti (2023)) and only change jobs a limited number of times in the course of their career.

and granular data makes tennis an ideal setting to test whether individuals deviate from outcome-maximizing behavior to engage in more intrinsically rewarding forms of play.

We propose a simple framework in which players derive utility from winning points -outcome utility– but also from the way in which they win points -process utility–. In particular, we allow process utility to depend on whether a point is won in one shot, immediately after the serve, or through multi-shot rallies, and model serving decisions as the outcome of utility maximization under this extended preference structure. We call the relative importance of process utility in total utility the *salience weight*. Our framework yields a natural interpretation of observed deviations from outcome-maximizing strategies[3] as reflecting heterogeneous salience weights, which we recover for each player from point-level data.

Using optimizing serving decisions in professional tennis, we first develop a nonparametric identification strategy that exploits optimality conditions to derive player-specific bounds on the salience weight. We show that, under mild shape restrictions, these bounds are informative enough to establish that a large majority of players have a strictly positive salience weight. Further exploiting these optimality conditions, we adopt a parametric approach and propose an algorithm to structurally recover player-specific salience weights on process utility. Applying this framework to detailed point-level data from the Sackmann Charting Project, we estimate player-specific skills and salience weights. We find that 79% of professional players have a positive salience weight—of which 64% are statistically significant at the 5% level—and therefore place greater weight on winning multi-shot rallies, resulting in systematically more conservative (second-)serve strategies than predicted by outcome-maximizing models.[4] Counterfactual exercises show that eliminating process utility would increase point-winning probability on serve by about 0.4 percentage points, translating into a 2.4 percentage point increase in match-winning probability and an increase of $33,000 (13.5%) in expected prize money at a Grand Slam tournament (US Open 2025). These results provide field evidence that individuals are willing to forgo measurable extrinsic rewards to engage in intrinsically rewarding activities, and illustrate how small deviations from outcome-maximizing behavior can have economically meaningful consequences in high-stakes competitive settings.

Methodologically, our model fits naturally within the canonical framework of trade-offs between extrinsic and intrinsic motivation. To fix ideas, consider an agent choosing among alternatives leading to an outcome. Each alternative x is associated with outcome utility $p(x)$ and process utility $k(x)$. In its simplest form, utility is a weighted sum of the two, where $\delta \geq 0$ denotes the relative weight on process utility. When $\delta = 0$, the agent is purely outcome-maximizing, while larger values of δ reflect increasing salience of process utility.

This formulation is consistent with leading models of motivation, where process utility arises from civic duty (Frey and Oberholzer-Gee (1997)), identity (Loewenstein (1999), Akerlof and Kranton (2000)), implicit contracts (Gneezy and Rustichini (2000a,b)), beliefs (Bénabou and Tirole (2003)), procedures (Frey et al. (2004)), meaning (Ariely et al. (2008), Norton et al. (2012)), gambling (Menestrel (2001)), or the act of choice itself (Sen (1997)). In our setting, a player chooses a serve strategy x and trades off the probability of winning

---

[3] See Borghans (1995); Klaassen and Magnus (2009) for instance.

[4] For 119 out of 151 players in the data, the estimated salience weight for winning multi-shot rallies is larger than 0. For 76 of them, this weight is significantly different from 0 at the 5% level. For only 3 players, this weight is significantly lower than 0 at the 5% level.

a point on serve $p(x)$ against the probability of winning the point through multiple shots $k(x)$. By structurally recovering player-specific salience weights from optimal choices, we quantify the extent to which individuals deviate from outcome maximization to engage in intrinsically rewarding play.

Our analysis builds on the premise that winning multi-shot rallies is intrinsically more rewarding than winning one-shot rallies. We motivate this in three steps. First, from the psychological literature, we learn that enjoyment and self-determination are important drivers of intrinsic motivation. Research first formalized by Csikszentmihalyi (1975, 1990) indicates that *flow* is a state of complete absorption in an activity, often described as being "in the zone," that arises when the challenge of the task aligns with the performer's skill level, and therefore "[e]njoyment appears at the boundary between boredom and anxiety, when the challenges are just balanced with the person's capacity to act." Csikszentmihalyi (1990), pp. 52–53.[5] Flow theory connects closely with *Self-Determination Theory* (SDT) Deci and Ryan (2000), which posits that intrinsic motivation is fostered when autonomy, competence, and relatedness are satisfied.[6] Flow states and Self-Determination are therefore more likely to arise during multi-shot rallies than through the execution of a single shot, the serve.

Second, from players' testimonies, we learn that enjoyment is indeed an important component of their motivation. Professional players frequently highlight enjoyment as a central goal: for instance, Carlos Alcaraz stated, "I just want to step on court . . . and try to enjoy as much as I can."[7]

Third, in tennis, although the server wins roughly 45% of his points on unreturned serves (one-shot), top players spend most of their training time (about 90%) on baseline rallies as shown in O'Shannessy (2019) and Fitzpatrick (2024). This apparent paradox in the behavior of players during practice supports the idea that players attach more importance to multi-shot rallies, likely so because they enjoy playing such rallies more. For these reasons, we expect process utility to arise from winning multi-shot rallies, and player-specific salience weights for these rallies capture the preference for winning points through multi-shot rallies relative to one-shot points.

This paper contributes in several important ways. First, it extends the analysis of tennis serving strategies of Klaassen and Magnus (2009) to account for process utility and distinguish between one-shot and multi-shot points. It proposes a novel method to compute player-specific salience weights from point-level data and quantifies the economic consequences of process-driven strategy, demonstrating that small deviations from outcome maximization can have substantial effects.

Second, more broadly, our results speak to a general class of preferences in which individuals derive utility not only from outcomes but also from the process by which these outcomes are achieved. This idea is closely related to the compensating differentials literature in labor economics, which documents that workers are willing to accept lower wages

---

[5] Csikszentmihalyi illustrates this with tennis: "One cannot enjoy doing the same thing [at tennis] at the same level for long. We grow either bored or frustrated, and then the desire to enjoy ourselves again pushes us to stretch our skills, or to discover new opportunities for using them." Csikszentmihalyi (1990), p.75.

[6] In tennis, autonomy arises from controlling shot selection and tactics; competence from improving skills and executing difficult shots; and relatedness from interactions with coaches, opponents, and the public.

[7] `https://www.atptour.com/en/news/alcaraz-lehecka-us-open-2025-qf`. A broader set of players' quotes is collected in Online Appendix (B1).

in exchange for non-pecuniary job attributes such as meaningful or intrinsically rewarding tasks, with estimated trade-offs on the order of 8–20% depending on the job attributes considered (e.g., Stern (2004); Mas and Pallais (2017)). Similarly, in consumer markets, a large body of evidence on fair trade, ethical consumption, and product provenance shows that individuals are willing to pay substantial price premia for goods produced under socially or environmentally desirable conditions, at comparable product quality. Experimental and field evidence typically finds willingness-to-pay premia in the range of approximately 5% to 30% for fair trade or ethically certified products, depending on product category (e.g., see Maertens and Swinnen (2009); Dragusanu et al. (2014)). Across these settings, individuals appear willing to trade off outcome against process, be it about having autonomy, providing meaning, or being ethical. Our contribution is to show that this same trade-off can be identified from high-frequency, within-individual, continuous choices in a high-stakes, competitive strategic environment, in contrast to labor and consumer studies, which typically rely on low-frequency or discrete-choice settings.

The remainder of the paper is structured as follows. Section 2 presents the model, discusses non-parametric bounds and introduces a parametric approach together with an algorithm to recover player-specific salience weights from the data. Section 3 describes the data and the estimation method, while Section 4 presents the results. Section 5 provides robustness checks, and Section 6 concludes.

2 Model

2.1 Preliminary observations

In tennis, a central strategic decision for the server is how much risk to take. A "safe" serve increases the probability that the ball lands in the service box (i.e., a higher serve percentage) but reduces the likelihood of winning the point outright through an ace or unreturned serve. Conversely, a "risky" serve raises the probability of winning the point immediately, at the cost of a higher probability of a fault.

While this resembles a standard high-risk–high-reward trade-off, the tennis setting involves an additional margin: not only whether the point is won, but how it is won. A safe serve increases the likelihood that the point evolves into a multi-shot rally, forcing the player to win through multiple shots, whereas a risky serve increases the likelihood of a one-shot win.

The server's decision, therefore, determines both the overall probability of winning the point and the distribution of this probability over one-shot and multi-shot points. As a result, the serve strategy reflects a trade-off not only between high-risk–high-reward outcomes, but also between one-shot wins and multi-shot wins, which may be intrinsically valued differently.

In support of this distinction, we propose the following defintions.

Definition (Outcome Utility): Outcome utility captures preferences over the results of an action, independent of how they are achieved. In our setting, it corresponds to the probability of winning a point, regardless of whether it is won through one-shot or multi-shot rallies.

Definition (Process Utility): Process utility captures preferences over how outcomes are achieved. In our setting, it is represented by a preference for winning points through multi-shot rallies rather than one-shot points.

2.2 Set up

Let the probability of a serve being in be denoted by x. This probability reflects the choice of the server. As depicted in the above observations, if the server wants to take more risk, he will choose a lower value of x, hence a lower serve percentage. In contrast, if he wants to take fewer risks, he will choose a higher value of x, a higher serve percentage. Since in tennis, players can serve a second serve if the first is out, the server's strategy consists, in fact, of two numbers x_1 and x_2 reflecting respectively the probability of the first and second serve to be in. A player's probability to win a point on his serve, denoted $p(x_1, x_2)$, depends on his strategy (x_1, x_2). Denote $y(x)$ the probability of winning a point conditional on the serve being in as a function of the serve probability, x. $y(x)$ reflects the skills of the player, encompassing both his serving and rally skills (relative to his opponent). In condition (1) below we assume that $y(x)$ is twice differentiable and in particular, strictly decreasing, i.e., $y'(x) < 0$, so that the safer the serve, that is, the higher the probability that it is in, the lower the probability of winning the point, conditional on the serve being in, and strictly concave, $y''(x) < 0$.

With these definitions, the probability of winning a point on one's own serve, the outcome utility, reads as,

$$p(x_1, x_2) = w(x_1) + (1 - x_1)\, w(x_2),$$

where $w(x) := xy(x)$ is the unconditional probability of winning a point, and a server aiming at maximizing his probability of winning a point on his serve then does $\max_{x_1, x_2} p(x_1, x_2)$.

This setting corresponds to the basis of the model presented in Klaassen and Magnus (2009) and discussed in more detail in Online Appendix (B2). In this paper, we depart from Klaassen and Magnus (2009) by assuming that players are perfect maximizers and allowing them to care about process utility, and hence possibly put different weights to the various possible ways of winning a point. This requires distinguishing between 4 possible ways to win a point: with one shot on the first or second serve, i.e., an ace or an unreturned serve, or with multiple shots on the first or second serve, i.e., a rally of more than 2 shots. We hence decompose the conditional probability to win a point into a conditional probability to win in one shot, say $f(x)$, and in multiple shots, say $k(x)$. By definition, one has $y(x) = f(x) + k(x)$. It seems natural to expect that $f'(x) < 0$ and $f''(x) \leq 0$ so f is concave, meaning that the conditional probability of winning a one-shot point is decreasing with the probability of the serve to be in (increasing with risk), more so as the level of risks decreases (concave).[8] In contrast, the conditional probability of winning a multi-shot point may be increasing or decreasing with the probability that the serve is in, depending on the skills of the player. Hence, we impose that, if it is decreasing, it is also concave, i.e., $k'(x) < 0$ and $k''(x) \leq 0$, whereas, if it is increasing, it is convex, i.e., $k'(x) > 0$ and $k''(x) > 0$.

To summarize, we assume that the following standing assumptions hold.

---

[8] This assumption is supported in the data. Indeed, for all professional players in the data, the percentage of first serves in is lower than the percentage of second serves in ($x_1 < x_2$), but only for two players, the share of one-shot points won on first serve ($f(x_1)$) is lower than that on the second serve ($f(x_2)$).

Condition 1 *The conditional probability of winning a one-shot point continuous and twice differentiable, strictly decreasing* $f'(x) < 0$ *and concave* $f''(x) < 0$*, whereas the conditional probability to win a multi-shot point is continuous and either strictly decreasing* $k'(x) < 0$ *and concave* $k''(x) \leq 0$*, or increasing* $k'(x) \geq 0$ *and convex* $k''(x) \geq 0$ *and so that* $y(x) := f(x) + k(x)$ *is so that* $y'(x) < 0$ *and* $y''(x) < 0$*.*

The second departure from the model in Klaassen and Magnus (2009), is that we consider the case where a player may attach more or less, but not necessarily the same, importance to winning a one-shot point rather than a multi-shot point. In the model, there are four possible (winning) outcomes for the server, listed below with their specific probability to occur and specific utility:

1. One shot on first serve: probability $x_1 f(x_1)$ and utility of α,
2. Multiple shots on first serve: probability $x_1 k(x_1)$ and utility of β,
3. One shot on second serve: probability $(1 - x_1) x_2 f(x_2)$ and utility of α,
4. Multiple shots on second serve: probability $(1 - x_1) x_2 k(x_2)$ and utility of β.

Hence, we assume that the player maximizes, not his probability to win a point $p(x_1, x_2)$, but rather, $\tilde{p}(x_1, x_2)$ defined as the weighted average of the probability to win a one-shot point and the probability to win a multi-shot point, which reads as

$$\tilde{p}(x_1, x_2) = \tilde{w}(x_1) + (1 - x_1)\tilde{w}(x_2).$$

where $\tilde{w}(x) = x\tilde{y}(x)$, $\tilde{y}(x) = \alpha f(x) + \beta k(x)$, α is the utility attached to winning a one-shot point, and β is the utility attached to winning a multi-shot point.

A first important remark is that normalizing the utility of winning one-shot rallies to $\alpha = 1$ is without loss of generality, as it does not affect the optimal solution of a player. From now on, we therefore set $\alpha = 1$ and interpret β as the *relative* preference parameter for multi-shot rallies. Moreover, in the case $\beta = 1$, distinguishing between f and k is irrelevant since all that matters is the conditional probability of winning a point $y(x) = f(x) + k(x)$ and not how. In terms of notation, we call $\tilde{y}$ the *perceived* conditional probability of winning a point because of the presence of preference parameter β in the expression. It only coincides with the true probability when $\beta = 1$. A similar interpretation and notation is used for $w(x)$ and $p(x_1, x_2)$.

A second important remark is that this utility rewrites as

$$\tilde{p}(x_1, x_2) = \underset{\text{Outcome Utility}}{p(x_1, x_2)} + (\beta - 1)\underset{\text{Process Utility}}{(x_1 k(x_1) + (1 - x_1) x_2 k(x_2))}$$

clearly distinguishing the outcome utility, i.e., the probability of winning a point on one's own serve, and process utility, i.e., the probability of winning a multi-shot point on one's own serve. Presented this way, it is clear that our model relates to the simple model of outcome and process utility presented in the introduction, where the salience weight δ obtains as $\delta = \beta - 1$.

To summarize, our setting distinguishes between one-shot and multi-shot rallies and allows preference weights to be different for these two types of rallies, replacing $w(x)$ by $\tilde{w}(x)$ and decomposing $y(x)$ into the constituants $f(x)$ and $k(x)$ to compute $\tilde{y}(x)$.

The FOCs to this problem are obtained as

$$\begin{aligned} \tilde{w}'(x_1) - \tilde{w}(x_2) &= 0, \\ \tilde{w}'(x_2) &= 0. \end{aligned}$$

The second order conditions require that the expected utility is concave. For this, the Hessian of the expected utility needs to be semi-definite negative and since at optimum, the Hessian is diagonal (see Online Appendix (B3.1)), the SOCs therefore are

$$\begin{aligned} 2\tilde{y}'(x_1) + x_1\tilde{y}''(x_1) &\leq 0, \\ (1-x_1)\left(2\tilde{y}'(x_2^*) + x_2^*\tilde{y}''(x_2^*)\right) &\leq 0. \end{aligned}$$

Interestingly, the SOCs provide restrictions on the curvature of the perceived conditional probability of winning a point. Indeed, rearranging both SOCs, one obtains:

$$x_j^* \frac{\tilde{y}''\left(x_j^*\right)}{\tilde{y}'\left(x_j^*\right)} \geq -2,$$

provided that $\tilde{y}'\left(x_j^*\right) < 0$, $\forall j = 1, 2$.[9]

Note that $x\frac{\tilde{y}''(x)}{\tilde{y}'(x)}$ is the elasticity of the perceived marginal probability $\tilde{y}'(x)$ and the SOCs, in fact, indicate that this elasticity should be larger than -2 on the interval $x \in [x_1^*, x_2^*]$.

Moreover, the FOCs reveal important information about the shape of the perceived conditional probability $\tilde{y}$ at the optimum. Indeed, consider the FOC associated with the optimal second serve strategy. Rearranging, one has

$$\tilde{y}'(x_2^*) = \frac{\tilde{y}(x_2^*)}{x_2^*}.$$

By a similar procedure, the FOC for the first serve strategy obtains as

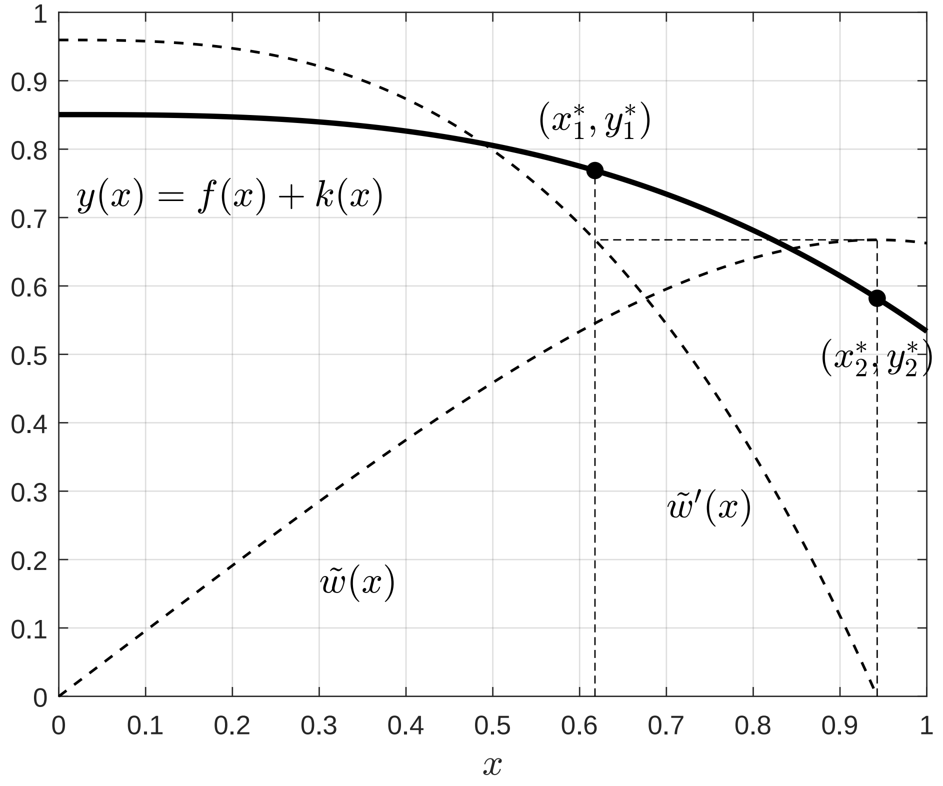
$$\tilde{y}'(x_1^*) = \frac{x_2^*\tilde{y}(x_2^*) - \tilde{y}(x_1^*)}{x_1^*}.$$

It follows that the optimal serve strategy (x_1^*, x_2^*) pins down the slope of the perceived conditional probability of winning a point at both x_2^* and x_1^* to respectively $\frac{\tilde{y}\left(x_2^*\right)}{x_2^*}$ and $\frac{x_2^*\tilde{y}\left(x_2^*\right)-\tilde{y}\left(x_1^*\right)}{x_1^*}$. Importantly, this means that the average curvature of $\tilde{y}(x)$ in the interval

---

[9] Note that, while it is possible that β is so that $\tilde{y}'(x) > 0$, even if $y'(x) < 0$, in that case, the optimal second serve strategy would be $x_2^* = 1$ since $\tilde{w}'(x) > 0$. For all players in the data, the observed second serve percentage is strictly lower than 1, and hence this situation never occurs.

Figure 1: Optimal service strategy: Roger Federer



$[x_1^*, x_2^*]$ is known and equal to

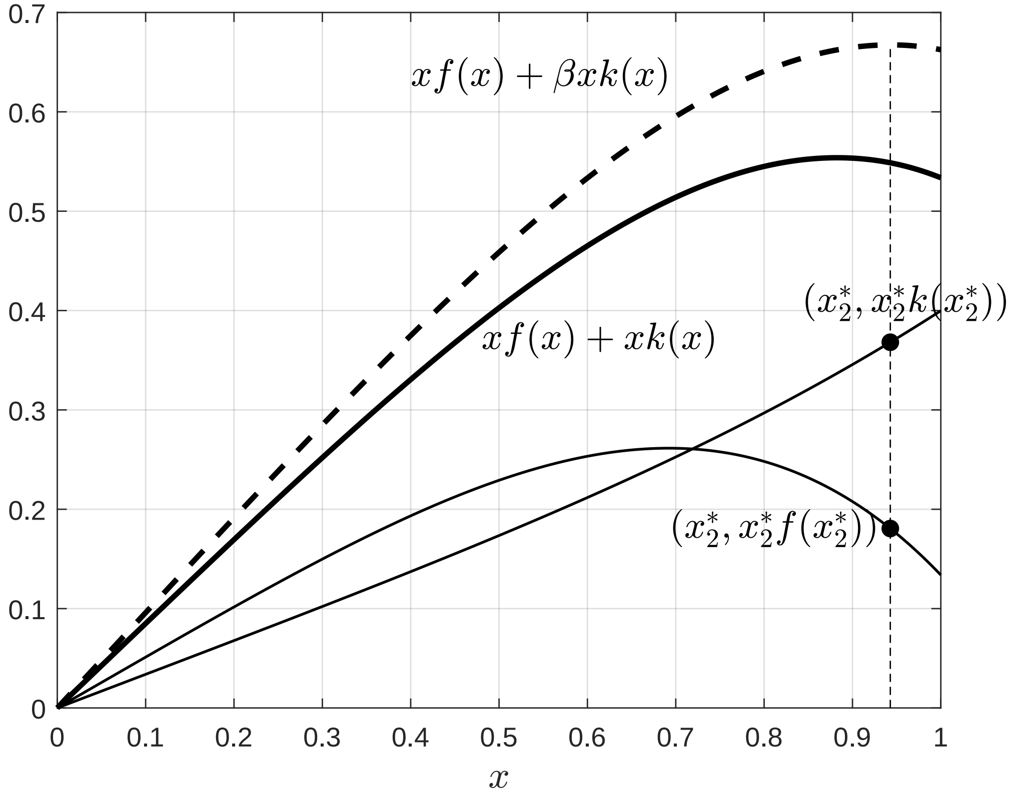
$$\begin{aligned}
\frac{1}{x_2^* - x_1^*} \int_{x_1^*}^{x_2^*} \tilde{y}''(x)\, dx &= \frac{\tilde{y}'(x_2^*) - \tilde{y}'(x_1^*)}{x_2^* - x_1^*} \\
&= \frac{\frac{\tilde{y}(x_2^*)}{x_2^*} - \frac{x_2^* \tilde{y}(x_2^*) - \tilde{y}(x_1^*)}{x_1^*}}{x_2^* - x_1^*}.
\end{aligned}$$

Figure (1) shows how the optimal service strategy is determined given the skills parameters of the player (the shapes of f and k) and his relative preference for winning multi-shot rallies (β). We herewith use the example of Roger Federer.

First, the second serve strategy is found by looking at the value of x for which $\tilde{w}'(x) = 0$, say x_2^*. Then, the first serve strategy is derived by looking at the value x for which $\tilde{w}'(x) = \tilde{w}(x_2^*)$, say x_1^*. The conditional probabilities of winning a point given that the first (second) serve is in are indicated on the curve $y(x)$, for the respective values of x_1^* and x_2^*. This delivers two observed points (x_1^*, y_1^*) and (x_2^*, y_2^*).

The additional source of identification provided by distinguishing between one-shot and multi-shot rallies can be best illustrated using Figure (2). This figure shows how the optimal second service strategy is determined given the skills parameters of the player (the shapes of f and k) and his preference parameter for multi-shot rallies (β) and how this generates one point on the conditional probability of one-shot rallies and one point on the conditional probability of multi-shot rallies. We herewith use the example of Roger Federer.

Figure 2: Optimal second serve strategy: Roger Federer.



The optimal second serve strategy is obtained as the value x_2^* that maximizes $\tilde{w}(x) = xf(x) + \beta xk(x)$. Note that it does not maximize $w(x) = xf(x) + xk(x)$. Although changing the value of β might result in a non-negligible change in the second serve percentage, we see that the probability of winning a point on one's serve might actually only change marginally, $w(x_2^*)$ being close to $\max_x(w(x))$. This marginal change, however, because of the scoring system in tennis, magnifies into a larger change in the probability to win a game, a set, and eventually a match. This, in turn, transforms into significantly larger expected prize money.

2.3 Nonparametric approach

Our aim is to nonparametrically bound the salience weight δ. The key idea is that optimal serve choices impose inequality restrictions on $\tilde{p}(.,.)$, which translate into player-specific bounds on δ. To show this, we first need to briefly introduce the data. Suppose that, for each player $i = 1, ..., N$, we observe the probabilities $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$ where x_{1i} and x_{2i} are the probabilities of first and second serves in, $f_{1i} = f\left(x_{1i}\right)$ and $f_{2i} = f\left(x_{2i}\right)$ are the conditional probabilities of winning a point with one shot on first and second serves, respectively, and $k_{1i} = k\left(x_{1i}\right)$ and $k_{2i} = k\left(x_{2i}\right)$ are the conditional probabilities of winning a point with multiple shots on first and second serves, respectively.[10]

We assume that each player i is a perfect maximizer of his perceived probability of

---

[10] In the data section, we discuss how to estimate these probabilities $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$ given data on rallies played, in possibly multiple matches, on the serve of each player i.

winning a point $\tilde{p}(x_1, x_2)$ so that the observed data reflect the optimum of each player, i.e., $(x_{1i}, x_{2i}) = (x_{1i}^*, x_{2i}^*)$. Dropping the index i for notational simplicity, optimality implies for each player

$$\tilde{p}(x_1^*, x_2^*) \geq \tilde{p}(x_1^*, x_1^*), \tag{1}$$

$$\tilde{p}(x_1^*, x_2^*) \geq \tilde{p}(x_2^*, x_2^*), \tag{2}$$

$$\tilde{p}(x_1^*, x_2^*) \geq \tilde{p}(x_2^*, x_1^*). \tag{3}$$

The first inequality compares the observed strategy to a deviation in which the player uses the first-serve strategy in both serves; the second compares it to always using the second-serve strategy; the third swaps the two strategies. Note that inequality (1) implies $\tilde{p}(x, x_2^*) \geq \tilde{p}(x, x_1^*)$ for all x, so that inequality (3) is in fact implied by inequalities (1-2).[11] Each inequality above can be written in linear form

$$\begin{aligned} A + \delta B &\geq 0 \quad \Rightarrow \quad \delta \geq -\frac{A}{B} \quad \text{if } B > 0, \\ A + \delta B &\geq 0 \quad \Rightarrow \quad \delta \leq -\frac{A}{B} \quad \text{if } B < 0, \end{aligned}$$

where the parameter A is in fact the difference in the probability of winning a point ($p(.,.)$ not $\tilde{p}(.,.)$) between the two strategies being compared in the inequality whereas the parameter B is a similar difference but for the probability of winning a multi-shot point. By construction, $A - B := C$ is the same difference but for the probability of winning a one-shot point.

The optimality conditions can be used to define a lower bound L (when $B > 0$) and an upper U (when $B < 0$) on the value of the salience weight $\delta \in [L, U]$ for each player. For all players we note that the lower bound L is determined by inequality (1) whereas the upper bound is either unrestricted ($+\infty$) by the optimality conditions, which is the case for 23 players, or determined by inequality (2). Regarding the sign of the bounds, the data actually show that for 8 players the lower bound is positive ($B > 0$ and $A < 0$) so that we can conclude that the salience weight of these players must be positive. For the remaining players, the lower bound is negative ($B > 0$ and $A > 0$), and the upper bound positive ($B < 0$ and $A > 0$). Even though the optimality conditions alone are not enough to determine the sign of the salience weight δ for all but 8 players, interestingly, the lower bound L is determined by the upper bound of $\tilde{p}(x_1^*, x_1^*)$, i.e., inequality (1).

Building on this observation, let $x \in [x_1^*, x_2^*]$ and consider the inequality

$$\tilde{p}(x, x_2^*) \geq \tilde{p}(x_1^*, x_1^*). \tag{4}$$

Associated with this inequality are the values

$$\begin{aligned} A(x) &= p(x, x_2^*) - p(x_1^*, x_1^*) = B(x) + C(x), \\ B(x) &= xk(x) + (1 - x) x_2^* k_2^* - (2 - x_1^*) x_1^* k_1^*, \\ C(x) &= xf(x) + (1 - x) x_2^* f_2^* - (2 - x_1^*) x_1^* f_1^*. \end{aligned}$$

---

[11] Indeed, note that, as long as one maintains the second serve strategy constant, the first serve strategy contributes the same term $\tilde{w}(x)$ and a slope $(1 - x)$ to $\tilde{p}(x, x_2)$ and $\tilde{p}(x, x_2)$ so that if the inequality $\tilde{p}(x, x_2) > \tilde{p}(x, x_2)$ holds for x, it holds for all $x' \neq x$.

This inequality delivers a lower bound for the player's salience weight if and only if $B(x) > 0$. This lower bound is negative if $A(x) > 0$ and positive if $A(x) < 0$.[12] We derive our main result from the following lemma, which provides a sufficient condition under which the lower bound on δ is positive. The conditions (b)–(c) in the lemma are sign restrictions directly verifiable in the data, while condition (d) restricts the curvature of $f(.)$.

Lemma 1 *If a): Condition (1) is satisfied, b):* $B(x_1^*) > A(x_1^*) > 0$*, c):* $B(x_2^*) > 0 > A(x_2^*)$*, and d):* $C'(x_1^*) < -\frac{C(x_1^*)}{x_2^* - x_1^*}$*, then there exists a unique* $x_0 \in [x_1^*, x_2^*]$ *such that* $A(x_0) = 0$ *and* $B(x_0) > 0$*.*

Proof. (Sketch) Existence follows from continuity and the Intermediate Value Theorem. Uniqueness follows from the strict concavity of $A(x)$. Finally, concavity of $C(x)$ implies that at the root x_0, $C(x_0) < 0$ and hence $B(x_0) > 0$. See Appendix (A2) for the full proof. ■

Conditions (b)–(c) of Lemma (1) are met for all but one player in the data. Condition (d) is better understood once rearranged using that $C'(x) = f(x) + xf'(x) - x_2^* f_2^*$ by definition to obtain

$$f'(x_1^*) < \frac{x_2^* f_2^* - f_1^* - \frac{C(x_1^*)}{x_2^* - x_1^*}}{x_1^*}.$$

The left-hand side is unknown, but the right-hand side offers an upper bound for Lemma (1) to apply, which can be computed from the data. The higher this upper bound, the less restrictive this condition is and the more likely the lemma's conditions are satisfied. To ease interpretation, however, one can express it as a fraction F of the average slope on $[x_1^*, x_2^*]$ which too can be computed on the data as $\frac{f_2^* - f_1^*}{x_2^* - x_1^*}$. A similar implication holds for F, i.e., the higher the fraction, the less restrictive the condition of the lemma.

We first note that for the eight players for whom the salience weight is revealed as positive from optimality conditions, the fraction F ranges from 0.02 to 0.31 with a mean/median value of 0.14. For the remaining 143 players, F ranges from 0.12 to 0.69 with a mean/median value of 0.33, indicating that roughly 50% of the players have a fraction at least as large as the maximum value observed among the players with revealed positive salience weight. Finally, 133 players (93%) have a fraction higher than 0.20, corresponding to the second highest value in the group of eight.

2.4 Parametric approach

Our objective is to estimate the salience weight δ (or equivalently, the relative preference for winning a multi-shot point β) to conduct counterfactual analyses. This requires imposing a parametric structure on the model, in particular on the functions $f(x)$ and $k(x)$. Recall that the SOCs of the optimization problem give a restriction on the elasticity of $\tilde{y}'(x)$, the perceived marginal probability of winning a point. Since $\tilde{y}'(x) = f'(x) + \beta k'(x)$, we propose to parametrize the elasticity of the marginal probabilities $f'(x)$ and $k'(x)$ as in the following condition.

---

[12] We already know that for 8 players in the data when $x = x_1^*$, which corresponds to Inequality (1), $A(x_1^*) < 0$ and $B(x_1^*) > 0$.

Condition 2 $x\frac{f''(x)}{f'(x)} = x\frac{k''(x)}{k'(x)} = \lambda - 1 > 0.$[13]

Condition (2) imposes that the elasticities of the marginal probability of winning a point with one-shot and with multi-shot, conditional on the serve being in, are equal to each other and to a strictly positive constant $\lambda - 1$.

This condition has three main implications. First, the elasticity of both the perceived and true marginal probability of winning a point is also equal to $\lambda - 1$, as under condition (2) one has $x\frac{\tilde{y}''(x)}{\tilde{y}'(x)} = x\frac{y''(x)}{y'(x)} = \lambda - 1$.[14] Furthermore, since, by assumption $y'(x) < 0$, this means that $y(x)$ follows the law of diminishing marginal *returns* (read conditional probability of winning a point).

Second, it means that $f(x)$ and $k(x)$ are power functions of the form $f(x) = \frac{a_f - x^\lambda}{\tau_f}$ and $k(x) = \frac{a_k - x^\lambda}{\tau_k}$, offering great flexibility with only 5 unknown parameters.

Third, as a by product of the two preceeding remarks, the conditional probability of winning a point $y(x)$ is itself a power function with power λ, as indeed $y(x) = f(x) + k(x) = \frac{a - x^\lambda}{\tau}$, where $\tau = \frac{\tau_f \tau_k}{\tau_f + \tau_k}$ and $a = \frac{a_f \tau_k + a_k \tau_f}{\tau_f + \tau_k}$, see Online Appendix (B4).[15]

Associated with this parametric choice, Conditions (1) are met with the following restrictions on the parameters of $f(x)$ and $k(x)$.

Condition 3 *i)* $\lambda > 1$, *ii)* $\tau_f > 0$ *and either* $\tau_k > 0$ *or* $-\tau_k > \max(\tau_f, \beta\tau_f)$, *iii)* $0 < \frac{a_f \tau_k + a_k \tau_f}{\tau_k + \tau_f} < \lambda + 1$.

Note that Condition (3.ii) garantees that $y(x) = f(x) + k(x)$ is decreasing, as indeed it leads to $\tau = \frac{\tau_f \tau_k}{\tau_f + \tau_k} > 0$, and $f'(x) < 0$ from $\tau_f > 0$.[16]

With these parametric choices, the optimal first and second serve strategies can be derived in closed form from the first-order conditions (see Online Appendix (B3.2)). One obtains for the second serve strategy

$$x_2^{*\lambda} = \frac{a_f + a_k \beta \frac{\tau_f}{\tau_k}}{(1+\lambda)\left(1 + \beta\frac{\tau_f}{\tau_k}\right)},$$

yielding an interior solution, i.e., $0 < x_2^* < 1$, if and only if

$$\lambda + 1 > \frac{a_f + a_k \beta \frac{\tau_f}{\tau_k}}{1 + \beta\frac{\tau_f}{\tau_k}} > 0.$$

---

[13] Note that this assumption corresponds to the CRRA assumption of the von Neumann-Morgenstern utility function in the context of a classical problem of expected utility. In Online Appendix (B7) we present the model using the CARA assumption instead, i.e., $\frac{f''(x)}{f'(x)} = \frac{k''(x)}{k'(x)} = \lambda > 0$. The associated functions are of the form $f(x) = a_f + \tau_f \exp(\lambda x)$.

[14] Indeed, since one has $x\frac{f''(x)}{f'(x)} = x\frac{k''(x)}{k'(x)} = \lambda - 1$, it follows that $xk''(x) = (\lambda - 1)k'(x)$ and $xf''(x) = (\lambda - 1)f'(x)$ and since $xk''(x) + xf''(x) = xy''(x)$ and $xk''(x) + x\beta f''(x) = x\tilde{y}''(x)$, one has

$$\begin{aligned} x\tilde{y}''(x) &= (\lambda - 1)\tilde{y}'(x), \\ xy''(x) &= (\lambda - 1)y'(x), \end{aligned}$$

which for $\tilde{y}'(x), y'(x) > 0$ yields the result in the text.

[15] This corresponds to the functional shape assumed in Klaassen and Magnus (2009) for $y(x)$.

[16] In our data, these conditions are met for all players except for the condition $-\tau_k > \beta\tau_f$, which is not met for 3 of them.

And, the first serve strategy, therefore obtains as

$$\begin{aligned} x_1^{*\lambda} &= \frac{a_f + a_k\beta\frac{\tau_f}{\tau_k} - x_2^*\left(a_f + a_k\beta\frac{\tau_f}{\tau_k} - \left(1+\beta\frac{\tau_f}{\tau_k}\right)x_2^{*\lambda}\right)}{(1+\lambda)\left(1+\beta\frac{\tau_f}{\tau_k}\right)} \\ &= x_2^{*\lambda}\left(1 - \frac{\lambda}{1+\lambda}x_2^*\right). \end{aligned}$$

where the second line follows by simple substitution.

A remarkable result is that at optimum,[17] one has

$$\left(\frac{x_1^*}{x_2^*}\right)^{\lambda} = 1 - \frac{\lambda}{1+\lambda}x_2^*$$

meaning that, when players are optimizers, data on first and second serve percentages, and in particular $\frac{x_1^*}{x_2^*}$ and x_2^*, uniquely identify the curvature parameter λ as shown in section (2.6).

To summarize, under our standing assumptions, at optimality, the server adopts the following service strategy

$$\left(x_1^{*\lambda}, x_2^{*\lambda}\right) = \left(x_2^{*\lambda} - \frac{\lambda}{1+\lambda}x_2^*x_2^{*\lambda}, \frac{a_f + a_k\beta\frac{\tau_f}{\tau_k}}{(1+\lambda)\left(1+\beta\frac{\tau_f}{\tau_k}\right)}\right),$$

and enjoys the following conditional probabilities of winning a point in one and multiple shots on first and second serve:

$$\begin{aligned} f(x_1^*) &= \frac{a_f - x_1^{*\lambda}}{\tau_f}, k(x_1^*) = \frac{a_k - x_1^{*\lambda}}{\tau_k}, \\ f(x_2^*) &= \frac{a_f - x_2^{*\lambda}}{\tau_f}, k(x_2^*) = \frac{a_k - x_2^{*\lambda}}{\tau_k}. \end{aligned}$$

An important remark is that although the preference parameter β and the relative skills $\frac{\tau_f}{\tau_k}$ only enter the expressions of the optimal serve strategy on first and second serve, through the term $\beta\frac{\tau_f}{\tau_k}$, the expressions for the conditional probabilities of winning a one-shot or multi-shot point at the optimum depend respectively, only (directly) on the skills parameters τ_f and τ_k. At same value of $\beta\frac{\tau_f}{\tau_k}$, i.e., at same optimal serve strategy, players with different skills τ_f and τ_k have different optimal conditional probabilities of winning a one-shot or multi-shot point. This is the source for the separate identification of the preference parameter and the skills parameters to exploit in the data.

(x_1^*, x_2^*) is an optimum if the expected utility is concave at (x_1^*, x_2^*). With the parametric shapes assumed above, the conditions for the expected utility to be concave are given as

$$\begin{aligned} -\tau_f\lambda(1+\lambda)(x_1^*)^{\lambda-1}\left(1+\beta\frac{\tau_f}{\tau_k}\right) &\leq 0 \\ -\tau_f(1-x_1)\lambda(1+\lambda)(x_2^*)^{\lambda-1}\left(1+\beta\frac{\tau_f}{\tau_k}\right) &\leq 0 \end{aligned}$$

---

[17] This result arises not only with power functions but also with softmax functions, see Online Appendix (B7).

Since $\tau_f > 0$, $\lambda > 0$, $x_1^* > 0$, and from Condition (3.ii) one has either $\tau_k > 0$ or $-\tau_k > \beta\tau_f$ so that $1 + \beta\frac{\tau_f}{\tau_k} > 0$, we conclude that the expected utility is concave.

It is easy to show by simple substitution that when $\beta = 1$, the optimal strategy is

$$\left(x_1^{*\lambda}, x_2^{*\lambda}\right) = \left(x_2^{*\lambda}\left(1 - \frac{\lambda}{\lambda+1}x_2^*\right), \frac{a}{\lambda+1}\right)$$

and the SOC reads as $\tau \geq 0$.

2.5 Comparative statics

Note that for $\lambda > 1$, as estimated in the data, the signs of the derivatives $\frac{\partial x_j^*}{\partial\beta}$ for $j = 1, 2$, agree with those of $\frac{\partial x_j^{*\lambda}}{\partial\beta}$ which are simpler to derive. Hence, w.l.o.g., we produce comparative statics on the preference parameter β on the optimal strategy of the server using $x_j^{*\lambda}$ rather than x_j^*.

Consider a small change in the preference parameter for winning multi-shot rallies on the optimal probability of serving first and second serves in, one has

$$\frac{\partial x_1^{*\lambda}}{\partial\beta} = -\tau_f\tau_k\frac{a_f - a_k}{(1+\lambda)\left(\tau_k + \beta\tau_f\right)^2}\left(1 - x_2^*\right)$$

so that $sign\left(\frac{\partial x_1^{*\lambda}}{\partial\beta}\right) = -sign\left(\tau_k\left(a_f - a_k\right)\right)$ since $x_2^* < 1$ and

$$\begin{aligned}\frac{\partial x_2^{*\lambda}}{\partial\beta} &= \frac{\partial\frac{a_f+\beta a_k\frac{\tau_f}{\tau_k}}{(1+\lambda)\left(1+\beta\frac{\tau_f}{\tau_k}\right)}}{\partial\beta}\\ &= -\tau_f\tau_k\frac{a_f - a_k}{(1+\lambda)\left(1+\beta\tau_f\right)^2}\end{aligned}$$

so that $sign\left(\frac{\partial x_2^{*\lambda}}{\partial\beta}\right) = -sign\left(\tau_k\left(a_f - a_k\right)\right) = sign\left(\frac{\partial x_2^*}{\partial\beta}\right)$.

Hence, the optimal probability of serving first or second serves in, is increasing with the preference parameter put on winning multi-shot points if $\tau_k\left(a_f - a_k\right) < 0$ and vise versa.[18]

Proposition 2 *If $\tau_k\left(a_f - a_k\right) < 0$ (resp. $\tau_k\left(a_f - a_k\right) > 0$), the optimal probability of serving first or second serves in is increasing (resp. decreasing) with the preference for winning multi-shot rallies.*

Note that, as can already be anticipated from the above comparative statics, whenever $a_f = a_k$, the optimal serving strategy does not depend on the preference parameter for winning multi-shot points of the player. This is comfirmed by looking at this strategy for $a_f = a_k = a$. One indeed has

$$\left(x_1^{*\lambda}, x_2^{*\lambda}\right) = \left(a\frac{1 - x_2^*\left(1 - \frac{1}{a}x_2^{*\lambda}\right)}{1+\lambda}, a\frac{1}{1+\lambda}\right).$$

These expressions do not depend on β.

---

[18] In our data, all but one players meet the condition $\tau_k\left(a_f - a_k\right) < 0$.

2.6 Identification and computation

For each player $i = 1, ..., N$, we observe the probabilities $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$ and have 5 unknown skill parameters $(\lambda_i, a_{fi}, \tau_{fi}, a_{ki}, \tau_{ki})$ and 1 unknown preference parameter β_i. The two points (x_{1i}, f_{1i}) and (x_{2i}, f_{2i}) can be used to identify 2 of the 3 parameters of the function $f(x)$, whereas the two points (x_{1i}, k_{1i}) and (x_{2i}, k_{2i}) can be used to identify 2 of the 3 parameters of the function $k(x)$. Since $f(x)$ and $k(x)$ have one parameter in common, i.e., λ_i, this means that the four points (x_{1i}, f_{1i}), (x_{1i}, k_{1i}), (x_{2i}, f_{2i}) and (x_{2i}, k_{2i}) together only identify 4 of the 5 skills parameters. Assuming that these players are perfect optimizers, the optimality conditions imply that $(x_{1i}, x_{2i}) = (x^*_{1i}, x^*_{2i})$ which provides 2 restrictions to the system of equations. Hence, we have 6 parameters to be identified by 4 data points and 2 optimality restrictions.

2.6.1 Identification

To show the identification of parameters $(\lambda_i, a_{fi}, \tau_{fi}, a_{ki}, \tau_{ki}, \beta_i)$ given data $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$, we first show identification of parameters $(a_{fi}, \tau_{fi}, a_{ki}, \tau_{ki}, \beta_i)$ given data $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$ conditional on λ_i, and propose a bisection algorithm that searches for the curvature parameter λ given the data (x_{1i}, x_{2i}).

First, the slope and constant terms of the functions $f(x)$ and $k(x)$ are identified given $\lambda_i = \lambda$. Indeed, as soon as the value of λ is known, data (x_{1i}, x_{2i}) can be used to compute $(z_{1i}, z_{2i}) = \left(x^{\lambda}_{1i}, x^{\lambda}_{2i}\right)$. It follows that, using the functional form for $f(x)$ and $k(x)$, from the points (z_{1i}, f_{1i}) and (z_{2i}, f_{2i}), by simply rearranging terms, one can deduce the slopes

$$\begin{aligned} \tau_{fi} &= -\frac{z_{2i}}{f_{2i}}\frac{\Delta z_i}{\Delta f_i}, \\ \tau_{ki} &= -\frac{z_{2i}}{k_{2i}}\frac{\Delta z_i}{\Delta k_i}, \end{aligned}$$

where $\Delta l_i = \frac{l_{1i}-l_{2i}}{l_{2i}}$ $\forall l = z, f, k$, and then the constants

$$\begin{aligned} a_{fi} &= \tau_{fi} f_{1i} + z_{1i}, \\ a_{ki} &= \tau_{ki} k_{1i} + z_{1i}. \end{aligned}$$

Interestingly, note that the relative slope does not depend on the value of λ, and is solely determined by the observed data as

$$\frac{\tau_{fi}}{\tau_{ki}} = \frac{k_{2i}}{f_{2i}}\frac{\Delta k_i}{\Delta f_i}.$$

Second, one can uncover the preference parameter β_i once the slope and constant parameters of $f(x)$ and $k(x)$ are known. This is done by using the previous results together with the equation for the optimal second serve strategy to isolate

$$\beta_i = -\frac{\tau_{ki}}{\tau_{fi}}\frac{a_{fi} - z_{2i}(1+\lambda)}{a_{ki} - z_{2i}(1+\lambda)}$$

provided $z_{2i}(1+\lambda) - a_{ki} \neq 0$. Using the expressions previously obtained, for $x_{2i} > 0$, $f_{2i} > 0$ and $k_{2i} > 0$, it can be shown that in fact

$$\beta_i = -\frac{f_{2i}}{k_{2i}} \frac{1 + \lambda \frac{\Delta f_i}{\Delta z_i}}{1 + \lambda \frac{\Delta k_i}{\Delta z_i}},$$

and it follows that one must have $\Delta z_i + \lambda \Delta k_i \neq 0$.

Note that after following these steps, the curvature parameter is the only remaining unknown parameter. We can then use the last remaining condition, i.e., the FOC for the optimal first serve strategy, to implicitly solve for the curvature parameter and obtain

$$\lambda_i = \frac{a_{fi} + a_{ki}\beta_i \frac{\tau_{fi}}{\tau_{ki}} - x_{2i}\left(a_{fi} + a_{ki}\beta_i \frac{\tau_{fi}}{\tau_{ki}} - \left(1 + \beta_i \frac{\tau_{fi}}{\tau_{ki}}\right) z_{2i}\right)}{\left(1 + \beta_i \frac{\tau_{fi}}{\tau_{ki}}\right) z_{1i}} - 1$$

since $z_{1i} > 0$.

Importantly, as shown in Online Appendix (B5), this expression simplifies considerably to read as

$$\lambda_i = \frac{z_{2i}}{z_{1i}} \left(1 + \lambda \left(1 - x_{2i}\right)\right) - 1.$$

This expression actually shows that the curvature parameter only depends on data x_{1i} and x_{2i} through z_{1i} and z_{2i} and the initial value of λ selected. Of course, the value λ_i herewith obtained might be different than the λ used to compute z_{1i} and z_{2i}. However, as it turns out, this equation together with the structure of the problem, provides a fixed-point, so that there exists, for each player i, a λ such that $\lambda_i = \lambda$. We shall see below that a relatively simple algorithm allows us to recover this value, for all the players in the data, and that this value is larger than unity for all players.

2.6.2 Computation

We propose the following algorithm to compute the parameter λ_i of each player i, from the associated data (x_{1i}, x_{2i}).

Algorithm 1 *Data* (x_{1i}, x_{2i})*, tolerance* ε*, Output* λ_i*.*

1. *Initializing step. Set* $t = 0$ *and initialize the interval* $\left[l^{(t)}, u^{(t)}\right]$ *so that* $\Lambda_i\left(l^{(t)}\right) < l^{(t)}$ *and* $\Lambda_i\left(u^{(t)}\right) > u^{(t)}$*. In practice,* $l^{(t)}$ *should be close to* 0 *and* $u^{(t)}$ *large enough.*
2. *Solving step. Set* $\lambda = \frac{l^{(t)} + u^{(t)}}{2}$ *and using data* (x_{1i}, x_{2i}) *compute*
$$\lambda^{new} = \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda} \left(1 + \lambda \left(1 - x_{2i}\right)\right) - 1.$$
3. *Incrementation step. If* $\lambda^{new} < \lambda$, $l^{(t)} = \lambda$*, else* $u^{(t)} = \lambda$*. If* $\left|l^{(t)} - u^{(t)}\right| < \varepsilon$*, stop, else, set* $t \leftarrow t + 1$ *and go back to solving step.*

Note that the iteration step of Algorithm (1) defines a map $\Lambda_i : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, associating for each $\lambda^{(t)}$ a new value $\lambda^{(t+1)}$ for the curvature parameter given the data (x_{1i}, x_{2i}). If it exists, a solution is therefore a fixed-point $\Lambda_i\left(\lambda_i^*\right) = \lambda_i^*$. The following theorem shows that such a solution exists and is unique under mild conditions for the structure of the data (x_{1i}, x_{2i}).

Theorem 3 (Existence and uniqueness of the solution) *For each player i, with data (x_{1i}, x_{2i}) so that*

(A1) $x_{1i}, x_{2i} > 0$ *with* $0 < x_{1i} < x_{2i} < 1$,

(A2) $\ln \frac{x_{2i}}{x_{1i}} < x_{2i}$,

(A3) $2\left(x_{2i} - x_{1i}\right) < \left(x_{2i}\right)^2$,

there exists a unique $\lambda_i^* \geq 1$ *such that*

$$\lambda_i^* = \Lambda_i(\lambda_i^*).$$

Proof. See Appendix (A3) ■

We illustrate the theorem using the example of Roger Federer. Figure (3) plots the map $\Lambda_i(\lambda)$ for Roger Federer. We see that the properties of the map exploited in the proof, are so that the map crosses the 45 degree line only ones at λ^*, the point $(\lambda^*, \Lambda_i(\lambda^*))$ forming the unique (non zero) fixed-point of the map. Suppose that the algorithm starts with lower bound $l^{(0)} = 1$ and upper bound $u^{(0)} = 4$. The mid point value is 2.5. At $\lambda = 2.5$, the value of the map is below the 45 degree line and the lower bound is updated to 2.5. The termination condition is not satisfied (at conventional levels of precision) and the algortihm goes back to the iteration step with lower bound 2.5 and upper bound 4. The mid point is now 3.75 so that the value of the map is higher than the 45 degree line. The upper bound is updated to 3.75, etc.. The algorithm converges very fast to the value $\lambda^* = 2.81$.

Once the curvature parameter λ_i is obtained from the algorithm, one can compute the remaining parameters using the steps outlined in the Section (2.6.1).

3 Data

We use the Match Charting Project by Jeff Sackmann,[19] which collects information about professional tennis matches, encoded by dozens of contributors. In particular, we use the point-by-point data for men's matches with information on more than 1,200,000 rallies of over 7,100 matches by the end of January 2026.

The unit of observation is a rally in a match. For each rally, we know who is serving, whether it is a rally on first or second serve, and the length of the rally, i.e., the number of shots that were recorded "in" the court. A rally of length 1 is necessarily ending with either an ace or an unreturned serve. The server wins all rallies of odd length, whereas the returner wins rallies of even length. For our analysis, for each player, we need to observe a large number of rallies on their own serve. For this reason, we only select those players with at least 20 matches charted in the data. 151 players satisfy this criterion.

Table (1) presents service statistics for selected players, along with sample descriptive statistics. The selected players include the four players with the most Grand Slam titles[20],

---

[19] https://github.com/JeffSackmann/tennis_MatchChartingProject

[20] The four most prestigious tournaments of the year are the Australian Open, Roland Garros, Wimbledon, and the US Open. These tournaments also have the most generous prize money distributions.

Figure 3: Fixed-point of the map $\Lambda(\lambda)$: Roger Federer.

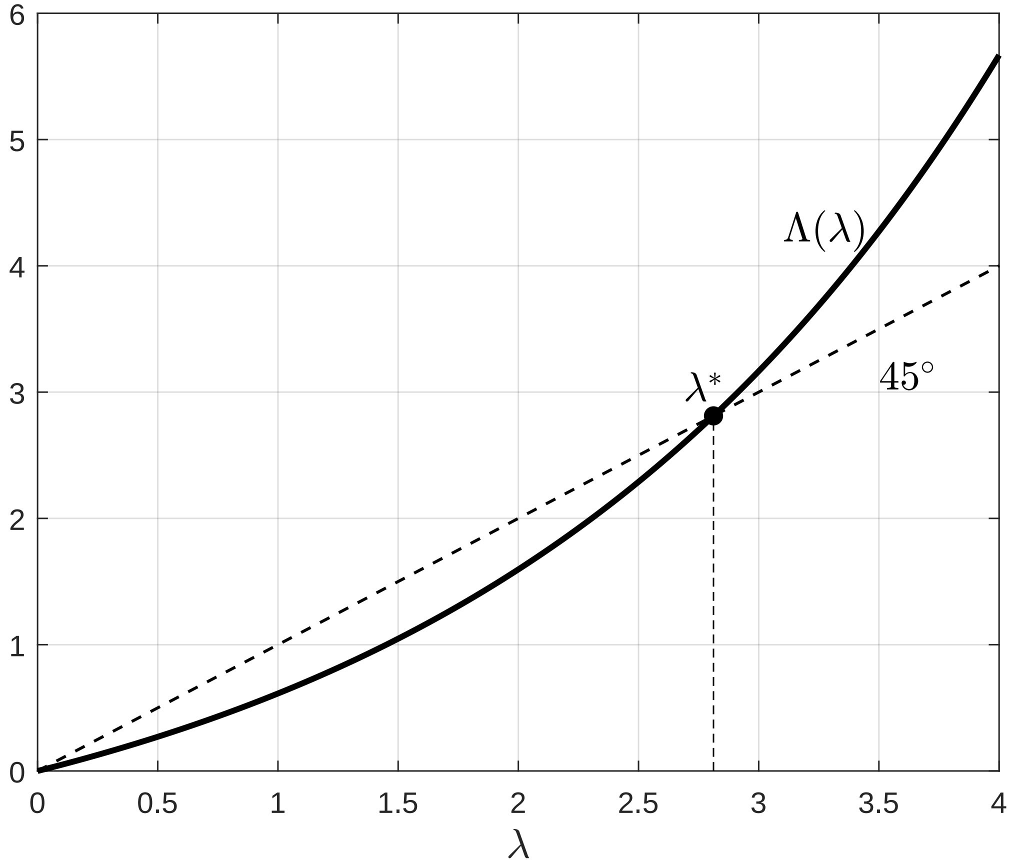


Table 1: First and second serves statistics for selected players and aggregate sample.

		First serve				Second serve			
Player	Nb pt. ($\times$1000)	% in x_1	% Multi	% One won $f(x_1)$	% Multi won $k(x_1)$	% in x_2	% Multi	% One won $f(x_2)$	% Multi won $k(x_2)$
Novak Djokovic	44.8	64.9	65.9	34.1	39.2	92.0	83.2	16.8	42.0
Rafael Nadal	34.0	68.6	72.4	27.6	42.9	92.6	83.9	16.1	43.4
Roger Federer	58.0	61.8	58.5	41.5	35.4	94.3	80.8	19.2	39.0
Pete Sampras	20.3	56.5	46.4	53.6	26.7	89.7	78.5	21.5	34.2
...									
Boris Becker	15.7	55.0	52.3	47.7	29.5	88.2	80.5	19.5	32.9
...									
Carlos Alcaraz	16.5	64.6	68.0	32.0	40.0	90.9	82.2	17.8	42.0
Jannik Sinner	20.1	60.3	61.6	38.4	37.4	93.3	83.0	17.0	42.2
...									
Ivo Karlovic	3.6	65.4	43.0	57.0	23.8	86.3	69.0	31.0	28.8
John Isner	7.9	69.5	45.9	54.1	23.5	91.9	73.6	26.4	32.1
Reilly Opelka	4.4	65.6	45.0	55.0	22.4	91.9	74.0	26.0	33.8
...									
David Ferrer	6.0	61.7	76.0	24.0	40.2	91.3	86.8	13.2	38.9
Diego Schwartzman	3.9	63.9	78.4	21.6	42.0	89.6	85.7	14.3	41.2
...									
Mean	6.4	61.1	63.6	36.4	35.2	90.9	83.1	16.9	37.4
Std	7.5	4.1	8.6	8.6	4.8	3.1	3.9	3.9	3.2
Min	1.3	51.5	39.4	19.5	21.3	74.0	59.3	10.2	27.1
Median	3.7	61.4	63.7	36.3	35.5	91.6	83.7	16.3	37.6
Max	58.0	71.3	80.5	60.6	43.3	96.8	89.8	40.7	43.7

i.e., the Greatest Of All Times (GOATs): Novak Djokovic, Rafael Nadal, Roger Federer, and Pete Sampras; Boris Becker; the best two players of the current generation, Carlos Alcaraz and Jannik Sinner; three players known for their big serves, John Isner, Reilly Opelka, and Ivo Karlovic; and two players known for their baseline game, David Ferrer and Diego Schwartzman.

The data include between 1,300 and 58,000 rallies per player. On average, the first serve percentage is about 61%, while the second serve percentage is roughly 91%. Players win approximately 72% of points played on their first serve, with an equal distribution between one-shot and multi-shot rallies, whereas 63% of points on the first serve are multi-shot rallies. On the second serve, being more conservative, players win fewer one-shot points (17%) and about 37% of multi-shot rallies. The percentage of multi-shot rallies on second serves is 83%, roughly 20 percentage points higher than on the first serve.

There are also notable disparities across players. Big servers, such as John Isner, win about twice as many one-shot points as multi-shot points on their first serve, whereas the reverse is nearly true for baseline specialists like David Ferrer. Comparing newer top players to earlier ones, Carlos Alcaraz's statistics resemble those of Rafael Nadal, and Jannik Sinner's resemble those of Roger Federer. Finally, the percentage of points won on one-shot rallies ranges from 10% to 41% on the second serve and from 20% to 61% on the first serve, while for multi-shot rallies, the corresponding ranges are 27% to 44% on the second serve and 21% to 43% on the first serve.

3.1 Maximum likelihood estimation of $(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$

Let $\theta_i = (x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i})$ be the unknown probabilities of interest. From the data, for each player $i = 1, ..., N$, we observe N_i points played on his serve. We can compute the number of points played on the first serve, i.e., $n_{x_1 i}$ and on the second serve $n_{x_2 i} = N_i - n_{x_1 i}$ and since the data identifies for each point, the length of the rally, i.e., the number of shots played into the court during the point, we can also compute the number of rallies of length 1 on first and second serves, i.e., $n_{f_1 i}$ and $n_{f_2 i}$ and the number of other rallies of odd length on the first and second serves as well, i.e., $n_{k_1 i}$ and $n_{k_2 i}$.

Note that Klaassen and Magnus (2001) showed that even though points in Tennis are not i.i.d., the deviations from the i.i.d. hypothesis are small. As a result, the i.i.d. hypothesis can still be used as a good approximation when aggregating over a large number of points as we do in this paper, where we use "averages" over points by players. We therefore maintain the assumption that points are i.i.d. so that each of the aforementioned variables follows a binomial distribution.[21] For instance, on the first serve, one has:

1. the number of first serves in, $n_{x_1 i}$, represents the number of successes resulting from N_i independent draws with unknown probability of success x_{1i}, meaning that

$$n_{x_1 i} \sim B\left(N_i, x_{1i}\right),$$

2. the number of rallies of length 1 on first serves, $n_{f_1 i}$, represents the number of successes resulting from $n_{x_1 i}$ independent draws with unknown probability of success f_{1i}, meaning

---

[21] See Online Appendix (B6) for a tree representation of the statistics for each player.

that

$$n_{f_1 i} \sim B\left(n_{x_1 i}, f_{1i}\right),$$

3. the number of rallies won with multiple shots on first serve, $n_{k_1 i}$, represents the number of successes resulting from $n_{x_1 i} - n_{f_1 i}$ independent draws with unknown probability of success k_{1i}, meaning that[22]

$$n_{k_1 i} \sim B\left(n_{x_1 i} - n_{f_1 i}, k_{1i}\right).$$

It implies, for instance, that the log-likelihood of observing data $\left(n_{f_1 i}, n_{x_1 i}\right)$ given probability f_{1i} is

$$l_{f_1 i}\left(n_{f_1 i}, n_{x_1 i} | \theta_i\right) = \log \binom{n_{x_1 i}}{n_{f_1 i}} + n_{f_1 i} \log f_{1i} + \left(n_{x_1 i} - n_{f_1 i}\right) \log\left(1 - f_{1i}\right).$$

Applying the same logic to all data and collecting the associated terms of the log-likelihood, obtains

$$\begin{aligned} l_{.i}\left(N_i, n_{x_1 i}, n_{x_2 i}, n_{f_1 i}, n_{f_2 i}, n_{k_1 i}, n_{k_2 i} | \theta_i\right) &= l_{x_1 i}\left(n_{x_1 i}, N_i | \theta_i\right) + l_{x_2 i}\left(n_{x_2 i}, N_i - x_{1i} | \theta_i\right) \\ &\quad + l_{f_1 i}\left(n_{f_1 i}, n_{x_1 i} | \theta_i\right) + l_{f_2 i}\left(n_{f_2 i}, n_{x_2 i} | \theta_i\right) \\ &\quad + l_{k_1 i}\left(n_{k_1 i}, n_{x_1 i} - n_{f_1 i} | \theta_i\right) + l_{k_2 i}\left(n_{k_2 i}, n_{x_2 i} - n_{f_2 i} | \theta_i\right). \end{aligned}$$

Consider the derivative of the log-likelihood with respect to f_{1i}. It reads as

$$\begin{aligned} \frac{\partial l_{.i}\left(N_i, n_{x_1 i}, n_{x_2 i}, n_{f_1 i}, n_{f_2 i}, n_{k_1 i}, n_{k_2 i} | \theta_i\right)}{\partial f_{1i}} &= \frac{\partial l_{f_1 i}\left(n_{f_1 i}, n_{x_1 i} | \theta_i\right)}{\partial f_{1i}} \\ &= \frac{n_{f_1 i}}{f_{1i}} - \frac{n_{x_1 i} - n_{f_1 i}}{1 - f_{1i}}. \end{aligned}$$

The first order condition of the maximum likelihood requires that this derivative is set equal to 0, implying that

$$\begin{aligned} \frac{n_{f_1 i}}{f_{1i}} - \frac{n_{x_1 i} - n_{f_1 i}}{1 - f_{1i}} &= 0 \\ &\Leftrightarrow \\ n_{f_1 i}\left(1 - f_{1i}\right) &= f_{1i}\left(n_{x_1 i} - n_{f_1 i}\right) \\ &\Leftrightarrow \\ f_{1i} &= \frac{n_{f_1 i}}{n_{x_1 i}}. \end{aligned}$$

Hence, the frequency $\frac{n_{f_1 i}}{n_{x_1 i}}$ is the maximum likelihood estimate of probability f_{1i}. A similar argument holds for all other probabilities. Let $\hat{f}_{1i} = \frac{n_{f_1 i}}{n_{x_1 i}}$ denote the empirical of frequency of f_{1i} with a similar notation for the other terms. Then, by maximum likelihood, the empirical frequencies $\left(\hat{x}_{1i}, \hat{x}_{2i}, \hat{f}_{1i}, \hat{f}_{2i}, \hat{k}_{1i}, \hat{k}_{2i}\right)$ are the estimates of $\theta_i = \left(x_{1i}, x_{2i}, f_{1i}, f_{2i}, k_{1i}, k_{2i}\right)$.

---

[22] Note that, conditional on the serve being in, a multi-shot rally only occurs if a one-shot rally fails, giving tennis rallies data, as structured in this paper, a very natural representation in terms of (conditional, sequential) binomial distributions.

3.2 Checking the data

Theorem (3) applies on data satisfying conditions (A1)-(A3), which garantees the convergence of Algorithm (1) to a solution λ larger than unity. As indicated in Table (1), for all professional tennis players in our data, the first serve percentage (x_1) ranges between 0.51 and 0.72 while the second serve percentage ranges between[23] 0.73 and 0.97. This trivially shows that condition (A1) is always met. Since $x_{1i} > 0.5$, one has

$$\max_i \ln\left(\frac{x_{2i}}{x_{1i}}\right) < \ln 2 \simeq 0.69 < \min_i x_{2i} = 0.73,$$

which also garantees that Condition (A2) is satisfied for the ranges of values for x_1 and x_2 observed in the data. It follows that, for all players in the data, there exists a unique solution λ_i^*. Last, note that Condition (A3) rewrites as $x_{2i}\left(2 - x_{2i}\right) < 2x_{1i}$ and in the data,

$$\max_i x_{2i}\left(2 - x_{2i}\right) < 1 < \min_i 2x_{1i} = 1.02$$

so that this condition is also satisfied for all players in the data. Hence, the unique solution λ_i^* is strictly greater than unity for all $i = 1, .., N$.

---

[23] In fact, Maxime Cressy and Alexandre Bublik are the only two players with second serve percentages lower than 84%.

4 Results

Table 2: Parameter estimates with 95% bootstrap confidence intervals for selected players

Player	Salience weight δ	Curvature λ	Slope f τ_f	Const. f a_f	Slope k τ_k	Const. k a_k
Novak Djokovic	0.27^{***} [0.15, 0.41]	3.67^{***} [3.56, 3.81]	3.09^{***} [2.93, 3.26]	1.26^{***} [1.21, 1.30]	-19.58^{***} [-31.23, -14.32]	-7.48^{***} [-12.11, -5.36]
Rafael Nadal	0.21^{***} [0.03, 0.47]	4.89^{***} [4.70, 5.10]	4.62^{***} [4.23, 5.07]	1.43^{***} [1.34, 1.54]	-105.09 [-1079.64, 953.52]	-44.96 [-463.38, 410.43]
Roger Federer	0.32^{***} [0.23, 0.42]	2.81^{***} [2.73, 2.90]	2.64^{***} [2.55, 2.74]	1.35^{***} [1.32, 1.39]	-16.27^{***} [-21.06, -13.41]	-5.50^{***} [-7.26, -4.44]
Pete Sampras	0.42^{***} [0.31, 0.56]	1.93^{***} [1.82, 2.05]	1.49^{***} [1.43, 1.56]	1.13^{***} [1.10, 1.16]	-6.41^{***} [-7.86, -5.45]	-1.38^{***} [-1.80, -1.10]
Boris Becker	0.56^{***} [0.39, 0.79]	1.73^{***} [1.62, 1.85]	1.59^{***} [1.51, 1.68]	1.12^{***} [1.08, 1.15]	-13.23^{***} [-24.26, -9.18]	-3.55^{***} [-6.93, -2.30]
Carlos Alcaraz	0.05 [-0.12, 0.26]	3.66^{***} [3.45, 3.88]	3.56^{***} [3.22, 3.98]	1.34^{***} [1.24, 1.46]	-25.78^{***} [-115.53, -13.94]	-10.11^{***} [-46.84, -5.27]
Jannik Sinner	0.06 [-0.04, 0.19]	2.52^{***} [2.39, 2.66]	2.61^{***} [2.46, 2.77]	1.28^{***} [1.24, 1.33]	-11.65^{***} [-16.24, -9.01]	-4.08^{***} [-5.87, -3.04]
Ivo Karlovic	1.19^{***} [0.63, 2.36]	4.37^{***} [3.78, 5.11]	1.42^{***} [1.13, 1.74]	0.96^{***} [0.78, 1.15]	-7.41^{***} [-20.52, -4.43]	-1.61^{***} [-4.88, -0.88]
John Isner	0.79^{***} [0.52, 1.16]	5.33^{***} [4.89, 5.83]	1.78^{***} [1.59, 2.01]	1.11^{***} [1.00, 1.23]	-5.68^{***} [-7.92, -4.45]	-1.19^{***} [-1.75, -0.88]
Reilly Opelka	0.36^{***} [0.14, 0.65]	3.89^{***} [3.47, 4.33]	1.81^{***} [1.60, 2.06]	1.19^{***} [1.07, 1.32]	-4.63^{***} [-6.32, -3.61]	-0.84^{***} [-1.29, -0.58]
David Ferrer	-0.01 [-0.32, 0.61]	2.90^{***} [2.62, 3.24]	4.82^{***} [4.04, 5.83]	1.41^{***} [1.24, 1.62]	40.72 [-431.32, 261.69]	16.62 [-172.34, 103.65]
Diego Schwartzman	-0.38 [-0.75, 0.30]	3.57^{***} [3.15, 4.07]	6.45^{***} [4.70, 9.97]	1.60^{***} [1.26, 2.24]	62.62 [-427.40, 425.94]	26.49 [-174.09, 177.09]
Mean	0.34	2.94	2.96	1.24	3.96	2.52
Std	0.70	0.98	1.36	0.19	96.83	38.20
Min	-3.86	1.28	1.04	0.62	-352.94	-129.32
Median	0.33	2.86	2.62	1.22	-10.22	-3.17
Max	2.79	6.27	7.99	1.69	900.62	359.60

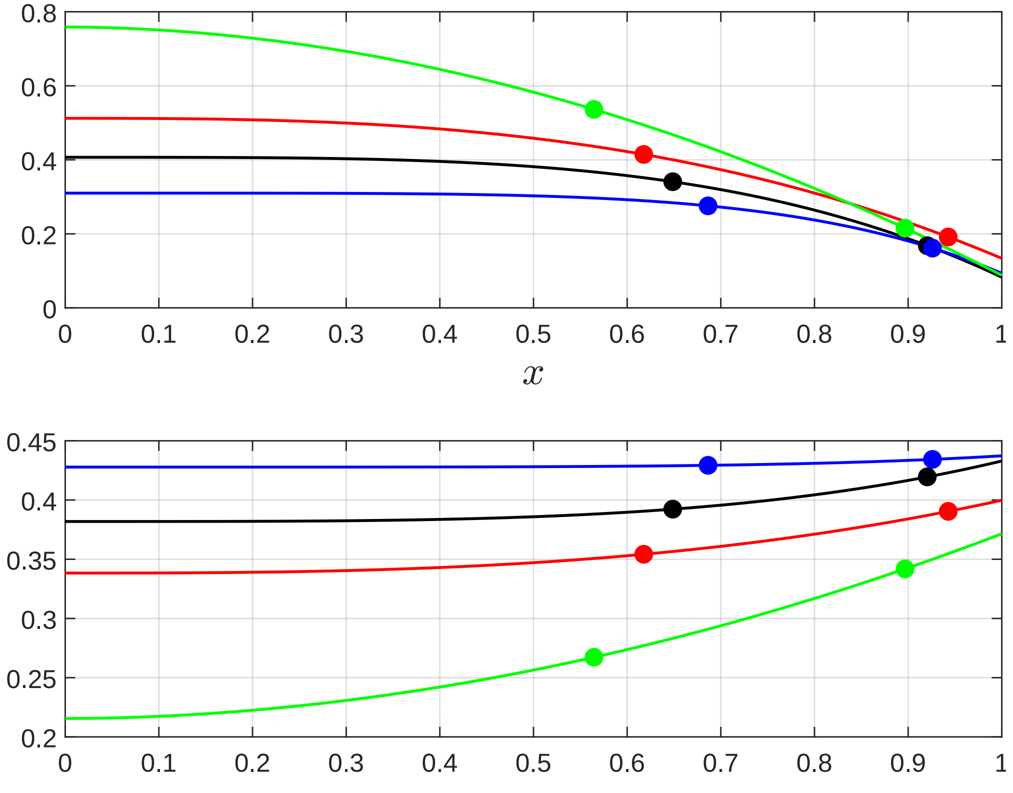
Notes: 95% confidence intervals obtained by bootstrap with 300 replications.*** indicates significance at the 5% level: for λ, CI excludes 1; for other parameters, CI excludes 0.

Table (2) presents the estimates of the parameters for our selected players as well as the summary statistics for all players (bottom rows). The mean and median salience weights in our sample are both 1/3, indicating a preference for winning multi-shot rallies. In fact, 79% (119/151) of the players have a positive salience weight, and for about 64% (76/119) of them, that coefficient is statistically significant at 5%. In contrast, there are 32 players with negative salience weight, and only 3 of them for whom that estimate is statistically significant at 5%. The first four listed players in the table are the GOATs. We see that for all of them, the salience weight is positive and significant. The salience weight is also positive and significant for big servers like John Isner, Ralley Opelka, and Ivo Karlovic. Interestingly, the salience weight for the new top players, Carlos Alcaraz and Jannik Sinner, is positive (0.05) but not significant. In particular, although the statistics in Table (1) for Jannik Sinner were quite close to those of Roger Federer and those of Carlos Alcaraz to those of Rafael

Nadal, their salience weights are quite different. This reflects the fact that the identification of the salience weight shown in Section (2.6.1) is non-trivial.

The curvature parameter λ is on average 2.9 and ranges from 1.3 to 6.3. It is significantly different from unity for all players in the data, so that the conditional probability of winning points of all players abbeys the law of diminishing marginal returns. The slope and constant parameters of its constituents ($f(x)$ and $k(x)$) are also mostly significantly different from 0. They can be best interpreted using a graphical representation, as in Figure (4). This figure shows the skills parameters of the four GOATs through the plot of f (top panel) and k (bottom panel) of these players. The figure clearly indicates that f is decreasing and concave for these players (true for all players) while k is increasing and convex (for a few players, i.e., David Ferrer and Diego Schwartzman, for instance, k is slightly decreasing and concave). Interestingly, we clearly see that Pete Sampras has the most efficient serve (measured as the probability to win one-shot rallies) for serve percentages between 0 and about 80% where Roger Federer's serve becomes more efficient. However, in terms of winning-point percentages on multi-shot rallies, Figure (4) shows that Rafael Nadal is the dominating player at all serve percentages, with a relatively flat profile, and the profile of Pete Sampras lies at the other extreme.

Figure 4: Conditional probabilities of winning one-shot (top) and multi-shot (bottom) rallies of the four GOATs.



Note: Dots represent the optimal serve strategy of each player. Novak Djokovic is in black, Rafael Nadal in blue, Roger Federer in red and Pete Sampras in green.

Having estimated the salience weight for each player, we can now ask the natural question: what would have been the optimal service strategy if players were mere outcome-maximizers, that is, if their salience weight were zero and they paid no attention to process utility? In other words, this exercise quantifies how much players are willing to sacrifice in point-winning probability to enjoy a more appealing style of play.

The model outlined in this paper allows us to answer this question through a simple counterfactual exercise. Using the estimated parameters for each player, we set the salience weight to zero and compute the optimal serve strategy, along with the corresponding probabilities of winning points on first and second serves. Table (3) presents the differences in optimal serve strategies between the observed and counterfactual scenarios ($\delta = 0$) in columns 3 and 4. As expected, the change in serve percentage is negative for almost all players, meaning their strategies would be more aggressive under the counterfactual, which is especially true on the second serve. Setting the salience weight to zero increases the point-winning probability on a player's own serve by approximately 0.39 %-pt. While this change appears small, the cumulative nature of tennis scoring amplifies its impact.

To illustrate, we compute the probability of winning a set and a match (best-of-five). For this exercise, we assume that the opponent's probability of winning their own serve equals the player's observed probability under the estimated δ.[24] Columns 7 and 8 show the corresponding changes in set- and match-winning probabilities. For instance, although a player increases the probability of winning a point on their serve by only 0.39 %-pt on average, their probability of winning a best-of-five match increases by 2.42 %-pt.

Finally, we can estimate the probability of reaching each round of a Grand Slam and the expected prize money under both observed and counterfactual strategies. Using the 2025 US Open prize distribution, Table (3) shows that, on average, players forego approximately \$33,000, i.e., 13.5%, in expected price money per Grand Slam by optimizing for both outcome and process utility, with 50% of the players forgoing more than \$12,880, i.e., 5.5%. When interpreting these results, one should also bear in mind that there are four Grand Slam tournaments per year, and while smaller tournaments offer lower prize money, this amount accumulates across tournaments over a professional career.

5 Robustness checks

Our results rely crucially on Condition (2), which has two components: first, the elasticity of the marginal probability of winning one-shot points equals that of winning multi-shot points, and second, this elasticity is constant. We conduct two robustness checks by relaxing these assumptions.

In the first robustness check, we allow the elasticity of the marginal conditional probabilities to vary linearly with x, i.e., proportional to $x\lambda$ with $\lambda > 0$ (Condition (B.4), Online Appendix (B7)) rather than being constant. Under this specification of the model, we find that the estimated salience weights are slightly larger than in the baseline, indicating that our earlier computations of trade-offs are robust and, if anything, conservative (see Table (B.4), Online Appendix (B7)).

---

[24] Hence, the player has a 50% chance of winning a set and a match under the estimated δ.

Table 3: Counterfactual optimal service strategy when $\delta = 0$, by player and aggregate effects.

Player	Salience weight δ	$\Delta\% - st$ in	$\Delta\% - nd$ in	$\Delta\% - pt$ won	$\Delta\% - gm$ won	$\Delta\% - set$ won	$\Delta\% - mat$ won	Δ Prize ($\times$\$1000) won
Novak Djokovic	0.27	-1.06	-4.31	0.12	0.18	0.39	0.73	8.07
Rafael Nadal	0.21	-0.63	-2.36	0.04	0.06	0.13	0.24	2.56
Roger Federer	0.32	-1.07	-6.01	0.20	0.27	0.64	1.20	13.51
Pete Sampras	0.42	-2.25	-9.94	0.57	0.82	1.82	3.41	42.01
...								
Boris Becker	0.56	-2.53	-10.79	0.64	1.14	2.13	3.99	50.26
...								
Carlos Alcaraz	0.05	-0.21	-0.89	0.00	0.01	0.02	0.03	0.31
Jannik Sinner	0.06	-0.23	-1.61	0.01	0.02	0.04	0.07	0.80
...								
Ivo Karlovic	1.19	-4.69	-10.79	1.53	1.61	4.64	8.66	132.87
John Isner	0.79	-3.87	-10.18	1.20	1.33	3.67	6.86	97.63
Reilly Opelka	0.36	-2.24	-7.71	0.48	0.57	1.48	2.78	33.36
...								
David Ferrer	-0.01	0.04	0.21	0.00	0.00	0.00	0.00	0.01
Diego Schwartzman	-0.38	1.14	6.25	0.17	0.38	0.60	1.12	12.56
Mean	0.34	-1.46	-4.94	0.39	0.66	1.29	2.41	33.22
Std	0.70	1.52	5.48	0.47	0.77	1.53	2.85	46.76
Min	-3.86	-6.33	-16.26	0.00	0.00	0.00	0.00	0.00
Median	0.33	-1.17	-5.54	0.18	0.33	0.61	1.15	12.88
Max	2.79	1.40	11.63	2.44	4.61	8.15	15.02	304.06

Notes: Δ means the difference in the variable concerned under the counterfactual $\delta = 0$ and the observed situation. To derive Δ Prize, we compute the probability that the player reaches each of round of the US Open, first, when he has 50% chances of winning each match (best-of-five), and then when his chances are $50+\Delta\% - mat$. We use the prize money distribution of the 2025 US open for this computation.

The second robustness check relaxes the equality of elasticities across one-shot and multi-shot points. Specifically, we let the former be $\lambda-1$ and the latter $t\lambda-1$ with $t > 0$ (Condition (B.5), Online Appendix (B8)). Note that, for most players, the conditional probabilities of winning multi-shot rallies on first and second serve, $k(x_1)$ and $k(x_2)$, are closer to each other than for one-shot points, so that $k(x)$ is likely to be "flatter" than $f(x)$, which would obtain for $t < 1$. Nevertheless, we evaluate the model over a grid of t values from 0.5 to 2, so that the curvature of k ranges from half to twice that of f. This grid approach is necessary because the data do not provide sufficient degrees of freedom to estimate t jointly with the other parameters. Across this range, the estimated salience weights are remarkably stable as shown in Table (B.5) of Online Appendix (B8).

We conclude from these results that the estimates of salience weights presented above are very robust to departure from Condition (2).

To further challenge our findings, we consider two alternative explanations for the apparent preference for winning multi-shot rallies and resulting conservative serve strategies adopted by tennis players. First, one could argue that players value one-shot and multi-shot rallies differently because these rallies require different levels of effort. One-shot rallies by nature are less demanding and more energy-efficient, so if effort considerations were driving deviations from outcome-maximization, players would be expected to put more importance on winning one-shot rallies, conserving energy for later points on their opponents' serve. We actually find the opposite: players put more weight on winning multi-shot rallies on their own serve, suggesting that enjoyment, rather than energy considerations, drives these choices.

Second, one might argue that conservative second-serve strategies arise due to risk aversion and, in particular, aversion to double faults, rather than process utility. To examine this, we develop and estimate a model (in Online Appendix (B9)) where players have a disutility for double faults, rather than a utility for process, and derive the associated optimal serve strategies. Importantly, this alternative model does not rely on the distinction between the conditional probabilities of winning one-shot and multi-shot rallies, only on the sum of the two. Therefore, if double-fault aversion was the true mechanism, the estimated disutility parameter should be unrelated to the probability of winning one-shot rallies conditional on serve strategy and point-winning probabilities. However, our analysis shows that $\mathbb{E}\left[\gamma \times f_j | x_1, x_2, y_1, y_2\right]$ for $j = 1, 2$, is systematically positive and significant at 5%. This indicates that there is additional information in the conditional probability of winning one-shot rallies that is being forced into the parameter of double-fault aversion and supports the relevance of process utility in explaining strategic deviations from outcome-maximizing behavior.

6 Conclusion and discussion

In this paper, we first develop a nonparametric identification strategy based solely on optimality conditions that delivers model-free bounds on the salience weight. Under mild shape restrictions, these bounds imply that a large majority of players have a positive salience weight. We then propose a parametric approach to estimate each player's salience weight and to conduct counterfactual analyses. We show that professional tennis players in our sample are willing to trade off a lower point-winning probability on their serve (outcome utility), on average 0.4%-pt, in exchange for a higher probability of winning multi-shot rallies (process utility). Although these differences are small in probability terms, due to the rules of tennis, they translate into a substantial forgone expected prize money. Using the 2025 US Open prize distribution as an example, we find that, on average, players forgo $\$33,000$, i.e., 13.5%, in prize money per tournament (Grand Slam), with 50% of the players sacrificing $\$13,000$ or more, i.e., 5.5%, per Grand Slam. These results demonstrate that process utility—i.e., the enjoyment of winning multi-shot rallies—is a significant component of tennis players' utility. That players are willing to sacrifice substantial expected prize money aligns well with the psychological literature on flow and intrinsic motivation. Indeed, according to Csikszentmihalyi (1990), when individuals enter a state of flow, they become fully absorbed in the activity itself, such that the experience is intrinsically rewarding even at a great material cost. This interpretation is further consistent with Self-Determination Theory, which emphasizes that intrinsic motivation is fostered when individuals experience autonomy and competence in the activity itself (e.g., Ryan and Deci (2024)). In this light, multi-shot rallies may provide a richer environment for such experiences than one-shot outcomes, helping to rationalize the observed willingness to trade off performance for process enjoyment.

As discussed in the introduction, the tennis setting offers several advantages for identifying and estimating individual-specific salience weights while capturing a mechanism likely to be present in many other economic contexts, particularly in labor markets. A growing body of evidence on compensating differentials suggests that individuals are willing to make substantial monetary sacrifices to access intrinsically rewarding job attributes. Existing evidence, however, is typically based on either revealed preferences over discrete job choices or on experimentally elicited stated or incentivized choices. For instance, Stern (2004) uses multiple job offers received by PhD students to measure the preference for independent research and shows that scientists are willing to forgo approximately 19% of their wages to engage in independent research, while Mas and Pallais (2017) estimates willingness to pay for non-wage job amenities using incentivized discrete-choice experiments over alternative work arrangements, finding wage trade-offs of approximately 8% to 20% for working from home and having regular time schedules respectively. More broadly, similar trade-offs between material outcomes and process-related attributes are also well documented in consumer markets, suggesting that the type of preferences identified in this paper may represent a general feature of economic behavior across domains. Consistent with these findings, this paper shows that such trade-offs can be identified from revealed repeated, high-frequency, continuous choices in a competitive field setting, thereby providing a novel approach to measuring process utility from observed behavior.

Our results are also important for policy interventions. For instance, in the tennis context, one might be tempted to conclude that, from a coaching perspective, a possible policy intervention would be to encourage players to set aside their desire for enjoyment and adopt strategies more closely aligned with outcome maximization. However, while such an approach may be effective in certain points of a match, the long-run implications of systematically neglecting process utility may be detrimental to performance. As suggested by flow and self-determination theories, suppressing the need for enjoyment and undermining autonomy and competence during play can lead to boredom, anxiety, stress, or “controled motivation” which in turn may reduce performance and contribute to adverse long-term outcomes, including burnout, disengagement and withdrawal (e.g., Ryan and Deci (2024)).

References

Akerlof, G. A. and R. E. Kranton (2000): “Economics and Identity,” *Quarterly Journal of Economics*, 115, 715–753.

Ariely, D., E. Kamenica, and D. Prelec (2008): “Man’s search for meaning: The case of legos,” *Journal of Economic Behavior and Organization*, 67, 671–677.

Bandiera, O., I. Barankay, and I. Rasul (2005): “Social Preferences and the Response to Incentives: Evidence from Personnel Data,” *Quarterly Journal of Economics*, 120, 917–962.

Borghans, L. (1995): “Keuzeprobleem op Centre Court,” *Economisch-Statistische Berichten*, 80, 658–661, analysis of the Wimbledon final and choice behavior in tennis.

Bryan, G., S. J. Hwang, and A. Said (2011): “Getting Paid to Care: The Effects of Financial Incentives on Voluntary Contributions,” *Journal of Public Economics*, 95, 334–342.

Bénabou, R. and J. Tirole (2003): “Intrinsic and extrinsic motivation,” *Review of Economic Studies*, 70, 489–520.

Corgnet, B., A. Espín, and R. Hernán-González (2016): “The Effect of Rewards on Performance: Evidence from a Real Effort Task,” *Management Science*, 62, 876–894.

Csikszentmihalyi, M. (1975): *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*, San Francisco, CA: Jossey-Bass.

——— (1990): *Flow: The Psychology of Optimal Experience*, New York, NY: Harper & Row.

Deci, E. L., R. Koestner, and R. M. Ryan (1999): “A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation,” *Psychological Bulletin*, 125, 627–668.

Deci, E. L. and R. M. Ryan (2000): “The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior,” *Psychological Inquiry*, 11, 227–268.

Dragusanu, R., D. Giovannucci, and N. Nunn (2014): “Ethical Consumption and Fair Trade: A Review of Evidence from Experimental and Field Studies,” *Journal of Economic Perspectives*, 28, 217–236.

Falk, A., E. Fehr, and C. Zehnder (2013): “Test-System Design and Incentive Effects,” *American Economic Journal: Microeconomics*, 5, 29–50.

Falk, A. and M. Kosfeld (2006): “The Hidden Costs of Control,” *American Economic Review*, 96, 1611–1630.

Fitzpatrick, A. (2024): “Important Performance Characteristics at Wimbledon: Implications for Coaches,” *ITF Coaching & Sport Science Review*, 32, 14–17, loughborough University, UK; performance analysis in elite grass court tennis.

Frey, B. S., M. Benz, and A. Stutzer (2004): "Introducing procedural utility: Not only what, but also how matters," *Journal of Institutional and Theoretical Economics*, 160, 377–401.

Frey, B. S. and F. Oberholzer-Gee (1997): "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out," *American Economic Review*, 87, 746–755.

Frey, B. S. and A. Stutzer (2002): *Happiness and Economics: How the Economy and Institutions Affect Well-Being*, Princeton, NJ: Princeton University Press.

Gneezy, U. and A. Rustichini (2000a): "A Fine is a Price," *Journal of Legal Studies*, 29, 1–17.

——— (2000b): "Pay enough or don't pay at all," *Quarterly Journal of Economics*, 115, 791–810.

Klaassen, F. J. and J. R. Magnus (2001): "Are Points in Tennis Independent and Identically Distributed? Evidence From a Dynamic Binary Panel Data Model," *Journal of the American Statistical Association*, 96, 500–509.

——— (2009): "The efficiency of top agents: An analysis through service strategy in tennis," *Journal of Econometrics*, 148, 72–85.

Kube, S., M. A. Maréchal, and C. Puppe (2013): "Sticking to Your Job: Non-Monetary Incentives and the Employment Relationship," *Quarterly Journal of Economics*, 128, 1627–1671.

Larkin, I., L. Pierce, and F. Gino (2012): "The Psychological Costs of Pay-for-Performance: Implications for the Strategic Compensation of Individuals," *Strategic Management Journal*, 33, 1194–1214.

Lavetti, K. (2023): "Compensating Wage Differentials in Labor Markets: Empirical Challenges and Applications," *Journal of Economic Perspectives*, 37, 189–212.

Loewenstein, G. (1999): "Because it is there: The challenge of mountaineering... for utility theory," *Kyklos*, 52, 315–344.

Maertens, M. and J. Swinnen (2009): "The Willingness to Pay for Fair Trade Products: A Meta-Analysis," *World Development*, 37, 1802–1811.

Mas, A. (2025): "Non-wage amenities." in *Handbook of Labor Economics*, ed. by C. Dustmann and T. Lemieux, Elsevier, vol. 6 of *Handbook of Labor Economics*, chap. 5, 373–446.

Mas, A. and A. Pallais (2017): "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–3759.

Menestrel, M. L. (2001): "A Process Approach to the Utility for Gambling," *Theory and Decision*, 50, 249–262.

Norton, M. I., D. Mochon, and D. Ariely (2012): "The IKEA Effect: When Labor Leads to Love," *Journal of Consumer Psychology*, 22, 453–460.

O'Shannessy, C. (2019): "Coaching analysis from the experts: first four shots are the key," https://tennishead.net/coaching-analysis-from-the-experts-first-four-shots-are-the-key/, accessed: 2023-04-02.

Rosen, S. (1986): "The Theory of Equalizing Differences," *Handbook of Labor Economics*, 1, 641–692.

Ryan, R. M. and E. L. Deci (2024): "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," in *Encyclopedia of Quality of Life and Well-Being Research*, Springer.

Sen, A. K. (1997): "Maximization and the Act of Choice," *Econometrica*, 65, 745–779.

Stern, S. (2004): "Do Scientists Pay to Be Scientists?" *Management Science*, 50, 835–853.

A Appendix

A1 Derivatives of the map $\Lambda(\lambda)$

Lemma 4 *The map $\Lambda(\lambda)$ defined by Algorithm (1) is strictly increasing and strictly convex on $\mathbb{R}_0^+$ for all $1 > x_{2i} > x_{1i} > 0$.*

Proof. Recall that

$$\Lambda(\lambda) = \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda} (1 + \lambda(1 - x_{2i})) - 1$$

and $\Lambda(\lambda) > 0$ on $\mathbb{R}_0^+$, for all $1 > x_{2i} > x_{1i} > 0$.

One has

$$\begin{aligned}
\Lambda'(\lambda) &= \frac{\partial\left(\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1+\lambda(1-x_{2i}))-1\right)}{\partial\lambda} \\
&= \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\left(\ln\frac{x_{2i}}{x_{1i}}(1+\lambda(1-x_{2i}))+1-x_{2i}\right) \\
&= \ln\frac{x_{2i}}{x_{1i}}\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1+\lambda(1-x_{2i}))+\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1-x_{2i}) \\
&= \ln\frac{x_{2i}}{x_{1i}}\left[\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1+\lambda(1-x_{2i}))-1+1\right]+\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1-x_{2i}) \\
&= (\Lambda(\lambda)+1)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}
\end{aligned}$$

Since $\Lambda(\lambda) > 0$ on $\mathbb{R}_0^+$ and $1 \geq x_{2i} > x_{1i}$, we conclude that $\Lambda'(\lambda) > 0$ on $\mathbb{R}_0^+$.

The second derivative is

$$\begin{aligned}
\Lambda''(\lambda) &= \Lambda'(\lambda)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\ln\frac{x_{2i}}{x_{1i}} \\
&= \left(\Lambda'(\lambda)+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\right)\ln\frac{x_{2i}}{x_{1i}}.
\end{aligned}$$

It follows that since $\Lambda'(\lambda) > 0$ on $\mathbb{R}_0^+$ and $x_{2i} > x_{1i}$, one has $\Lambda''(\lambda) > 0$ on $\mathbb{R}_0^+$.

Hence the map $\Lambda(\lambda)$ is strictly increasing and strictly convex on $\mathbb{R}_0^+$. ■

A2 Proof Lemma (1).

Proof. From Condition (1), since $f(x)$ and $k(x)$ are continuous, so are $B(x)$, $C(x)$ and hence $A(x)$. From (b) and (c), an application of the Intermediate Value Theorem, indicates that there exists a $x_0 \in [x_1^*, x_2^*]$ so that $A(x_0) = 0$. Under the assumption that $y(x)$ is strictly decreasing and concave, i.e., $y'(x) < 0$ and $y''(x) \leq 0$, $A(x)$ must also be strictly concave, i.e., $A''(x) = 2y'(x) + xy''(x) < 0$, so that x_0 is unique.

Next, we want to show that $B(x_0) > 0$. Note that since $A(x) := B(x) + C(x)$, this condition is equivalent to $C(x_0) < 0$. From Condition (1), $f(x)$ is strictly decreasing and concave, i.e., $f'(x) < 0$ and $f''(x) \leq 0$, so that $C(x)$ must also be strictly concave. By the strict concavity of $C(x)$ we know that $C'(x_1^*) > C'(x)$ for all $x > x_1^*$ meaning that the graph of $C(x)$ on $x > x_1^*$ lies below its tangent in x_1^* and hence

$$C(x) < C(x_1^*) + C'(x_1^*)(x - x_1^*)$$

for $x > x_1^*$. Using condition (d) to replace $C'(x_1^*)$ in the above inequality yields $C(x_0) < C(x_1^*) - \frac{C(x_1^*)}{x_2^* - x_1^*}(x_0 - x_1^*) < 0$ and noting that $-\frac{C(x_1^*)}{x_2^* - x_1^*}$ is the slope of the line passing through the points $(x_1^*, C(x_1^*))$ and $(x_2^*, 0)$, it follows that $C(x_0) < 0$ and hence $B(x_0) = -C(x_0) > 0$.
■

A3 Proof Theorem (3).

Proof. The proof proceeds in 6 steps.

1. *Step 1: Definition of the map* Λ_i and its properties.

   The map $\Lambda_i : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, defined as

   $$\Lambda_i(\lambda) = \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1 + \lambda(1 - x_{2i})) - 1,$$

   is strictly increasing and convex on $\mathbb{R}_0^+$ from Lemma (4) in Appendix (A1).

   Its first and second derivatives read as

   $$\begin{aligned}\Lambda_i'(\lambda) &= (\Lambda_i(\lambda) + 1)\ln\frac{x_{2i}}{x_{1i}} + (1 - x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda} > 0,\\ \Lambda''(\lambda) &= \left(\Lambda'(\lambda) + (1 - x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\right)\ln\frac{x_{2i}}{x_{1i}} > 0.\end{aligned}$$

   Define the function $g_i(\lambda)$ as

   $$g_i(\lambda) = \Lambda_i(\lambda) - \lambda = \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1 + \lambda(1 - x_{2i})) - 1 - \lambda, \qquad \lambda > 0.$$

2. *Step 2: Continuity of* g_i.

   Since Λ_i is continuously twice differentiable on $\mathbb{R}_0^+$ so is $g_i(\lambda)$.

3. *Step 3: Boundary behavior.*

   For $\lambda \to 0$, $\Lambda_i(\lambda) \to 0$, since we have

   $$\begin{aligned}\lim_{\lambda\to 0^+} g_i(\lambda) &= \lim_{\lambda\to 0^+}[\Lambda_i(\lambda) - \lambda]\\ &= \lim_{\lambda\to 0^+}\Lambda_i(\lambda)\\ &= \lim_{\lambda\to 0^+}\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1 + \lambda(1 - x_{2i})) - 1\\ &= 0.\end{aligned}$$

In contrast, as $\lambda \to \infty$, $g_i(\lambda) \to +\infty$ since

$$\begin{aligned}\lim_{\lambda\to\infty} g_i(\lambda) &= \lim_{\lambda\to\infty}\left[\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}(1+\lambda(1-x_{2i}))-1-\lambda\right]\\ &= (1-x_{2i})\lim_{\lambda\to\infty}\lambda\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\\ &= +\infty,\end{aligned}$$

using (A1).

4. *Step 4: First and second derivatives of* $g_i(\lambda)$*.*

   By definition

$$\begin{aligned}g_i'(\lambda) &= \Lambda_i'(\lambda)-1\\ &= (\Lambda_i(\lambda)+1)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}-1.\end{aligned}$$

   Hence,

$$\begin{aligned}\lim_{\lambda\to 0^+} g_i'(\lambda) &= \lim_{\lambda\to 0^+}\Lambda_i'(\lambda)-1\\ &= \lim_{\lambda\to 0^+}\left[(\Lambda_i(\lambda)+1)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\right]-1\\ &= \left[\left(\lim_{\lambda\to 0^+}\Lambda_i(\lambda)+1\right)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\lim_{\lambda\to 0^+}\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\right]-1\\ &= \ln\frac{x_{2i}}{x_{1i}}-x_{2i}\end{aligned}$$

   This means that for $\ln\frac{x_{2i}}{x_{1i}} < x_{2i}$, $\lim_{\lambda\to 0^+} g_i'(\lambda) < 0$.

   One also has

$$\begin{aligned}\lim_{\lambda\to +\infty} g'(\lambda) &= \lim_{\lambda\to +\infty}\Lambda_i'(\lambda)-1\\ &= \lim_{\lambda\to +\infty}\left[(\Lambda_i(\lambda)+1)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\right]\\ &= \left(\lim_{\lambda\to +\infty}\Lambda_i(\lambda)+1\right)\ln\frac{x_{2i}}{x_{1i}}+(1-x_{2i})\lim_{\lambda\to +\infty}\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\\ &= \ln\frac{x_{2i}}{x_{1i}}\lim_{\lambda\to +\infty}\Lambda_i(\lambda)+(1-x_{2i})\lim_{\lambda\to +\infty}\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\\ &= +\infty.\end{aligned}$$

   Finally, note that

$$g_i''(\lambda) = \Lambda_i''(\lambda) > 0.$$

5. *Step 5:* $g_i\left(1\right)$.

   Note that

$$\begin{aligned} g_i\left(1\right) &= \Lambda_i(1) - 1 \\ &= \left(\frac{x_{2i}}{x_{1i}}\right)\left(1 + \left(1 - x_{2i}\right)\right) - 2. \end{aligned}$$

   Hence, $g_i\left(1\right) < 0$ is implied by (A3) as indeed

$$\begin{aligned} g_i\left(1\right) < 0& \\ &\Leftrightarrow \\ x_{2i}\left(2 - x_{2i}\right) &< 2x_{1i} \\ &\Leftrightarrow \\ 2\left(x_{2i} - x_{1i}\right) &< \left(x_{2i}\right)^2. \end{aligned}$$

6. *Step 6: conclusion.*

   The function $g_i\left(\lambda\right)$ tends to 0 when $\lambda \to 0$, is first decreasing for values of λ close to 0 under (A2), i.e., $\ln \frac{x_{2i}}{x_{1i}} < x_{2i}$, hence becomes negative as λ increases away from 0, but tends to $+\infty$ as $\lambda \to \infty$. Since g_i is continuous, this means that $g_i\left(\lambda\right)$ must have at least one root on $\mathbb{R}_0^+$. However, since it is strictly convex on $\mathbb{R}_0^+$, there can be at most one root.

   So, under (A1)-(A2), for all i, the associated map $\Lambda_i(\lambda)$ has exactly one solution $\lambda_i^* = \Lambda_i(\lambda_i^*)$. Moreover, under (A3), this solution λ_i^* is strictly greater than unity.

■

B Online Appendix

B1 Players' Testimonies

Professional tennis players frequently emphasize enjoyment, intrinsic motivation, and the pursuit of flow as key drivers of their performance. The following quotes illustrate that top athletes prioritize the process of play alongside outcome maximization:

> "*I just want to step on court, try to do my things, follow my goals, and try to enjoy as much as I can.*" — Carlos Alcaraz, before his 2022 US Open semifinal vs. Novak Djokovic.
> https://www.atptour.com/en/news/alcaraz-lehecka-us-open-2025-qf
>
> "*I love the winning, I can take the losing, but most of all I love to play.* " — Boris Becker https://www.allgreatquotes.com/quote-56608
>
> "*My personal goal is to have fun and enjoy the moment, not put too much pressure on myself.*" — Venus Williams, interview by Rory Carroll, 2025.
> https://www.reuters.com/sports/tennis/venus-williams-prioritising-fun-she-returns-after-16-month-absence-2025-07-20/
>
> "*Every match that I've played since the beginning, I was just trying to win every point, [...] but the most important thing is to just have fun and enjoy.*" — Mirra Andreeva, Tennis.com, Wimbledon 2025.
> https://www.tennis.com/news/articles/carlos-alcaraz-joy-wimbledon-final
>
> "*[...] I spend every single moment enjoying the play, enjoying the fact that you're competing and just make it a fun game [...] every decision I went for felt like it was absolutely the right decision at the right time. It's what I like to call flow.*" — Stefanos Tsitsipas, Interview by James Richardson, 2023.
> https://www.tennis365.com/tennis-news/stefanos-tsitsipas-reveals-what-means-flow
>
> "*I felt like I was moving in a flow.*" — Iga Swiatek, post-match interview, Australian Open 2023.
> https://www.tntsports.co.uk/tennis/australian-open/2023/i-felt-like-i-was-moving-in-a-flow-iga-swiateks-post-match-interview-after-latest-win-at-australian-open_vid1815515/video.sht
>
> "*I go out there because I love tennis and I love playing.*" — Bethanie Mattek-Sands.
> https://www.edwardssports.co.uk/news/post/practice-tennis-like-a-pro
>
> "*The glory is being happy. The glory is not winning here or winning there. The glory is enjoying practicing, enjoy every day, enjoying to work hard, trying to be a better player than before.*" — Rafael Nadal, Sports Illustrated (interview by Jon Wertheim), 2010.
> https://www.si.com/more-sports/2010/07/16/nadal-interview
>
> "*I play tennis for the love of the game, having fun and what pleasure the game affords me.*" — Evonne Goolagong.
> https://www.tennis-prose.com/articles/scoop/some-fantastic-tennis-quotes

"*What separates good tennis players from the truly great? Their ability to enter flow state.*"
https://www.streetfamilytennis.com/flow-state-in-tennis-mastering-the-mental-game

Taken together, these testimonies highlight a consistent theme: elite athletes derive substantial intrinsic utility from the act of playing itself, aligning closely with the concept of *autotelic experience* in flow theory.

B2 Relation to Klaassen and Magnus (2009)

In the Klaassen and Magnus (2009) (KM from now on) model, the probability of a serve being in is denoted x. This probability reflects the choice of the server. As depicted in the above observations, if the server wants to take more risk, he will choose a lower value of x, hence a lower serve percentage. In contrast, if he wants to take fewer risks, he will choose a higher value of x, a higher serve percentage. Since in tennis, players can serve a second serve if the first is out, the server's strategy consists, in fact, of two numbers x_1 and x_2 reflecting respectively the probability of the first and second serve to be in. A player's probability to win a point on his serve, denoted $p\left(x_1, x_2\right)$, depends on his strategy $\left(x_1, x_2\right)$. Denote $y\left(x\right)$ the probability of winning a point conditional on the probability of the serve being in, x. $y\left(x\right)$ reflects the skills of the player, encompassing both his serving and rally skills. $y\left(x\right)$ is assumed to be twice differentiable and in particular to be strictly decreasing, i.e., $y'\left(x\right) < 0$, so that the easier the serve, that is, the higher the probability that it is in, the lower the probability to win the point conditional on the serve being in and concave, $y''\left(x\right) < 0$.

With these definitions, the probability to win a point on one's own serve, the outcome utility, reads as,

$$\begin{aligned} p\left(x_1, x_2\right) &= x_1 y\left(x_1\right) + \left(1 - x_1\right) x_2 y\left(x_2\right) \\ &= w\left(x_1\right) + \left(1 - x_1\right) w\left(x_2\right), \end{aligned}$$

where $w\left(x\right) := xy\left(x\right)$ is the unconditional probability to win a point when the serve is in.

A server aiming at maximizing his probability to win a point on his serve then does

$$\max_{x_1, x_2} p\left(x_1, x_2\right).$$

With mild regular conditions on the conditional probability $y\left(x\right)$, KM show that a unique solution exists.

The FOCs of this problem read as

$$\begin{aligned} \frac{\partial p\left(x_1, x_2\right)}{\partial x_1} &= w'\left(x_1\right) - w\left(x_2\right) = 0, \\ \frac{\partial p\left(x_1, x_2\right)}{\partial x_2} &= \left(1 - x_1\right) w'\left(x_2\right) = 0. \end{aligned}$$

Denote the unique optimal solution $\left(x_1^*, x_2^*\right)$. KM have shown that this solution is so that

$$\begin{aligned} x_1^* &< x_2^*, \\ y\left(x_1^*\right) &> y\left(x_2^*\right), \\ w\left(x_1^*\right) &< w\left(x_2^*\right). \end{aligned}$$

For further analysis, a parametric functional form must be adopted for the conditional probability $y(x)$. A particularily attractive shape is the power function, which, with only 3 parameters offers a very good trade-off between tractability and flexibility[25]

$$y(x) = \frac{a - x^{\lambda}}{\tau}$$

where the regularity conditions are met when $1 \leq a \leq \tau + x_0^{\lambda} < \tau + x^{*\lambda}$.

The parameters a, τ and λ are skills parameters in the sense that conditional on the server's decision x, they determine the probability of winning a point on own serve. Interestingly, if the curve $y(x)$ was linear, i.e., $\lambda = 1$, it would suffice to observe two points on the curve to identify the parameters (τ, α), i.e., slope and constant. Hence, data for the first and second serve would be enough to identify the skills parameters. However, the presence of the curvature parameter requires at least a third point or an additional restriction for its identification.

To understand the identification strategy of KM, let us first assume that players differ in their skills and hence in their conditional probability to win a point and if i indexes players, then (a_i, τ_i, λ_i) are the skills parameters of player i. The KM approach to identify the parameters (a_i, τ_i, λ_i) for all players $i = 1, .., N$, using data on $(x_{1i}, x_{2i}, y_{1i}, y_{2i})$ for $i = 1, ..., N$, is the following. First, note that the data for each player offer two points on the curve $y_i(x)$, namely (x_{1i}, y_{1i}) and (x_{2i}, y_{2i}). This means that given a curvature parameter λ_i, one can, for each player i, identify the slope τ_i and constant a_i. Noting this, KM therefore first randomly draw, for each player, a curvature parameter λ_i from a Gaussian distribution with mean μ_{λ} and standard deviation σ_{λ} and then use the two data points (x_{1i}, y_{1i}) and (x_{2i}, y_{2i}) for each player i, to compute (α_i, τ_i) given λ_i. Given the draw of λ_i, the shape of $y_i(x)$ is fully determined and the optimal solution (x_{1i}^*, x_{2i}^*) can be derived from the FOCs. This yields two additional points (x_{1i}^*, y_{1i}^*) and (x_{2i}^*, y_{2i}^*) that can be used to pin down the (distribution of the) curvature parameters λ_i, $i = 1, ..., N$. The KM strategy consists in picking the value of the mean and standard deviation $(\mu_{\lambda}, \sigma_{\lambda})$ that maximizes average efficiency across players and whose definition is given by

$$\begin{aligned} eff &= E\left(\frac{p(x_1, x_2)}{p(x_1^*, x_2^*)}\right) \\ &\simeq \frac{1}{n}\sum_{i=1}^{n} \frac{x_{1i}y_{1i} + (1 - x_{1i})\, x_{2i}y_{2i}}{x_{1i}^* y_{1i}^* + (1 - x_{1i}^*)\, x_{2i}^* y_{2i}^*}. \end{aligned}$$

With respect to our project that are 4 important findings from their analysis:

1. λ is virtually the same across (male) players as σ_{λ} is not statistically different from 0, the mean is $\mu_{\lambda} = 3.07$.
2. players are not perfect optimizers but close enough (the average/median player has an inefficiency of 1.1%, $eff = 0.989$),

---

[25] Power functions appear to fit well sports performance data in general. The relationship between world record time and distance in running, speed skating and swimming for instance is best fit with a power function.

3. there is more inefficiency on the second serve than on the first and,
4. better players are better optimizers.

B3 Deriving the first and second order conditions

B3.1 Theory

The first order condition, after substitution of the expression $\tilde{y}(x)$ yields

$$\begin{aligned}
\frac{\partial \tilde{w}(x_1)}{\partial x} - \tilde{w}(x_2) \quad &: \quad = \tilde{y}(x_1) + x_1 \frac{\partial \tilde{y}(x_1)}{\partial x} - x_2 \tilde{y}(x_2) \\
&= 0 \\
&\Leftrightarrow \\
\left[f(x_1) + \beta k(x_1)\right] + x_1 \left[\frac{\partial f(x_1)}{\partial x} + \beta \frac{\partial k(x_1)}{\partial x}\right] &= x_2 \tilde{y}(x_2)
\end{aligned}$$

while proceeding similarly for the second FOC yields

$$\begin{aligned}
\frac{\partial \tilde{w}(x_2)}{\partial x} \quad : \quad &= \frac{\partial x_2 \tilde{y}(x_2)}{\partial x} \\
&= \tilde{y}(x_2) + x_2 \frac{\partial \tilde{y}(x_2)}{\partial x} \\
&= \left[f(x_2) + \beta k(x_2)\right] + x_2 \left[\frac{\partial f(x_2)}{\partial x} + \beta \frac{\partial k(x_2)}{\partial x}\right] \\
&= 0.
\end{aligned}$$

To derive the Hessian we need the second order partial derivatives of the expected utility function. One has

$$\begin{aligned}
\frac{\partial^2 E(x_1, x_2)}{\partial x_1^2} &= \frac{\partial^2 \left(\tilde{w}(x_1) + (1 - x_1)\tilde{w}(x_2)\right)}{\partial x_1^2} \\
&= \frac{\partial \left(\tilde{y}(x_1) + x_1 \frac{\partial \tilde{y}(x_1)}{\partial x_1} - x_2 \tilde{y}(x_2)\right)}{\partial x_1} \\
&= 2\frac{\partial \tilde{y}(x_1)}{\partial x_1} + x_1 \frac{\partial^2 \tilde{y}(x_1)}{\partial x_1^2} \\
&= 2\left(\frac{\partial f(x_1)}{\partial x_1} + \beta \frac{\partial k(x_1)}{\partial x_1}\right) + x_1 \left(\frac{\partial^2 f(x_1)}{\partial x_1^2} + \beta \frac{\partial^2 k(x_1)}{\partial x_1^2}\right)
\end{aligned}$$

and,

$$\begin{aligned}
\frac{\partial^2 E\left(x_1, x_2\right)}{\partial x_1 \partial x_2} &= \frac{\partial^2\left(x_1 \tilde{y}\left(x_1\right)+\left(1-x_1\right) x_2 \tilde{y}\left(x_2\right)\right)}{\partial x_1 \partial x_2} \\
&= \frac{\partial\left(\left(1-x_1\right)\left(\tilde{y}\left(x_2\right)+x_2 \frac{\partial \tilde{y}\left(x_2\right)}{\partial x_2}\right)\right)}{\partial x_1} \\
&= -\left(\tilde{y}\left(x_2\right)+x_2 \frac{\partial \tilde{y}\left(x_2\right)}{\partial x_2}\right) \\
&= -\left(f\left(x_2\right)+\beta k\left(x_2\right)+x_2\left(\frac{\partial f\left(x_2\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2\right)}{\partial x_2}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 E\left(x_1, x_2\right)}{\partial x_2^2} &= \frac{\partial\left(\left(1-x_1\right)\left(\tilde{y}\left(x_2\right)+x_2 \frac{\partial \tilde{y}\left(x_2\right)}{\partial x_2}\right)\right)}{\partial x_2} \\
&= \left(1-x_1\right)\left(2 \frac{\partial \tilde{y}\left(x_2\right)}{\partial x_2}+x_2 \frac{\partial^2 \tilde{y}\left(x_2\right)}{\partial x_2^2}\right) \\
&= \left(1-x_1\right)\left(2\left(\frac{\partial f\left(x_2\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2\right)}{\partial x_2}\right)+x_2\left(\frac{\partial^2 f\left(x_2\right)}{\partial x_2^2}+\beta \frac{\partial^2 k\left(x_2\right)}{\partial x_2^2}\right)\right).
\end{aligned}$$

Collecting terms, this yields

$$\nabla^2 E\left(x_1, x_2\right)=\left(\begin{array}{cc}
\begin{array}{c} 2\left(\frac{\partial f\left(x_1\right)}{\partial x_1}+\beta \frac{\partial k\left(x_1\right)}{\partial x_1}\right) \\ +x_1\left(\frac{\partial^2 f\left(x_1\right)}{\partial x_1^2}+\beta \frac{\partial^2 k\left(x_1\right)}{\partial x_1^2}\right) \end{array} & -\left(\begin{array}{c} f\left(x_2\right)+\beta k\left(x_2\right) \\ +x_2\left(\frac{\partial f\left(x_2\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2\right)}{\partial x_2}\right) \end{array}\right) \\
-\left(\begin{array}{c} f\left(x_2\right)+\beta k\left(x_2\right) \\ +x_2\left(\frac{\partial f\left(x_2\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2\right)}{\partial x_2}\right) \end{array}\right) & \left(1-x_1\right)\left(\begin{array}{c} 2\left(\frac{\partial f\left(x_2\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2\right)}{\partial x_2}\right) \\ +x_2\left(\frac{\partial^2 f\left(x_2\right)}{\partial x_2^2}+\beta \frac{\partial^2 k\left(x_2\right)}{\partial x_2^2}\right) \end{array}\right)
\end{array}\right)$$

However, note that, by the second FOC, one has $\frac{\partial^2 E\left(x_1, x_2^*\right)}{\partial x_1 \partial x_2}=0$ so that the Hessian at optimum second serve reads as

$$\nabla^2 E\left(x_1, x_2^*\right)=\left(\begin{array}{cc}
\begin{array}{c} 2\left(\frac{\partial f\left(x_1\right)}{\partial x_1}+\beta \frac{\partial k\left(x_1\right)}{\partial x_1}\right) \\ +x_1\left(\frac{\partial^2 f\left(x_1\right)}{\partial x_1^2}+\beta \frac{\partial^2 k\left(x_1\right)}{\partial x_1^2}\right) \end{array} & 0 \\
0 & \left(1-x_1\right)\left(\begin{array}{c} 2\left(\frac{\partial f\left(x_2^*\right)}{\partial x_2}+\beta \frac{\partial k\left(x_2^*\right)}{\partial x_2}\right) \\ +x_2^*\left(\frac{\partial^2 f\left(x_2^*\right)}{\partial x_2^2}+\beta \frac{\partial^2 k\left(x_2^*\right)}{\partial x_2^2}\right) \end{array}\right)
\end{array}\right).$$

B3.2 Parametric

Using the parametric assumptions of the paper, the FOC for second serve strategy reads as

$$
\begin{aligned}
\left[f\left(x_2\right)+\beta k\left(x_2\right)\right]+x_2\left[\frac{\partial f\left(x_2\right)}{\partial x}+\beta\frac{\partial k\left(x_2\right)}{\partial x}\right] &= 0 \\
&\Leftrightarrow \\
\left[\frac{a_f-x_2^\lambda}{\tau_f}+\beta\frac{a_k-x_2^\lambda}{\tau_k}\right]-\lambda\left[\frac{x_2^\lambda}{\tau_f}+\beta\frac{x_2^\lambda}{\tau_k}\right] &= 0 \\
&\Leftrightarrow \\
\left[\frac{a_f-x_2^\lambda}{\tau_f}+\beta\frac{a_k-x_2^\lambda}{\tau_k}\right] &= \lambda\left[\frac{x_2^\lambda}{\tau_f}+\beta\frac{x_2^\lambda}{\tau_k}\right] \\
&\Leftrightarrow \\
\frac{a_f}{\tau_f}+\beta\frac{a_k}{\tau_k}-\frac{x_2^\lambda}{\tau_f}-\beta\frac{x_2^\lambda}{\tau_k} &= \lambda\frac{x_2^\lambda}{\tau_f}+\lambda\beta\frac{x_2^\lambda}{\tau_k} \\
&\Leftrightarrow \\
\left(\lambda\frac{\tau_k}{\tau_f\tau_k}+\frac{\tau_k}{\tau_f\tau_k}+\lambda\beta\frac{\tau_f}{\tau_f\tau_k}+\beta\frac{\tau_f}{\tau_f\tau_k}\right)x_2^\lambda &= \frac{a_f\tau_k}{\tau_f\tau_k}+\beta\frac{a_k\tau_f}{\tau_f\tau_k} \\
&\Leftrightarrow \\
\left(\lambda\tau_k+\tau_k+\lambda\beta\tau_f+\beta\tau_f\right)x_2^\lambda &= a_f\tau_k+\beta a_k\tau_f \\
&\Leftrightarrow \\
x_2^{*\lambda} &= \frac{1}{\lambda+1}\frac{a_f\tau_k+\beta a_k\tau_f}{\tau_k+\beta\tau_f} \\
&\Leftrightarrow \\
x_2^* &= \left(\frac{1}{\lambda+1}\frac{a_f\tau_k+\beta a_k\tau_f}{\tau_k+\beta\tau_f}\right)^{\frac{1}{\lambda}}
\end{aligned}
$$

The expression for the FOC for the first serve optimal strategy reads as

$$
\begin{aligned}
\left[f\left(x_1\right)+\beta k\left(x_1\right)\right]+x_1\left[\frac{\partial f\left(x_1\right)}{\partial x}+\beta\frac{\partial k\left(x_1\right)}{\partial x}\right] &= x_2\tilde{y}\left(x_2\right)\\
&\Leftrightarrow\\
\left[f\left(x_1\right)+\beta k\left(x_1\right)\right]+x_1\left[\frac{\partial f\left(x_1\right)}{\partial x}+\beta\frac{\partial k\left(x_1\right)}{\partial x}\right] &= x_2\left(f\left(x_2\right)+\beta k\left(x_2\right)\right)\\
&\Leftrightarrow\\
\left[\frac{a_f-x_1^\lambda}{\tau_f}+\beta\frac{a_k-x_1^\lambda}{\tau_k}\right]-\lambda\left[\frac{x_1^\lambda}{\tau_f}+\beta\frac{x_1^\lambda}{\tau_k}\right] &= x_2\left(\frac{a_f-x_2^\lambda}{\tau_f}+\beta\frac{a_k-x_2^\lambda}{\tau_k}\right)\\
&\Leftrightarrow\\
\left[\frac{a_f}{\tau_f}-\frac{x_1^\lambda}{\tau_f}+\beta\frac{a_k}{\tau_k}-\beta\frac{x_1^\lambda}{\tau_k}\right]-\left[\lambda\frac{x_1^\lambda}{\tau_f}+\lambda\beta\frac{x_1^\lambda}{\tau_k}\right] &= x_2\left(\frac{a_f-x_2^\lambda}{\tau_f}+\beta\frac{a_k-x_2^\lambda}{\tau_k}\right)\\
&\Leftrightarrow\\
\left[\frac{a_f\tau_k}{\tau_f\tau_k}-\frac{\tau_k x_1^\lambda}{\tau_f\tau_k}+\beta\frac{a_k\tau_f}{\tau_f\tau_k}-\beta\frac{\tau_f x_1^\lambda}{\tau_f\tau_k}\right]-\left[\lambda\frac{\tau_k x_1^\lambda}{\tau_f\tau_k}+\lambda\beta\frac{\tau_f x_1^\lambda}{\tau_f\tau_k}\right] &= x_2\left(\frac{a_f\tau_k-\tau_k x_2^\lambda}{\tau_f\tau_k}+\beta\frac{a_k\tau_f-\tau_f x_2^\lambda}{\tau_f\tau_k}\right)\\
&\Leftrightarrow\\
\left(\tau_k+\beta\tau_f+\lambda\tau_k+\lambda\beta\tau_f\right)x_1^\lambda &= a_f\tau_k+\beta a_k\tau_f\\
&\quad -x_2\left(a_f\tau_k-\tau_k x_2^\lambda+\beta a_k\tau_f-\beta\tau_f x_2^\lambda\right)\\
&\Leftrightarrow\\
\frac{a_f\tau_k+\beta a_k\tau_f-x_2\left(a_f\tau_k+\beta a_k\tau_f-\left(\tau_k+\beta\tau_f\right)x_2^\lambda\right)}{\left(1+\lambda\right)\left(\tau_k+\beta\tau_f\right)} &= x_1^\lambda\\
&\Leftrightarrow\\
\left(\frac{a_f\tau_k+\beta a_k\tau_f-x_2\left(a_f\tau_k+\beta a_k\tau_f-\left(\tau_k+\beta\tau_f\right)x_2^\lambda\right)}{\left(1+\lambda\right)\left(\tau_k+\beta\tau_f\right)}\right)^{\frac{1}{\lambda}} &= x_1
\end{aligned}
$$

The expression for the Hessian of the expected utility at optimum reads as

$$
\nabla^2 E\left(x_1^*,x_2^*\right)=\begin{pmatrix}-\lambda\left(1+\lambda\right)\left(x_1^*\right)^{\lambda-1}\left(\frac{1}{\tau_f}+\frac{\beta}{\tau_k}\right) & 0\\ 0 & -\left(1-x_1\right)\lambda\left(1+\lambda\right)\left(x_2^*\right)^{\lambda-1}\left(\frac{1}{\tau_f}+\frac{\beta}{\tau_k}\right)\end{pmatrix}
$$

Interestingly, as shown below, the optimal service strategy of the player only depends on the relative slope $\frac{\tau_f}{\tau_k}$ as indeed both expressions rewrite as

$$
\begin{aligned}
x_2^{*\lambda} &= \frac{a_f+\beta a_k\frac{\tau_f}{\tau_k}}{\left(1+\lambda\right)\left(1+\beta\frac{\tau_f}{\tau_k}\right)}\\
x_1^{*\lambda} &= \frac{a_f+\beta a_k\frac{\tau_f}{\tau_k}-x_2^*\left(a_f+\beta a_k\frac{\tau_f}{\tau_k}-\left(1+\beta\frac{\tau_f}{\tau_k}\right)x_2^{*\lambda}\right)}{\left(1+\lambda\right)\left(1+\beta\frac{\tau_f}{\tau_k}\right)}
\end{aligned}
$$

and,

$$\begin{aligned} x_2^{*\lambda} &= \frac{a_f\tau_k+\beta a_k\tau_f}{(1+\lambda)(\tau_k+\beta\tau_f)} \\ &= \frac{\frac{1}{\tau_k}}{\frac{1}{\tau_k}}\frac{a_f\tau_k+\beta a_k\tau_f}{(1+\lambda)(\tau_k+\beta\tau_f)} \\ &= \frac{a_f+\beta a_k\frac{\tau_f}{\tau_k}}{(1+\lambda)\left(1+\beta\frac{\tau_f}{\tau_k}\right)} \end{aligned}$$

and for $x_1^{*\lambda}$:

$$\begin{aligned} x_1^{*\lambda} &= \frac{a_f\tau_k+\beta a_k\tau_f-x_2^*\left(a_f\tau_k+\beta a_k\tau_f-(\tau_k+\beta\tau_f)\,x_2^{*\lambda}\right)}{(1+\lambda)(\tau_k+\beta\tau_f)} \\ &= \frac{\frac{1}{\tau_k}}{\frac{1}{\tau_k}}\frac{a_f\tau_k+\beta a_k\tau_f-x_2^*\left(a_f\tau_k+\beta a_k\tau_f-(\tau_k+\beta\tau_f)\,x_2^{*\lambda}\right)}{(1+\lambda)(\tau_k+\beta\tau_f)} \\ &= \frac{a_f+\beta a_k\frac{\tau_f}{\tau_k}-x_2^*\left(a_f+\beta a_k\frac{\tau_f}{\tau_k}-\left(1+\beta\frac{\tau_f}{\tau_k}\right)x_2^{*\lambda}\right)}{(1+\lambda)\left(1+\beta\frac{\tau_f}{\tau_k}\right)} \end{aligned}$$

B4 Sum of power functions with same curvature

We aim to show that the sum of two power functions with same curvature parameter λ, yield a power function with curvature parameter λ. Let $f(x)=\frac{a_f-x^\lambda}{\tau_f}$ and $k(x)=\frac{a_k-x^\lambda}{\tau_k}$. Then one has

$$\begin{aligned} f(x)+k(x) &= \frac{a_f-x^\lambda}{\tau_f}+\frac{a_k-x^\lambda}{\tau_k} \\ &= \frac{a_f}{\tau_f}+\frac{a_k}{\tau_k}-\left(\frac{1}{\tau_f}+\frac{1}{\tau_k}\right)x^\lambda \\ &= \frac{a_f\tau_k+a_k\tau_f}{\tau_f\tau_k}-\frac{\tau_f+\tau_k}{\tau_f\tau_k}x^\lambda \\ &= \frac{\frac{a_f\tau_k+a_k\tau_f}{\tau_f+\tau_k}}{\frac{\tau_f\tau_k}{\tau_f+\tau_k}}-\frac{1}{\frac{\tau_f\tau_k}{\tau_f+\tau_k}}x^\lambda \end{aligned}$$

so that letting $\tau=\frac{\tau_f\tau_k}{\tau_f+\tau_k}$ and $a=\frac{a_f\tau_k+a_k\tau_f}{\tau_f+\tau_k}$ one indeed has

$$\begin{aligned} y(x) &= f(x)+k(x) \\ &= \frac{a_f-x^\lambda}{\tau_f}+\frac{a_k-x^\lambda}{\tau_k} \\ &= \frac{a-x^\lambda}{\tau} \end{aligned}$$

and

$$\begin{aligned}\frac{\partial y(x)}{\partial x} &= \frac{\partial f(x)}{\partial x}+\frac{\partial k(x)}{\partial x}\\ &= -\lambda\left(\frac{x^{\lambda-1}}{\tau_f}+\frac{x^{\lambda-1}}{\tau_k}\right)\\ &= -\lambda x^{\lambda-1}\left(\frac{\tau_k+\tau_j}{\tau_f\tau_k}\right)\\ &= -\lambda\frac{x^{\lambda-1}}{\tau}\end{aligned}$$

with in particular

$$x\frac{\partial y(x)}{\partial x}=-\lambda\frac{x^{\lambda}}{\tau}.$$

B5 Derivation of expression of $\lambda^{(t+1)}$

Note first that the slopes, constants and preference parameter have the following expressions in terms of λ and z_{1i} and z_{2i} where we recall that

$$\begin{aligned}z_{1i} &= (x_{1i})^{\lambda},\\ z_{2i} &= (x_{2i})^{\lambda}.\end{aligned}$$

The slopes parameters read as

$$\begin{aligned}\tau_{fi} &= -\frac{z_{1i}-z_{2i}}{f_{1i}-f_{2i}},\\ \tau_{ki} &= -\frac{z_{1i}-z_{2i}}{k_{1i}-k_{2i}}.\end{aligned}$$

with a relative slope independent of λ

$$\frac{\tau_{ki}}{\tau_{fi}}=\frac{f_{1i}-f_{2i}}{k_{1i}-k_{2i}}$$

The constants read as

$$\begin{aligned}a_{fi} &= -\frac{z_{1i}-z_{2i}}{f_{1i}-f_{2i}}f_{1i}+z_{1i},\\ a_{ki} &= -\frac{z_{1i}-z_{2i}}{k_{1i}-k_{2i}}k_{1i}+z_{1i},\end{aligned}$$

and the attention parameter as

$$\beta_i=-\frac{\tau_{ki}}{\tau_{fi}}\frac{a_{fi}-z_{2i}(1+\lambda)}{a_{ki}-z_{2i}(1+\lambda)}.$$

Substituting the epressions of the previous terms into β_i yields

$$\begin{aligned}
\beta_i &= -\frac{\tau_{ki}}{\tau_{fi}}\frac{\tau_{fi}f_{2i}+z_{2i}\left(1-1-\lambda\right)}{\tau_{ki}k_{2i}+z_{2i}\left(1-1-\lambda\right)} \\
&= -\frac{f_{1i}-f_{2i}}{k_{1i}-k_{2i}}\frac{\left(-\frac{z_{1i}-z_{2i}}{f_{1i}-f_{2i}}\right)f_{2i}-\lambda z_{2i}}{\left(-\frac{z_{1i}-z_{2i}}{k_{1i}-k_{2i}}\right)k_{2i}-\lambda z_{2i}} \\
&= -\frac{\left(z_{1i}-z_{2i}\right)f_{2i}+\lambda z_{2i}\left(f_{1i}-f_{2i}\right)}{\left(z_{1i}-z_{2i}\right)k_{2i}+\lambda z_{2i}\left(k_{1i}-k_{2i}\right)} \\
&= -\frac{z_{1i}f_{2i}-z_{2i}f_{2i}+\lambda z_{2i}f_{1i}-\lambda z_{2i}f_{2i}}{z_{1i}k_{2i}-z_{2i}k_{2i}+\lambda z_{2i}k_{1i}-\lambda z_{2i}k_{2i}} \\
&= -\frac{z_{1i}f_{2i}+\lambda z_{2i}f_{1i}-z_{2i}f_{2i}\left(1+\lambda\right)}{z_{1i}k_{2i}+\lambda z_{2i}k_{1i}-z_{2i}k_{2i}\left(1+\lambda\right)} \\
&= -\frac{f_{2i}}{k_{2i}}\frac{\Delta z_i+\lambda\Delta f_i}{\Delta z_i+\lambda\Delta k_i}
\end{aligned}$$

where $a_{fi}=\tau_{fi}f_{2i}+z_{2i}$, $a_{ki}=\tau_{ki}k_{2i}+z_{2i}$, $\frac{\tau_{fi}}{\tau_{ki}}=\frac{k_{1i}-k_{2i}}{f_{1i}-f_{2i}}$, $\tau_{fi}=-\frac{z_{1i}-z_{2i}}{f_{1i}-f_{2i}}$, $\tau_{ki}=-\frac{z_{1i}-z_{2i}}{k_{1i}-k_{2i}}$ and $\Delta l_i=\frac{l_{1i}-l_{2i}}{l_{2i}}$ $\forall l=x,f,k$.

Note also that rearranging the expression of β_i offers the following interesting expression that is used below

$$\beta_i\frac{\tau_{fi}}{\tau_{ki}}=-\frac{a_{fi}-z_{2i}\left(1+\lambda\right)}{a_{ki}-z_{2i}\left(1+\lambda\right)}$$

Recall now the expression of λ_i and rearranging yields

$$\begin{aligned}
\lambda_i &= \frac{a_{fi}+a_{ki}\beta_i\frac{\tau_{fi}}{\tau_{ki}}}{\left(1+\beta_i\frac{\tau_{fi}}{\tau_{ki}}\right)z_{1i}}\left(1-x_{2i}\right)+x_{2i}\frac{z_{2i}}{z_{1i}}-1 \\
&= \frac{a_{fi}+a_{ki}\beta_i\frac{\tau_{fi}}{\tau_{ki}}}{1+\beta_i\frac{\tau_{fi}}{\tau_{ki}}}\frac{1-x_{2i}}{z_{1i}}+x_{2i}\frac{z_{2i}}{z_{1i}}-1 \\
&= \Gamma\frac{1-x_{2i}}{z_{1i}}+x_{2i}\frac{z_{2i}}{z_{1i}}-1
\end{aligned}$$

where

$$\Gamma=\frac{a_{fi}+a_{ki}\beta_i\frac{\tau_{fi}}{\tau_{ki}}}{1+\beta_i\frac{\tau_{fi}}{\tau_{ki}}}$$

Let us now rewrite the term Γ by using the interesting expression derived from that of β_i

$$\beta_i\frac{\tau_{fi}}{\tau_{ki}}=-\frac{a_{fi}-z_{2i}\left(1+\lambda\right)}{a_{ki}-z_{2i}\left(1+\lambda\right)}$$

so that

$$\begin{aligned}
\Gamma &= \frac{a_{fi}\left(a_{ki}-z_{2i}\left(1+\lambda\right)\right)-a_{ki}\left(a_{fi}-z_{2i}\left(1+\lambda\right)\right)}{a_{ki}-z_{2i}\left(1+\lambda\right)-\left(a_{fi}-z_{2i}\left(1+\lambda\right)\right)} \\
&= \frac{a_{fi}a_{ki}-a_{fi}a_{ki}+a_{ki}z_{2i}\left(1+\lambda\right)-a_{fi}z_{2i}\left(1+\lambda\right)}{a_{ki}-a_{fi}+z_{2i}\left(1+\lambda\right)-z_{2i}\left(1+\lambda\right)} \\
&= \frac{a_{ki}-a_{fi}}{a_{ki}-a_{fi}}z_{2i}\left(1+\lambda\right) \\
&= z_{2i}\left(1+\lambda\right)
\end{aligned}$$

Substituting this into the expression of λ_i, it follows that

$$\begin{aligned}
\lambda_i &= \Gamma\frac{1-x_{2i}}{z_{1i}}+x_{2i}\frac{z_{2i}}{z_{1i}}-1 \\
&= z_{2i}\left(1+\lambda\right)\frac{1-x_{2i}}{z_{1i}}+x_{2i}\frac{z_{2i}}{z_{1i}}-1 \\
&= \frac{z_{2i}}{z_{1i}}\left(\left(1+\lambda\right)\left(1-x_{2i}\right)+x_{2i}\right)-1 \\
&= \frac{z_{2i}}{z_{1i}}\left(\left(1+\lambda-x_{2i}-x_{2i}\lambda\right)+x_{2i}\right)-1 \\
&= \frac{z_{2i}}{z_{1i}}\left(1+\lambda-x_{2i}\lambda\right)-1 \\
&= \frac{z_{2i}}{z_{1i}}\left(1+\lambda\left(1-x_{2i}\right)\right)-1
\end{aligned}$$

Hence the compact expression of λ_i as

$$\begin{aligned}
\lambda_i &= \Lambda\left(\lambda\right) \\
&= \frac{z_{2i}}{z_{1i}}\left(1+\lambda\left(1-x_{2i}\right)\right)-1 \\
&= \left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\left(1+\lambda\left(1-x_{2i}\right)\right)-1.
\end{aligned}$$

Note that $\Lambda\left(\lambda\right)>0$ since

$$\begin{aligned}
\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\left(1+\lambda\left(1-x_{2i}\right)\right)-1 &> 0 \\
&\Leftrightarrow \\
\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda}\left(1+\lambda\left(1-x_{2i}\right)\right) &> 1 \\
&\Leftrightarrow \\
\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda} &> \frac{1}{1+\lambda\left(1-x_{2i}\right)}
\end{aligned}$$

since $\left(\frac{x_{2i}}{x_{1i}}\right)^{\lambda} > 1$ for $x_{2i} > x_{1i}$ and $\frac{1}{1+\lambda(1-x_{2i})} < 1$ since

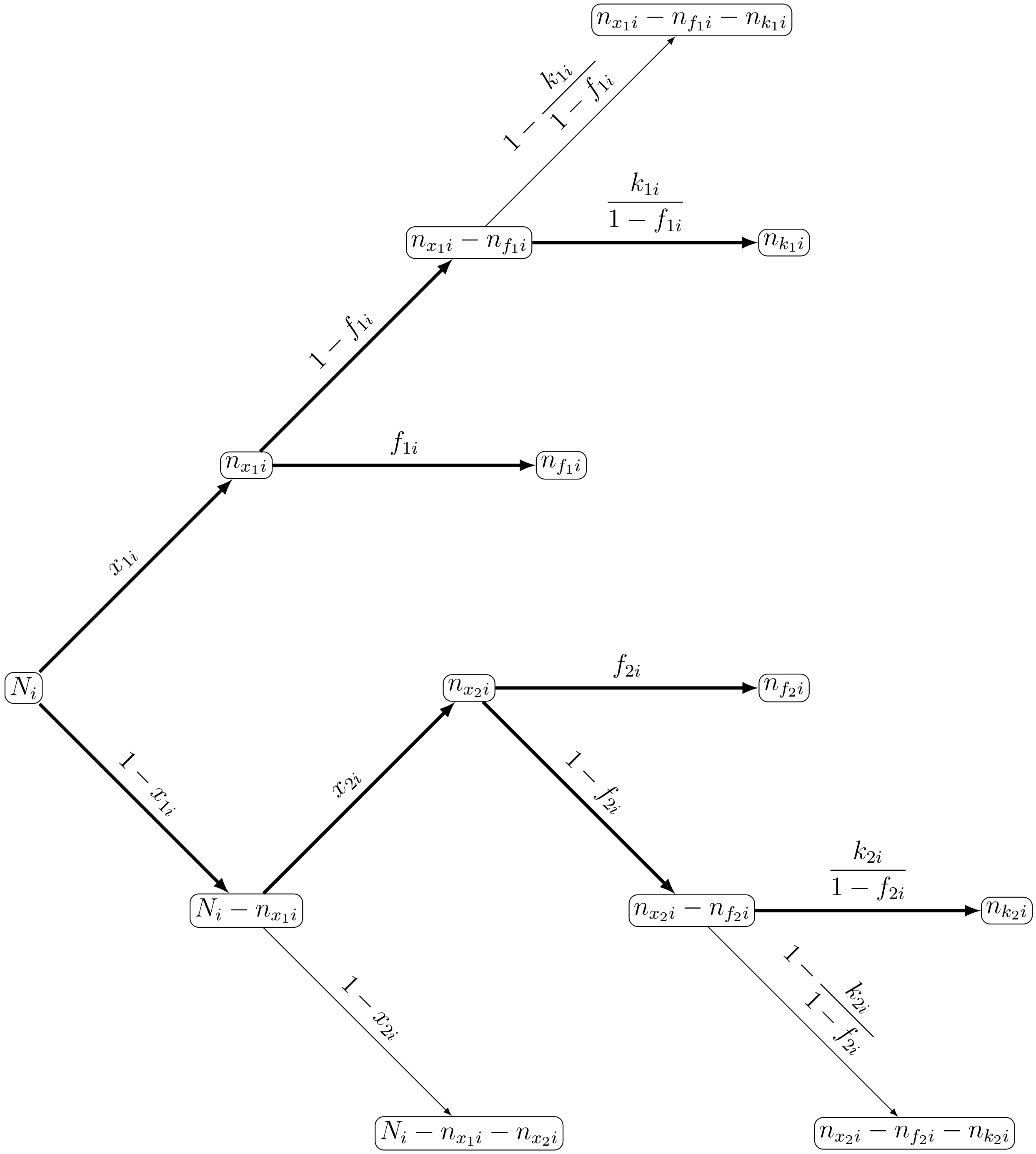
$$\begin{array}{rcl} \dfrac{1}{1+\lambda\left(1-x_{2i}\right)} & < & 1 \\ & \Leftrightarrow & \\ 1 & < & 1+\lambda\left(1-x_{2i}\right) \\ & \Leftrightarrow & \\ 0 & < & \lambda\left(1-x_{2i}\right) \end{array}$$

which obtains from $\lambda > 0$ and $0 < x_{2i} < 1$.

B6 Tree representation of the statistics for each player i.

In Figure (B.5) below, each node indicates the mass of points observed for player i, and each edge represents a probability of success or failure, conditional on having reached the previous node. At the staring node N_i, the two edges represent the probability that the first serve is in or out. If the service is in, we move to node $n_{x_1 i}$ where the two edges indicate the probability that the first serve is returned in or not, where the former occurs with probability f_{1i} and the latter with probability $1 - f_{1i}$. The terminal nodes that are accessed with horizontal edges, correspond to the masses of points won by the server, and the paths leading to them are highlighted in bold.

Figure B.5: Tree representation of tennis statistics for player i.



B7 Model with softmax parametrization

Suppose that we replace Condition (2) used to parametrize the elasticity of the marginal probabilities $f'(x)$ and $k'(x)$, by the following condition:

Condition B.4 $x\frac{f''(x)}{f'(x)} = x\frac{k''(x)}{k'(x)} = \lambda x > 0.$

This condition tells us that the elasticity of the marginal conditional probabilities f' and k' are not constant anymore but rather linear. In this case, the associated shape for the function f for instance is

$$f(x) = a_f + \tau_f \exp(\lambda x)$$

with

$$\begin{aligned} f'(x) &= \lambda \tau_f \exp(\lambda x) \\ f''(x) &= \lambda^2 \tau_f \exp(\lambda x) \end{aligned}$$

where $\tau_f = \frac{f_2 - f_1}{z_2 - z_1}$ and $a_f = f_1 - \tau_f z_1$ and $z_1 = \exp(\lambda x_1)$ and $z_2 = \exp(\lambda x_2)$.[26]

With the same logic for $k(x)$ and plugging these expressions into the solution for preference parameter β, one has

$$\begin{aligned} \beta &= -\frac{f_2 + x_2 f'(x_2)}{k_2 + x_2 k'(x_2)} \\ &= -\frac{f_2 + x_2 \lambda \tau_f \exp(\lambda x_2)}{k_2 + x_2 \lambda \tau_k \exp(\lambda x_2)} \\ &= -\frac{f_2 + x_2 \lambda \frac{f_2 - f_1}{z_2 - z_1} z_2}{k_2 + x_2 \lambda \frac{k_2 - k_1}{z_2 - z_1} z_2} \\ &= -\frac{f_2 (z_2 - z_1) + \lambda x_2 z_2 (f_2 - f_1)}{k_2 (z_2 - z_1) + \lambda x_2 z_2 (k_2 - k_1)}. \end{aligned}$$

Similarily, using the FOC for the first serve strategy one obtains an expression for the curvature parameter λ as follows.

Step 1:

Remember that the solution for the curvature parameter obtains from:

$$[f(x_1) + \beta k(x_1)] + x_1 [f'(x_1) + \beta k'(x_1)] = x_2 [f(x_2) + \beta k(x_2)]$$

which given the parametric assumptions become

$$[f_1 + \beta k_1] (\exp(\lambda (x_2 - x_1)) - 1) + \lambda x_1 [f_2 - f_1 + \beta (k_2 - k_1)] = x_2 [f_2 + \beta k_2] (\exp(\lambda (x_2 - x_1)) - 1).$$

---

[26] Indeed one has

$$\begin{aligned} x\frac{f''(x)}{f'(x)} &= x\frac{\lambda^2 \tau_f \exp(\lambda x)}{\lambda \tau_f \exp(\lambda x)} \\ &= x\lambda. \end{aligned}$$

Using the notation $z_j = \exp(\lambda x_j)$ this rewrite as

$$[f_1 + \beta k_1]\left(\frac{z_2}{z_1} - 1\right) + \lambda x_1 [f_2 - f_1 + \beta(k_2 - k_1)] = x_2 [f_2 + \beta k_2]\left(\frac{z_2}{z_1} - 1\right).$$

Note that this writes compactly as

$$A(E-1) + \lambda B = 0$$

where $E = \frac{z_2}{z_1}$, $A = f_1 + \beta k_1 - x_2 [f_2 + \beta k_2]$ and $B = x_1 [f_2 - f_1 + \beta(k_2 - k_1)]$.

Step 2: Note that we can rewrite the preference parameter more compactly too as

$$\begin{aligned} \beta &= -\frac{f_2(z_2 - z_1) + \lambda x_2 z_2 (f_2 - f_1)}{k_2(z_2 - z_1) + \lambda x_2 z_2 (k_2 - k_1)} \\ &= -\frac{N}{M} \end{aligned}$$

where $N = f_2 + \lambda(f_2 - f_1)Q$ and $M = k_2 + \lambda(k_2 - k_1)Q$.

Substituting this last expression of β into A and B of the compact expression of the FOC for x_1 yields

$$\begin{aligned} A &= f_1 - \frac{f_2 + \lambda(f_2 - f_1)Q}{k_2 + \lambda(k_2 - k_1)Q} k_1 - x_2\left[f_2 - \frac{f_2 + \lambda(f_2 - f_1)Q}{k_2 + \lambda(k_2 - k_1)Q} k_2\right] \\ B &= x_1\left[f_2 - f_1 - \frac{f_2 + \lambda(f_2 - f_1)Q}{k_2 + \lambda(k_2 - k_1)Q}(k_2 - k_1)\right], \end{aligned}$$

where $Q = \frac{x_2 E}{E-1}$ and which rewrites compactly as

$$\begin{aligned} A &= \frac{Mf_1 - Nk_1 - x_2[Mf_2 - Nk_2]}{M} \\ B &= x_1 \frac{M(f_2 - f_1) - N(k_2 - k_1)}{M}, \end{aligned}$$

Note that

$$\begin{aligned} M(f_2 - f_1) - N(k_2 - k_1) &= (k_2 + \lambda(k_2 - k_1)Q)(f_2 - f_1) - (f_2 + \lambda(f_2 - f_1)Q)(k_2 - k_1) \\ &= k_2(f_2 - f_1) - f_2(k_2 - k_1) \\ &= k_2 f_2 - k_2 f_1 - f_2 k_2 + f_2 k_1 \\ &= f_2 k_1 - k_2 f_1, \end{aligned}$$

so that

$$B = x_1 \frac{f_2 k_1 - k_2 f_1}{M}.$$

Note further that

$$A = \frac{M(f_1 - x_2 f_2) - N(k_1 - x_2 k_2)}{M},$$

and after substitution and simple rearranging, one has

$$\begin{aligned}
M\left(f_1 - x_2 f_2\right) &= k_2\left(f_1 - x_2 f_2\right) + \left(f_1 - x_2 f_2\right)\left(k_2 - k_1\right)\lambda Q, \\
N\left(k_1 - x_2 k_2\right) &= f_2\left(k_1 - x_2 k_2\right) + \left(k_1 - x_2 k_2\right)\left(f_2 - f_1\right)\lambda Q.
\end{aligned}$$

Hence, the numerator of A reads as

$$\begin{aligned}
M\left(f_1 - x_2 f_2\right) - N\left(k_1 - x_2 k_2\right) &= k_2\left(f_1 - x_2 f_2\right) - f_2\left(k_1 - x_2 k_2\right) \\
&\quad + \left[\left(f_1 - x_2 f_2\right)\left(k_2 - k_1\right) - \left(k_1 - x_2 k_2\right)\left(f_2 - f_1\right)\right]\lambda Q \\
&= \left(f_1 k_2 - x_2 f_2 k_2\right) - \left(f_2 k_1 - x_2 f_2 k_2\right) \\
&\quad + \left[f_1\left(k_2 - k_1\right) - x_2 f_2\left(k_2 - k_1\right) - \left(k_1\left(f_2 - f_1\right) - x_2 k_2\left(f_2 - f_1\right)\right)\right]\lambda Q \\
&= f_1 k_2 - x_2 f_2 k_2 - f_2 k_1 + x_2 f_2 k_2 \\
&\quad + \left[f_1\left(k_2 - k_1\right) - \left(x_2 f_2 k_2 - x_2 f_2 k_1\right) - k_1\left(f_2 - f_1\right) + \left(x_2 k_2 f_2 - x_2 k_2 f_1\right)\right] \\
&= f_1 k_2 - f_2 k_1 \\
&\quad + \left[f_1\left(k_2 - k_1\right) - k_1\left(f_2 - f_1\right) + x_2 f_2 k_1 - x_2 k_2 f_1\right]\lambda Q \\
&= f_1 k_2 - f_2 k_1 \\
&\quad + \left[f_1 k_2 - f_2 k_1 + x_2 f_2 k_1 - x_2 f_1 k_2\right]\lambda Q \\
&= f_1 k_2 - f_2 k_1 + \left[f_1 k_2\left(1 - x_2\right) - f_2 k_1\left(1 - x_2\right)\right]\lambda Q \\
&= \left(f_1 k_2 - f_2 k_1\right)\left(1 + \lambda Q\left(1 - x_2\right)\right),
\end{aligned}$$

and one obtains

$$A = \frac{\left(f_1 k_2 - f_2 k_1\right)\left(1 + \lambda Q\left(1 - x_2\right)\right)}{M}$$

Now, substituting these expressions of A and B back into the compact expression of the FOC for x_1 gives

$$\begin{aligned}
A\left(E - 1\right) + \lambda B &= 0 \\
\frac{\left(f_1 k_2 - f_2 k_1\right)\left(1 + \lambda Q\left(1 - x_2\right)\right)}{M}\left(E - 1\right) + \lambda x_1 \frac{f_2 k_1 - k_2 f_1}{M} &= 0 \\
\left(f_1 k_2 - f_2 k_1\right)\left(1 + \lambda Q\left(1 - x_2\right)\right)\left(E - 1\right) - \lambda x_1\left(f_1 k_2 - f_2 k_1\right) &= 0 \\
\left(f_1 k_2 - f_2 k_1\right)\left(\left(1 + \lambda Q\left(1 - x_2\right)\right)\left(E - 1\right) - \lambda x_1\right) &= 0 \\
\left(1 + \lambda Q\left(1 - x_2\right)\right)\left(E - 1\right) &= \lambda x_1 \\
\left(E - 1\right) &= \lambda\left(x_1 - \left(E - 1\right)Q\left(1 - x_2\right)\right) \\
\lambda &= \frac{E - 1}{x_1 - \left(E - 1\right)Q\left(1 - x_2\right)}
\end{aligned}$$

provided $f_1 k_2 \neq f_2 k_1$.

We hence obtain:

$$\begin{aligned}
\lambda &= \frac{E-1}{x_1 - (E-1) Q (1-x_2)} \\
&= \frac{E-1}{x_1 - (E-1) \frac{x_2 E}{E-1} (1-x_2)} \\
&= \frac{\frac{z_2}{z_1} - 1}{x_1 - x_2 (1-x_2) \frac{z_2}{z_1}} \\
&= \frac{z_2 - z_1}{x_1 z_1 - x_2 (1-x_2) z_2} \\
&= \frac{\exp(\lambda x_2) - \exp(\lambda x_1)}{x_1 \exp(\lambda x_1) - x_2 (1-x_2) \exp(\lambda x_2)}
\end{aligned}$$

where we made use of the definitions: $Q = \frac{x_2 E}{E-1}$ and $E = \exp(\lambda (x_2 - x_1)) = \frac{z_2}{z_1}$.

To summarize, in the softmax model we have

$$\begin{aligned}
\lambda &= \frac{\exp(\lambda x_2) - \exp(\lambda x_1)}{x_1 \exp(\lambda x_1) - x_2 (1-x_2) \exp(\lambda x_2)}, \\
\beta &= -\frac{f_2 (\exp(\lambda x_2) - \exp(\lambda x_1)) + \lambda (f_2 - f_1) x_2 \exp(\lambda x_2)}{k_2 (\exp(\lambda x_2) - \exp(\lambda x_1)) + \lambda (k_2 - k_1) x_2 \exp(\lambda x_2)}.
\end{aligned}$$

Table (B.4) shows the estimates of the parameters of the model with this parametric specification.

Table B.4: Parameter estimates with 95% bootstrap confidence intervals for selected players

Player	Salience weight δ	Curvature λ	Slope f τ_f	Const. f a_f	Slope k τ_k	Const. k a_k
Novak Djokovic	0.32^{***} [0.20, 0.47]	3.54^{***} [3.40, 3.70]	-0.01^{***} [-0.01, -0.01]	0.45^{***} [0.44, 0.46]	-0.00^{***} [-0.00, -0.00]	0.41^{***} [0.40, 0.41]
Rafael Nadal	0.27^{***} [0.08, 0.54]	4.97^{***} [4.75, 5.21]	-0.00^{***} [-0.00, -0.00]	0.33^{***} [0.32, 0.33]	-0.00 [-0.00, 0.00]	0.43^{***} [0.43, 0.44]
Roger Federer	0.38^{***} [0.28, 0.49]	2.46^{***} [2.35, 2.57]	-0.04^{***} [-0.05, -0.04]	0.60^{***} [0.58, 0.61]	-0.01^{***} [-0.01, -0.00]	0.38^{***} [0.38, 0.39]
Pete Sampras	0.45^{***} [0.33, 0.60]	1.35^{***} [1.20, 1.51]	-0.26^{***} [-0.32, -0.21]	1.10^{***} [1.04, 1.18]	-0.06^{***} [-0.08, -0.05]	0.40^{***} [0.38, 0.43]
Boris Becker	2.73^{***} [0.50, 3.70]	-11.45 [-11.89, 1.25]	-0.00^{***} [-0.35, -0.00]	0.49^{***} [0.48, 1.12]	-0.00^{***} [-0.05, -0.00]	0.30^{***} [0.29, 0.39]
Carlos Alcaraz	0.10 [-0.08, 0.32]	3.55^{***} [3.28, 3.81]	-0.01^{***} [-0.01, -0.01]	0.41^{***} [0.39, 0.43]	-0.00^{***} [-0.00, -0.00]	0.41^{***} [0.40, 0.42]
Jannik Sinner	0.10 [-0.01, 0.24]	2.11^{***} [1.93, 2.28]	-0.06^{***} [-0.07, -0.05]	0.60^{***} [0.57, 0.63]	-0.01^{***} [-0.02, -0.01]	0.42^{***} [0.41, 0.43]
Ivo Karlovic	1.23^{***} [0.66, 2.44]	4.54^{***} [3.79, 5.49]	-0.01^{***} [-0.02, -0.00]	0.73^{***} [0.69, 0.78]	-0.00^{***} [-0.00, -0.00]	0.27^{***} [0.25, 0.29]
John Isner	0.82^{***} [0.54, 1.20]	5.50^{***} [4.98, 6.10]	-0.00^{***} [-0.00, -0.00]	0.66^{***} [0.63, 0.68]	-0.00^{***} [-0.00, -0.00]	0.27^{***} [0.26, 0.28]
Reilly Opelka	0.39^{***} [0.17, 0.68]	3.80^{***} [3.28, 4.35]	-0.01^{***} [-0.02, -0.01]	0.72^{***} [0.68, 0.76]	-0.01^{***} [-0.01, -0.00]	0.29^{***} [0.27, 0.31]
David Ferrer	0.04 [-0.29, 0.74]	2.61^{***} [2.24, 3.02]	-0.02^{***} [-0.03, -0.01]	0.33^{***} [0.30, 0.37]	0.00 [-0.00, 0.01]	0.39^{***} [0.37, 0.41]
Diego Schwartzman	-0.35 [-0.73, 0.38]	3.46^{***} [2.93, 4.09]	-0.01^{***} [-0.01, -0.00]	0.27^{***} [0.24, 0.30]	0.00 [-0.00, 0.00]	0.41^{***} [0.39, 0.43]
Mean	0.47	2.13	-0.08	0.58	-0.01	0.38
Std	0.78	3.15	0.15	0.23	0.03	0.05
Min	-3.65	-14.92	-1.24	0.22	-0.25	0.16
Median	0.41	2.56	-0.03	0.55	-0.00	0.38
Max	3.09	6.66	-0.00	1.98	0.17	0.64

Notes: 95% confidence intervals obtained by bootstrap with 2000 replications.*** indicates significance at the 5% level: for λ, CI excludes 1; for other parameters, CI excludes 0.

B8 Model with different curvature for $f(x)$ and $k(x)$

How stable are the estimates of the salience weight to departure from the assumption that the curvatures of f and k are the same? This question relates to the second aspect of Condition (2) which forces the elasticity of the marginal probabilities of winning one-shot and multi-shot points to be the same. We propose to do so by replacing Condition (2) by:

Condition B.5 $x\frac{f''(x)}{f'(x)} = \lambda - 1$, $x\frac{k''(x)}{k'(x)} = t\lambda - 1$, *where* $\lambda > 0$ *and* $t > 0$.

Note that we cannot estimate the parameter t together with the other parameters since we have already exhausted all the degrees of freedom offered by the data. However, one can still check the robustness of our estimates of the salience weight, by estimating them *given* t and repeating for various values of t. We follow this approach and given Condition (B.5) adopt the following parametrization:

$$\begin{aligned} f\left(x\right) &= \frac{a_f - x_1^{\lambda}}{\tau_f}, \\ k\left(x\right) &= \beta\frac{a_k - x_1^{t\lambda}}{\tau_k}, \end{aligned}$$

where the parameter $t > 0$ allows us to introduce different curvature for the two functions.

For $t = 1$, we recover the main specification from Condition (2). For $t < 1$, there is less curvature in $k\left(x\right)$ and for $t > 1$ there is more curvature in $k\left(x\right)$.

Before we can proceed to the estimation of the salience weights for different values of t, we need to recover the formulas necessary to search for the value of λ given the data and the value of t.

This is done by reworking the FOCs under this specification. Recall first that using the FOC for the second serve optimal strategy one recovers the value of the salience weight condition on λ (and now also t). This yields

$$\begin{aligned} \left[f\left(x_2\right) + \beta k\left(x_2\right)\right] + x_2\left[\frac{\partial f\left(x_2\right)}{\partial x} + \beta\frac{\partial k\left(x_2\right)}{\partial x}\right] &= 0 \\ &\Leftrightarrow \\ \left[\frac{a_f - x_2^{\lambda}}{\tau_f} + \beta\frac{a_k - x_2^{t\lambda}}{\tau_k}\right] - \lambda\left[\frac{x_2^{\lambda}}{\tau_f} + \beta t\frac{x_2^{t\lambda}}{\tau_k}\right] &= 0 \\ &\Leftrightarrow \\ \left[\frac{a_f - x_2^{\lambda}}{\tau_f} + \beta\frac{a_k - x_2^{t\lambda}}{\tau_k}\right] &= \lambda\left[\frac{x_2^{\lambda}}{\tau_f} + \beta t\frac{x_2^{t\lambda}}{\tau_k}\right] \\ &\Leftrightarrow \\ \frac{a_f}{\tau_f} + \beta\frac{a_k}{\tau_k} - \frac{x_2^{\lambda}}{\tau_f} - \beta\frac{x_2^{t\lambda}}{\tau_k} &= \lambda\frac{x_2^{\lambda}}{\tau_f} + \lambda t\beta\frac{x_2^{t\lambda}}{\tau_k} \\ &\Leftrightarrow \\ \frac{a_f}{\tau_f} - \frac{x_2^{\lambda}}{\tau_f} - \lambda\frac{x_2^{\lambda}}{\tau_f} &= \lambda t\beta\frac{x_2^{t\lambda}}{\tau_k} - \beta\frac{a_k}{\tau_k} + \beta\frac{x_2^{t\lambda}}{\tau_k} \\ &\Leftrightarrow \\ a_f - x_2^{\lambda} - \lambda x_2^{\lambda} &= \lambda t\beta\frac{\tau_f}{\tau_k}x_2^{t\lambda} - \beta\frac{a_k\tau_f}{\tau_k} + \beta\frac{\tau_f}{\tau_k}x_2^{t\lambda} \\ &\Leftrightarrow \\ a_f - x_2^{\lambda} - \lambda x_2^{\lambda} &= \beta\left(t\lambda\frac{\tau_f}{\tau_k}x_2^{t\lambda} - \frac{a_k\tau_f}{\tau_k} + \frac{\tau_f}{\tau_k}x_2^{t\lambda}\right) \\ &\Leftrightarrow \\ \beta &= \frac{a_f - x_2^{\lambda} - \lambda x_2^{\lambda}}{t\lambda\frac{\tau_f}{\tau_k}x_2^{t\lambda} - \frac{a_k\tau_f}{\tau_k} + \frac{\tau_f}{\tau_k}x_2^{t\lambda}} \\ &\Leftrightarrow \\ \beta &= -\frac{a_f\tau_k - \tau_k\left(1 + \lambda\right)x_2^{\lambda}}{a_k\tau_f - \tau_f\left(1 + t\lambda\right)x_2^{t\lambda}} \end{aligned}$$

hence

$$\beta = -\frac{a_f\tau_k - \tau_k\left(1 + \lambda\right)z_2}{a_k\tau_f - \tau_f\left(1 + t\lambda\right)w_2}$$

where $z_j = x_j^{\lambda}$ and $w_j = x_j^{t\lambda}$ for $j = 1, 2$.

Doing a similar exercise using the FOC for the first serve optimal strategy to recover the updating equation for λ conditional on t yields

$$\begin{array}{rcl}
\left[f\left(x_1\right)+\beta k\left(x_1\right)\right]+x_1\left[\dfrac{\partial f\left(x_1\right)}{\partial x}+\beta\dfrac{\partial k\left(x_1\right)}{\partial x}\right] & = & x_2\tilde{y}\left(x_2\right) \\
 & \Leftrightarrow & \\
\left[f\left(x_1\right)+\beta k\left(x_1\right)\right]+x_1\left[\dfrac{\partial f\left(x_1\right)}{\partial x}+\beta\dfrac{\partial k\left(x_1\right)}{\partial x}\right] & = & x_2\left(f\left(x_2\right)+\beta k\left(x_2\right)\right) \\
 & \Leftrightarrow & \\
\left[\dfrac{a_f-x_1^{\lambda}}{\tau_f}+\beta\dfrac{a_k-x_1^{t\lambda}}{\tau_k}\right]-\lambda\left[\dfrac{x_1^{\lambda}}{\tau_f}+\beta t\dfrac{x_1^{t\lambda}}{\tau_k}\right] & = & x_2\left(\dfrac{a_f-x_2^{\lambda}}{\tau_f}+\beta\dfrac{a_k-x_2^{t\lambda}}{\tau_k}\right) \\
 & \Leftrightarrow & \\
\left[\dfrac{a_f}{\tau_f}-\dfrac{x_1^{\lambda}}{\tau_f}+\beta\dfrac{a_k}{\tau_k}-\beta\dfrac{x_1^{t\lambda}}{\tau_k}\right]-\left[\lambda\dfrac{x_1^{\lambda}}{\tau_f}+t\lambda\beta\dfrac{x_1^{t\lambda}}{\tau_k}\right] & = & x_2\left(\dfrac{a_f-x_2^{\lambda}}{\tau_f}+\beta\dfrac{a_k-x_2^{t\lambda}}{\tau_k}\right)
\end{array}$$

We can now proceed and isolate λ,

$$\begin{array}{rcl}
\left[\dfrac{a_f}{\tau_f}-\dfrac{x_1^{\lambda}}{\tau_f}+\beta\dfrac{a_k}{\tau_k}-\beta\dfrac{x_1^{t\lambda}}{\tau_k}\right]-\left[\lambda\dfrac{x_1^{\lambda}}{\tau_f}+t\lambda\beta\dfrac{x_1^{t\lambda}}{\tau_k}\right] & = & x_2\left(\dfrac{a_f-x_2^{\lambda}}{\tau_f}+\beta\dfrac{a_k-x_2^{t\lambda}}{\tau_k}\right) \\
 & \Leftrightarrow & \\
\left[a_f-z_1+\beta\dfrac{\tau_f a_k}{\tau_k}-\beta\dfrac{\tau_f w_1}{\tau_k}\right]-\left[\lambda z_1+t\lambda\beta\dfrac{\tau_f w_1}{\tau_k}\right] & = & x_2\left(a_f-z_2+\beta\dfrac{\tau_f\left(a_k-w_2\right)}{\tau_k}\right) \\
 & \Leftrightarrow & \\
\left[a_f-z_1+R\left(a_k-w_1\right)\right]-\left[\lambda z_1+t\lambda R w_1\right] & = & x_2\left(a_f-z_2+R\left(a_k-w_2\right)\right)
\end{array}$$

where $R = \beta\frac{\tau_f}{\tau_k}$.

A simple rearrangement yields

$$\begin{array}{rcl}
\lambda\left[z_1+tw_1R\right] & = & \left[a_f-z_1+R\left(a_k-w_1\right)\right]-x_2\left(a_f-z_2+R\left(a_k-w_2\right)\right) \\
 & \Leftrightarrow & \\
\lambda & = & \dfrac{\left[a_f-z_1+R\left(a_k-w_1\right)\right]-x_2\left(a_f-z_2+R\left(a_k-w_2\right)\right)}{\left[z_1+tw_1R\right]}
\end{array}$$

It follows that the updating equation given t reads as

$$\lambda = \frac{a_f-z_1+R\left(a_k-w_1\right)-x_2\left(a_f-z_2+R\left(a_k-w_2\right)\right)}{z_1+tw_1R}.$$

We note that when $t = 1$, this equation reads as the one obtained under Condition (2), for then $z_j = w_j$ for $j = 1, 2$.

The algorithm presented in the paper can be amended to account for this more general structure. In particular, it takes as input the data x_1, x_2, f_1, f_2, k_1 and k_2 together with a value for t. The output is still the curvature parameter λ. However, the slopes and constants of $f(x)$ and $k(x)$ now obtain as follows. Let $z_j = x_j^{\lambda}$ and $w_j = x_j^{t\lambda}$ for $j = 1, 2$. Note that

$$\begin{aligned} \tau_f &= -\frac{z_2}{f_2}\frac{\frac{z_1 - z_2}{z_2}}{\frac{f_1 - f_2}{f_2}}, \\ \tau_k &= -\frac{w_2}{k_2}\frac{\frac{w_1 - w_2}{w_2}}{\frac{k_1 - k_2}{k_2}}, \end{aligned}$$

and

$$\begin{aligned} a_f &= \tau_f f_1 + z_1, \\ a_k &= \tau_k k_1 + w_1. \end{aligned}$$

The preference parameter is obtained as

$$\beta = -\frac{a_f \tau_k - \tau_k (1 + \lambda) z_2}{a_k \tau_f - \tau_f (1 + t\lambda) w_2},$$

and, writing

$$R = \beta \frac{\tau_f}{\tau_k},$$

the updating equation now becomes

$$\lambda = \frac{a_f - z_1 + R(a_k - w_1) - x_2 (a_f - z_2 + R(a_k - w_2))}{z_1 + t w_1 R}.$$

Table (B.5) shows the estimates of the salience weight and that Table (B.6) of the curvature of the model with this parametric specification. Note that for a few players, e.g., Pete Sampras and Reilly Opelka, for instance, there is no solution for the curvature parameter λ at $t = 0.5$.

Table B.5: Stability of salience weight estimates to differences in relative curvature between $f(x)$ and $k(x)$, for selected players.

Player	t						
	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Novak Djokovic	0.26	0.27	0.27	0.27	0.25	0.23	0.19
Rafael Nadal	0.21	0.21	0.21	0.21	0.21	0.20	0.19
Roger Federer	0.31	0.32	0.32	0.32	0.30	0.27	0.24
Pete Sampras	n.a.	0.42	0.42	0.42	0.39	0.35	0.29
Boris Becker	0.56	0.56	0.56	0.56	0.55	0.54	0.51
Carlos Alcaraz	0.05	0.05	0.05	0.05	0.04	0.03	0.01
Jannik Sinner	0.05	0.06	0.06	0.06	0.04	0.02	−0.01
Ivo Karlovic	1.15	1.18	1.19	1.17	1.08	0.95	0.81
John Isner	0.73	0.77	0.79	0.73	0.59	0.43	0.29
Reilly Opelka	n.a.	0.35	0.36	0.33	0.23	0.12	0.01
David Ferrer	−0.01	−0.01	−0.01	−0.01	−0.01	−0.00	0.00
Diego Schwartzman	−0.38	−0.38	−0.38	−0.38	−0.38	−0.38	−0.38
Mean	0.31	0.33	0.34	0.38	0.35	0.30	0.27
Std	0.91	0.72	0.70	0.55	0.58	0.61	0.59
Min	−5.25	−4.04	−3.86	−2.54	−1.85	−2.62	−2.48
Median	0.32	0.33	0.33	0.33	0.31	0.28	0.25
Max	2.92	2.81	2.79	2.80	2.84	2.90	2.96

Notes: Each column reports the estimate of the salience weight δ for a given value of the relative curvature t of multi-shot probability of winning a point, relative to one-shot. n.a. means non available because there is no solution for λ.

Table B.6: Stability of curvature estimates to differences in relative curvature between $f(x)$ and $k(x)$, for selected players.

Player	t						
	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Novak Djokovic	3.22	3.44	3.67	3.90	4.10	4.27	4.41
Rafael Nadal	4.74	4.82	4.89	4.96	5.01	5.06	5.10
Roger Federer	2.44	2.62	2.81	3.00	3.17	3.31	3.42
Pete Sampras	n.a.	1.71	1.93	2.18	2.44	2.67	2.84
Boris Becker	1.54	1.63	1.73	1.83	1.93	2.03	2.12
Carlos Alcaraz	3.34	3.50	3.66	3.81	3.95	4.07	4.17
Jannik Sinner	2.15	2.33	2.52	2.72	2.90	3.06	3.18
Ivo Karlovic	3.14	3.67	4.37	5.21	5.99	6.55	6.88
John Isner	3.11	3.95	5.33	7.07	8.14	8.54	8.66
Reilly Opelka	n.a.	2.97	3.89	5.07	5.89	6.25	6.36
David Ferrer	3.08	2.99	2.90	2.83	2.77	2.71	2.66
Diego Schwartzman	3.69	3.62	3.57	3.52	3.47	3.43	3.40
Mean	2.94	2.86	2.94	3.08	3.28	3.48	3.55
Std	1.30	1.10	0.98	1.09	1.38	1.55	1.55
Min	1.17	1.19	1.28	1.28	0.59	1.33	1.30
Median	2.73	2.74	2.86	3.00	3.07	3.17	3.26
Max	9.71	9.15	6.27	7.32	9.70	9.69	9.69

Notes: Each column reports the estimate of the curvature λ for a given value of the relative curvature t of multi-shot probability of winning a point, relative to one-shot. n.a. means non available because there is no solution for λ.

B9 Model of aversion for double faults

Suppose a player has a disutility from making double faults, and let $\gamma > 0$ capture the magnitude of this disutility. In this case, the utility the player would maximize is

$$x_1 y\left(x_1\right) + \left(1 - x_1\right)\left[x_2 y\left(x_2\right) - \gamma\left(1 - x_2\right)\right].$$

It is easy to show that, after some rearranging, this utility function rewrites as

$$x_1 \breve{y}\left(x_1\right) + \left(1 - x_1\right) x_2 \breve{y}\left(x_2\right) - \gamma$$

where $\breve{y}\left(x\right) = y\left(x\right) + \gamma$.

The FOCs of the maximization problem then read as

$$\begin{aligned} x_2 y'(x_2) + y(x_2) + \gamma &= 0 \\ x_1 y'(x_1) + y(x_1) + \gamma(1 - x_2) &= x_2 y(x_2) \end{aligned}$$

since $\check{y}'(x) = y'(x)$.

Assuming, as we did in the main model of the paper, that $y(x)$ is a power function, $y(x) = \frac{a}{\tau} - \frac{1}{\tau}x^\lambda$ where $\tau > 0$, so that $y'(x) = -\frac{1}{\tau}\lambda x^{\lambda-1} < 0$ and $xy'(x) = -\frac{1}{\tau}\lambda x^\lambda$, so that the FOCs read as

$$\begin{aligned} -\frac{1}{\tau}\lambda x_2^\lambda + \frac{a}{\tau} - \frac{1}{\tau}x_2^\lambda + \gamma &= 0, \\ -\frac{1}{\tau}\lambda x_1^\lambda + \frac{a}{\tau} - \frac{1}{\tau}x_1^\lambda + \gamma &= x_2\left(\frac{a}{\tau} - \frac{1}{\tau}x_2^\lambda + \gamma\right). \end{aligned}$$

We can then solve for the optimal serve strategies. Solving the first equation for x_2^λ yields

$$x_2^\lambda = \frac{a + \tau\gamma}{\lambda + 1}$$

while solving the second for x_1^λ yields

$$x_1^\lambda = x_2^\lambda\left(1 - \frac{\lambda}{\lambda + 1}x_2\right).$$

Comparing the optimal strategy from the aversion to the double fault model to that of the process utility model, one first notes that the first serve strategy has the same expression in both models. Since this equation is used to estimate the curvature parameter, when applied on the same data, i.e., given x_1 and x_2, both models deliver the same estimate of λ. It follows that the slope and constant of $y(x)$ that would obtain from data (x_1, x_2, y_1, y_2) are also the same in the two models and obtain as: $\tau = \frac{x_2^\lambda - x_1^\lambda}{y_1 - y_2}$ and $a = \frac{x_2^\lambda - x_1^\lambda}{y_1 - y_2}y_2 + x_2^\lambda$.

The differentce between the two models is apparent in the expression of the second serve strategies which read respectively as

$$\begin{aligned} x_2^\lambda &= \frac{a + \tau\gamma}{\lambda + 1}, && \text{(B.5)} \\ x_2^\lambda &= \frac{a_f + a_k\beta\frac{\tau_f}{\tau_k}}{(1 + \lambda)\left(1 + \beta\frac{\tau_f}{\tau_k}\right)}. && \text{(B.6)} \end{aligned}$$

Picking the right values of the parameters γ for the aversion to double faults model and (β, a_f and τ_f) for the process utility model, we conclude that the two models fit equally well the data (x_1, x_2, y_1, y_2) and share in common the curvature parameter λ and the slope and constant of $y(x)$, i.e., τ and a. However, note that from equation (B.5), the estimate of γ is invariant to data (f_1, f_2) conditional on (x_1, x_2, y_1, y_2). Stated otherwise, the optimal decision of the player only reflects his general ability to win points, i.e., parameters a and τ, and his aversion to double faults γ, regardless of his ability to win one-shot points, i.e., a_f and τ_f. If the true data-generating model were that where players had an aversion to making double faults, the parameter γ should be unrelated to the conditional probabilities of winning one-shot rallies on either the first or second serve, since all that matters for such a player's strategy is $y(x)$ (y_1 and y_2) and not their constituents (f_1, f_2, k_1 and k_2). However, if the true model is one where players value process utility, then the observed optimal strategies x_1 and x_2 would depend on the constituents of $y(x)$, i.e., $f(x)$ and $k(x)$. It follows that the parameters obtained from estimating the double-fault aversion model and, in particular, the computed value of γ, would be related to $f(x)$, since they are is obtained using the optimal strategies x_1 and x_2 that depend on a_f and τ_f.

In particular, comparing players with similar serve strategies (x_1, x_2) and conditional probabilities of winning points (y_1 and y_2), the conditional mixed moments $\mathbb{E}\left[\hat{\gamma} \times f(x_j) | x_1, x_2, y_1, y_2\right]$ for $j = 1, 2$ should be zero. To test these hypotheses, we first estimate, for each player, the aversion for double faults parameter $\hat{\gamma}$ using the procedure outlined in the paper, but applied to the model with aversion to double faults. Then, we estimate the conditional expectations $\mathbb{E}\left[\hat{\gamma} \times f_1 | x_1, x_2, y_1, y_2\right]$ and $\mathbb{E}\left[\hat{\gamma} \times f_2 | x_1, x_2, y_1, y_2\right]$ nonparametrically using locally weighted scatterplot smoothing (LOWESS) and apply a bootstrap procedure (300 replications) to construct 95% confidence intervals. The in-sample predictions of the conditional expectations $\mathbb{E}\left[\hat{\gamma}_i \times f_{ji} | x_{1i}, x_{2i}, y_{1i}, y_{2i}\right]$ for $j = 1, 2$ and $i = 1, ...N$, and their confidence interval can then be used to check if the associated values are significantly different from 0. We find that for 99 out of the 151 players, at least one of the conditional expectations is statistically significantly larger than 0, rejecting the null hypothesis that the parameter $\hat{\gamma}$ is orthogonal to the way rallies are won. The estimates of the aversion to the double fault parameter γ seem to be systematically dependent on the conditional probability of winning one-shot rallies on first and second serves, conditional on the serve strategy and probability of winning points. This suggests that there is information in the conditional probability of winning one-shot and multi-shot rallies that a model with aversion to double fault does not capture, and that is being forced into the parameter γ.

Moreover, the fact that these expectations are positive tells a counterintuitive story, where, holding the serve strategy and probabilities of winning a point constant, those that rely more on winning points through one-shot rallies, i.e., the better serves, are the most averse to making double faults according to the model.