Disequilibrium Play in Tennis

Axel Anderson, Georgetown University
Jeremy Rosen, Georgetown University
John Rust, Georgetown University
Kin-Ping Wong, Digonex

April 7, 2025

Bottom Line of Talk

- We study the serve direction choices of the world's best tennis pros.
- We decisively reject the hypothesis that the server is playing a **best response** to the receiver.
- Our estimates of the size of the deviations for these world-class tennis pros are quite large, on the order of hurting their odds of winning each service game by 10 percentage points or more.
- These results contradict the conventional wisdom that pros' behavior is close to optimal.

Key Axioms of Economics

1. Rationality/Optimality
2. (Nash) Equilibrium

- These make eminent sense as key assumptions for theories of *perfectly rational behavior*
- But how closely does observed behavior hew to the predictions of perfectly rational models?

The Standard Critique of Lab Experiments

- The laboratory offers a controlled environment for testing the predictions of Nash equilibrium. However, many laboratory tests have rejected Nash equilibrium.
- The disadvantages of laboratory tests include:
  1. The subjects are inexperienced with the games played.
  2. The payoffs are insufficient to motivate the subjects to learn Nash behavior.
  3. The “artificial” nature of the games vs. the “real” world.
  4. Subjects’ payoffs may include more than their own monetary rewards.

Empirical Analysis of “Games in the Wild”

- Games in the wild overcome many of these problems:
  1. The players are often experienced with the games.
  2. The stakes can be very high.
- Real world games have their own shortcomings:
  1. The environments are complex: impossible to model every element of most real world games.
     - Solve for all (or even any) NE in chess!
  2. Elements of payoffs and/or choices may be unobserved.

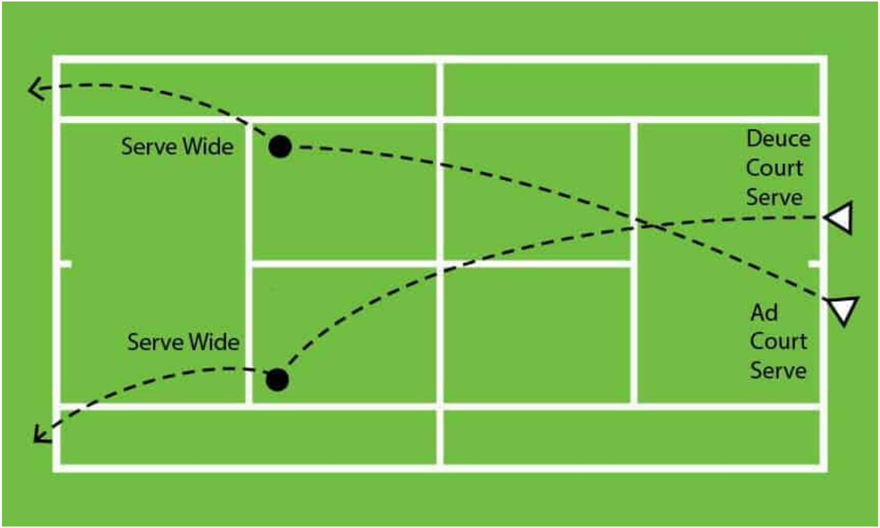
Empirical Analysis of Sports Competitions

- Sports competitions mitigate the disadvantages of laboratory games, while also avoiding many of the shortcomings of other real world games:
  1. The players are very experienced with the game.
  2. The stakes are typically very high.
  3. The rules are clear and simpler than many other games.
  4. The data on payoffs and choices is often very good.
  5. Zero-sum outcomes imply mild assumptions on payoffs: Often we only need *winning* preferred to *losing*.

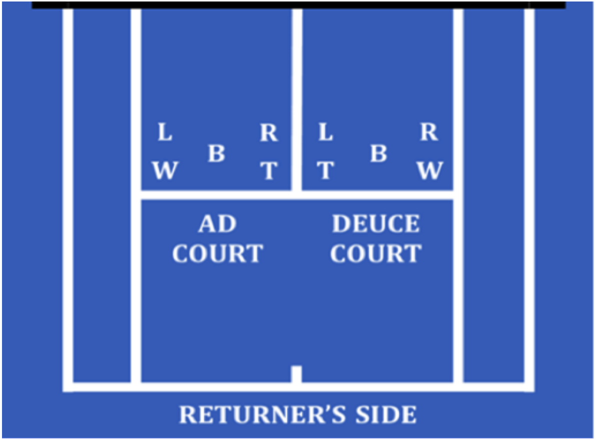
WW (2001): Empirical Analysis of Serve Direction

- A classic sports study is Walker and Wooders (2001): *Minimax Play at Wimbledon*
- They collected data on serve direction for first serves at top tournament finals, classifying all serves as “left” or “right.”
- In each of 40 server-receiver-**court** triples, both serve directions were observed in the data.
- For each server-receiver-court triple, they calculated the fraction of points won when serving in each direction.

Illustration of Serves by Court



Serve Directions from the Receiver's Perspective



WW (2001): Empirical Analysis of Serve Direction

- ***Null Hypothesis:*** *The probability the server wins a point in a given court is the same whether (s)he serves left or right.*
- The null hypothesis is valid as long as:
  1. The game payoff-matrix for each server-receiver pair is stable across points within each court.
  2. The server is best responding to the receiver ($\Leftarrow$ NE)
- WW fail to reject this hypothesis for 39 out of 40 of the server-receiver-court triples.
- WW Conclude: "The theory has performed far better in explaining the play of top professional tennis players in our data set [than in the laboratory]."

WW (2001): Negative Serial Correlation

- WW also test whether *the directions of successive serves are serially independent.*
- They find mild evidence of negative serial correlation.
- But the data are much closer to serial independence than in laboratory experiments.
- There is substantial evidence that humans are not so good at producing serially independent sequences.

Unlike WW, We Reject Equal Win Probabilities

- Walker and Wooders' general conclusions have been replicated in other sports, e.g. Palacios-Huerta (2003) study of soccer penalty kicks, *Professionals Play Minimax.*
- We deviate from WW on three dimensions:
  1. We enrich the strategy space by adding 2nd+body serves.
  2. We have a much larger data set.
  3. We model tennis as a *dynamic game*.
- We strongly reject the hypothesis of equal win probabilities and find substantial evidence of serial correlation in successive serve directions.
- This is not merely because we have a much larger data set: the deviations are large (in terms of win probabilities).

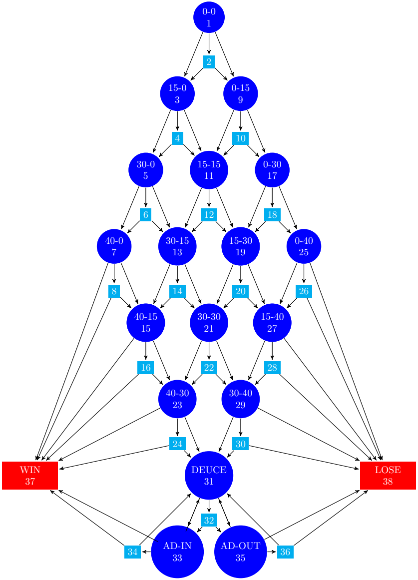
An Overview of the Rules of Tennis

- A *tennis match* consists of a series of *sets.*
- A *tennis set* consists of a sequence of *service games.*
- A *service game* consists of a sequence of *points.*

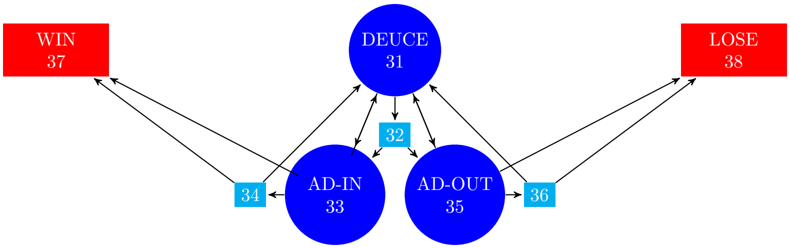
Unlike WW, our unit of analysis is the *service game:*

- Within each *service game* one of the players is the server. The server of the first game is chosen by a coin flip, and alternates in successive games thereafter.
- A service game consists of a sequence of stage games called *points.* A point consists of a first serve plus a second serve in the event of a faulted first serve.

Graph of State Transitions in the Service Game



Detail on the "Deuce Endgame"



Score States in the Service Game

- Let $x \in \{1, \ldots, 38\}$ denote the *score state.* $x = 1$ is the initial node of the game, corresponding to a score of 0–0.
- If the server wins the first serve, the state transits to $x = 3$ corresponding to a score of 15–0. If the server faults the first serve, the state transits to $x = 2$, corresponding to a second serve at score 0–0, etc.
- There are two *absorbing states:* $x = 37$ (server wins the game) and $x = 38$ (server loses the game).

Modeling Server and Receiver Choices

- At each state x except the two absorbing states, the server chooses a *serve type* $t = (s, d)$ where $s \in \mathcal{S} \subset R^2$ denotes serve *speed and spin*, and $d \in \{l, b, r\}$ is the serve direction (left, to the body, or right).
- The receiver *anticipates* (s, d), which includes observable choices (i.e. where to stand) and unobservable choices (i.e. where to focus their attention).
- We model anticipation with the vector $(a_l, a_b, a_r) \geq 0$ where a_d is the fraction of the receiver's "anticipation budget" for anticipating a serve to direction d. We normalize the budget to 1, so $1 = a_l + a_b + a_r$.

Accounting for (Short-Term) *Muscle Memory*

- We allow for the possibility that performance on the current serve could depend on recent serve history:
  - Maybe the server can improve his/her probability of making a serve by repeating a recent serve of a similar type.
  - Maybe the receiver is better able to handle a serve (s)he has recently faced.
- Adding muscle memory to the model gives the theory a better chance to match the data.
  - Serial dependence in serve directions is consistent with dynamic equilibrium behavior given MM.
- This makes our rejection of the theory more compelling.

The Muscle Memory State

- The *muscle memory state* $m = (d_1, d_2)$ encodes the direction of the previous two first serves.
  - On first serves, the “active" state is the direction of the previous first serve to the same court.
  - On second serves, the active state is the direction of the faulted first serve.
- Muscle memory starts in the null state, $m = (\emptyset, \emptyset)$.
- If the muscle memory is in state m and the server hits a first serve in direction d, then muscle memory updates to $m'(m, d) = (d, d_1)$.
- Muscle memory does not change after a second serve.

Score State Transition Probabilities

- Let $\ell(c(x), m, t)$ be the probability a serve lands in (i.e. is not a fault) given deuce/ad court $c(x)$, muscle memory m, and serve type $t = (s, d)$.
- Let $\omega(c(x), m, t, a)$ be the probability that the server wins the subsequent rally, given the serve type t is in, deuce/ad court is $c(x)$, muscle memory is m, and the receiver's anticipation vector is a.
- ℓ is continuous in s and ω is continuous in (s, a).

Stationarity and MPE

Assumption (Stationarity I)

The conditional probabilities ℓ and ω may vary across server-receiver pairs, but do not vary over time (independent of muscle memory *m*) or across service games.

- We assume stationarity in order to *pool observations across service games played by fixed server-receiver pairs.*
- For example, stationarity generically rules out fatigue.

Theorem (MPE Existence and Value Uniqueness)

All sub-games have a unique value (probability the server wins the game). There exists a Markov Perfect Equilibrium (MPE) in which strategies only depend on the current state (x, m).

Point Outcome Probabilities (POPs)

- We only observe serve direction d and the serve state x in our data from the Match Charting Project.
- We can "project" the full strategies of the players onto the data we do observe.
- Let $\rho(s|x, m)$ denote the probability distribution over speed and spin, and $\alpha(a|x, m)$ the probability distribution over receiver anticipation in a (potentially) mixed MPE of tennis.
- Let $\pi(in|x, m, d)$ and $\pi(win|x, m, d)$ denote the *point outcome probabilities* induced by these strategies.

Point Outcome Probabilities (POPs)

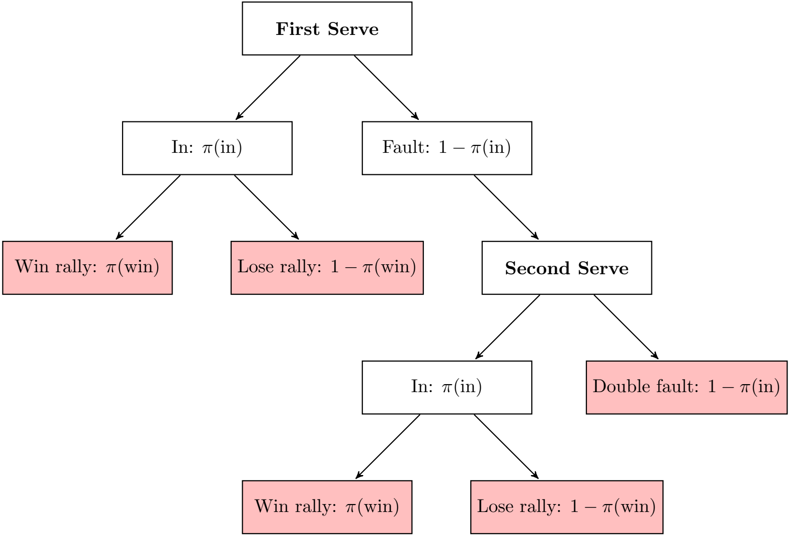
Definition (Point Outcome Probabilities)

If the stationarity assumption holds and (ρ, α) are MPE of the service game, then the *point outcome probabilities* (POPs) are:

$$\pi(in|x,m,d) \equiv \int \ell(c(x),m,d,s)\rho(ds|x,m)$$

$$\pi(win|x,m,d) \equiv \int\int \omega(c(x),m,d,s,a)\rho(ds|x,m)\alpha(da|x,m)$$

Detail on the "Point Stage Game"



The Induced Dynamic Program

- Assume the data is generated by some MPE with associated point outcome probabilities $\pi(in)$ and $\pi(win)$.
- Let $W_s(x, m)$ be the server's probability of winning the service game starting in state (x, m).
- Let $W_s(x, m, d)$ be the server's conditional win probability when serving to direction d in state (x, m).
- The Bellman equation relates these two win probabilities:

$$W_s(x, m) = \max_{d \in \{l,b,r\}} W_s(x, m, d)$$

- The optimal serve strategy, $\sigma_s(x, m)$ is the set of maximizers of the Bellman equation

$$\sigma_s(x, m) = \underset{d \in \{l,b,r\}}{argmax}\, W_s(x, m, d).$$

The Value at a Second Serve

- Let $x^+(x)$ ($x^-(x)$) denote the successor point states if the server wins (loses) the current point.
- If x is a second serve state, then W_s obeys:

$$W_s(x, m, d) = \pi(in)\pi(win)W_s(x^+(x), m) + (1 - \pi(in)\pi(win))W_s(x^-(x), m)$$

- Recall: Muscle memory only updates following first serves.

The Value at a First Serve

- Given a faulted first serve, score state x updates to $x + 1$.
- Recall $m'(m, d) = (d, d_1)$ denotes the muscle memory following a serve to direction d.
- Using these, we can express $W_s(x, m, d)$ recursively as:

$$
\begin{aligned}
W_s(x, m, d) \quad = \quad & \pi(in)\pi(win)W_s(x^+(x), m'(m, d)) \\
& +\pi(in)[1 - \pi(win)]W_s(x^-(x), m'(m, d)) \\
& +[1 - \pi(in)]W_s(x + 1, m'(m, d))
\end{aligned}
$$

Solving the Server's DP Problem

- While there is no time discounting, the POPs $(\pi(\textit{in}), \pi(\textit{win}))$ play an equivalent role, ensuring the existence of a unique solution to the Bellman equation via the Contraction Mapping Theorem.
- In practice, we calculate the server's DP problem by *backward induction*. However, the induction is not *over time* but rather *over states* using an algorithm Iskhakov, Schjerning and Rust (IRS 2016) refer to as *state recursion.*

Testable Implication: Equal Win Rates

- Recall that optimal serve directions $\sigma_s(x, m)$ obey

$$\sigma_s(x, m) = \underset{d \in \{l,b,r\}}{argmax}\, W_s(x, m, d).$$

Testable Implication of Equilibrium (Equal Win Probabilities)

If we observe serve directions d, d' chosen in state (x, m) then:

$$W_s(x, m, d) = W_s(x, m, d')$$

Data for Our Study

- From the Match Charting Project, a crowd-sourced database that records play by play outcomes of thousands of tennis matches, including classifying each serve direction as 1) wide, 2) down the T or 3) to the body
- We have $\approx$ ten times as many serves per server-receiver pair than Walker and Wooders used in their analysis.
- We focus on elite pro players who have all been ranked number one in the world and won multiple Grand Slams: Roger Federer, Rafael Nadal, Novak Djokovic, Andy Murray, Pete Sampras, and Andre Agassi.
- If we can show that they serve suboptimally, then even the best of the best are susceptible to strategic errors.

Mixed Serve Strategies: Stylized Facts

Server $\rightarrow$ receiver	Games, serves Serves/game	1st serves 2nd serves	Serve directions			Win prob (std) P-value: $P_1 = P_2$
			L	**B**	**R**	
Roger Federer $\rightarrow$	523, 4732	3208	.4402	.1007	.4592	.7686 (.0184)
Rafael Nadal	8.36	1164	.2174	.2698	.5129	5.1×10^{-60}
Rafael Nadal $\rightarrow$	519, 4081	3227	.6616	.2048	.1336	.8092 (.0172)
Roger Federer	7.86	854	.5937	.3208	.0855	6.3×10^{-12}
Roger Federer $\rightarrow$	411, 3501	2524	.4521	.0939	.4540	.8200 (.0190)
Novak Djokovic	8.52	977	.4084	.3408	.2508	6.7×10^{-68}
Novak Djokovic $\rightarrow$	407, 3653	2696	.4640	.1565	.3795	.8010 (.0198)
Roger Federer	8.98	957	.4389	.3365	.2247	1.0×10^{-33}
Rafael Nadal $\rightarrow$	346, 2937	2230	.3964	.2825	.3211	.7197 (.0241)
Novak Djokovic	8.49	707	.4073	.5403	.0523	2.4×10^{-64}
Novak Djokovic $\rightarrow$	356, 2877	2149	.4067	.1619	.4314	.7528 (.0222)
Rafael Nadal	8.08	728	.1484	.2940	.5577	1.2×10^{-40}
Novak Djokovic $\rightarrow$	230, 1958	1447	.4651	.1244	.4105	.7696 (.0278)
Andy Murray	8.51	511	.2192	.4618	.3190	3.0×10^{-53}
Andy Murray $\rightarrow$	230, 2141	1522	.3863	.0841	.5296	.7435 (.0288)
Novak Djokovic	9.31	619	.4233	.4782	.0985	5.8×10^{-122}
Pete Sampras $\rightarrow$	140, 1275	884	.4434	.0724	.4842	.9000 (.0254)
Andre Agassi	9.11	391	.4680	.1765	.3555	5.3×10^{-8}
Andre Agassi $\rightarrow$	135, 1125	825	.5127	.1115	.3758	.8666 (.0293)
Pete Sampras	8.33	300	.5766	.2700	.1533	7.2×10^{-16}

“Model-Free” Conclusions

- The “model-free” analysis reveals:
  - Servers use mixed strategies.
  - These strategies differ across first and second serves.
  - Servers adjust their serve strategy for different opponents.
- Service game win probabilities vary across server-receiver pairs (72–90%), and the server has an advantage.

Non-Parametric Estimation Problem

- To test equality of win probabilities across serve directions, we need accurate estimates of conditional win probabilities.
- Given sufficient observations, we could non-parametrically estimate conditional win probabilities.
- Accounting for muscle memory, there are a total of 894 conditional win probabilities $W_S(x, m, d)$.
- In order to estimate 894 probabilities with sufficient precision to have adequate power to test the hypothesis of equal win probabilities for all serve directions, we would need roughly 10,000 service games.
- Unfortunately in our data set we have between 135 and 523 service games per server-receiver pair.

Solution: Reduced-Form Approximation

- We partition the state and serve space into serve-court subsets and assume the conditional probabilities we wish to estimate are constant within each subset.
- We then estimate logit models for serve strategies $P(d|x, m)$ and the POPs $(\pi(in|x, m, d), \pi(win|x, m, d))$.
- We then use these estimates to recursively calculate the *implied conditional win probabilities* $W_P(x, m, d)$.
- Note that we cannot *assume* the estimated serve strategy is optimal, since this is what we ultimately want to test!
- Hausman tests show that the reduced-form model is a good approximation to non-parametric service game win probability estimates.

Allowing for Disequilibrium Serve Strategies

- Suppose the server employs a Markovian serve strategy, but perhaps *not a best response.*
- Let $P(s, d|x, m)$ denote the joint probability density that the server serves with speed/spin s and direction d in state (x, m), and let $P(d|x, m)$ be the marginal probability distribution over serve direction only.
- We maintain the following assumption about all (potentially disequilibrium) serve strategies and implied POPs

Assumption (Stationarity II)

The players use stationary, Markovian strategies, such that actual serve directions are given by a conditional probability $P(d|x, m)$ and the actual POPs are given by conditional probabilities $(\pi(in|x, m, d), \pi(win|x, m, d))$.

Valuing Disequilibrium Serve Strategies

- Let $W_P(x,m)$ be the win probability given Markovian strategy is $P(d|x,m)$:

$$W_P(x,m) = \sum_{d \in \{l,b,r\}} W_P(x,m,d) P(d|x,m)$$

  where $W_P(x,m,d)$ is the probability of winning given the server serves to direction d in state (x,m).

- The values $W_P(x,m,d)$ are given by the same recursive equations as for the optimal (equilibrium) serve strategies.

Testable Implication: No Profitable Deviations

- Suppose we estimate the win probabilities $W_P(x, m, d)$ *without imposing optimality/equilibrium.*

Testable Implication of Equilibrium (Server Best Response)

In all states (x, m) and for all serve directions d:

$$W_P(x, m, d) = W_s(x, m, d)$$

- This states that the win probabilities we calculate using the observed serve strategy P must equal the win probabilities for the dynamically optimal serve strategy, as long as the server is best responding.

The Reduced-Form Model for $P(d|x, m)$

- Let $f(x, m, d)$ be an indicator vector on the state space and direction d, and θ_P be the associated parameter vector.
- That is, f groups subsets of the state space and directions together, and θ_P applies scalar weights to each subset.
- The dot product $f(x, m, d)' \cdot \theta_P$ gives the relative weight given to direction d in state (x, m).
- The associated conditional serve directions probabilities are then:

$$P(d|x, m, \theta_P) = \frac{\exp\{f(x, m, d)'\theta_P\}}{\sum_{\delta \in \{l,b,r\}} \exp\{f(x, m, \delta)'\theta_P\}}$$

A Flexible, Agnostic Reduced-Form Model of Tennis

- Similarly for the POPs, let $g_{in}(x, m, d)$ and $g_{win}(x, m, d)$ be indicator vectors, and θ_{in} and θ_{win} be parameter vectors.
- Then the logit model for the conditional probabilities is:

$$\pi(in|x, m, d, \theta_{in}) = \frac{\exp\{g_{in}(x, m, d)'\theta_{in}\}}{1 + \exp\{g_{in}(x, m, d)'\theta_{in}\}}$$

$$\pi(win|x, m, d, \theta_{win}) = \frac{\exp\{g_{win}(x, m, d)'\theta_{win}\}}{1 + \exp\{g_{win}(x, m, d)'\theta_{win}\}}$$

Likelihood Function

- The likelihood function has 2 blocks: 1 for conditional choice probabilities (CCPs) and 1 for POPs
- For CCPs, there are 3 possible serve directions: $\{L, B, R\}$
- For POPs, there are 3 possible serve outcomes:
  - $o = 1$: serve is in and server wins the rally
  - $o = 2$: the serve is in and server loses the rally
  - $o = 3$: the serve is out, i.e. a fault
- Let $f(o|x, m, d, \theta_{in}, \theta_{win})$ be the conditional probability of these outcomes. We have

$$f(o|x, m, d, \theta_{in}, \theta_{win}) = \begin{cases} \pi(in|x, m, d, \theta_{in})\pi(win|x, m, d, \theta_{win}) & \text{if } o = 1 \\ \pi(in|x, m, d, \theta_{in})[1 - \pi(win|x, m, d, \theta_{win})] & \text{if } o = 2 \\ 1 - \pi(in|x, m, d, \theta_{in}) & \text{if } o = 3 \end{cases}$$

Parameter Estimation and Model Selection

- For any given set of partitions $\{f, g_{in}, g_{win}\}$ we estimate the associated parameter vector $\theta = (\theta_P, \theta_{in}, \theta_{win})$ by maximizing the associated log-likelihood function.
- Trivially, adopting a finer partition yields a better fit to the data, but results in less precise parameter estimates.
- In order to balance the tradeoff between model flexibility (bias) and overfitting (variance), we use the Akaike Information Criterion (AIC) to select our preferred serve-court specification.
- All of our qualitative results are robust to adopting several alternative specifications that we estimated.

Estimates of P, Djokovic and Federer

	Parameter Name	Djokovic $\to$ Federer Estimate	Djokovic $\to$ Federer Std Error	Federer $\to$ Djokovic Estimate	Federer $\to$ Djokovic Std Error
1	1st serve, deuce court, $d = l$	.836	(.102)	1.234	(.145)
2	1st serve, deuce court, $d = r$	.795	(.106)	1.402	(.143)
3	1st serve, deuce court, $d = d_{-2}$	.566	(.077)	.483	(.103)
4	2nd serve, deuce court, $d = l$	$-.208$	(.145)	$-.293$	(.132)
5	2nd serve, deuce court, $d = r$	$-.650$	(.151)	$-.531$	(.159)
6	2nd serve, deuce court, $d = d_{-1}$	$-.002$	(.168)	.137	(.144)
7	1st serve, ad court, $d = l$	.861	(.095)	1.367	(.133)
8	1st serve, ad court, $d = r$	.601	(.100)	1.319	(.132)
9	1st serve, ad court, $d = d_{-2}$	.294	(.079)	.396	(.086)
10	2nd serve, ad court, $d = l$	.462	(.146)	.551	(.152)
11	2nd serve, ad court, $d = r$	$-.386$	(.183)	$-.215$	(.171)
12	2nd serve, ad court, $d = d_{-1}$	$-.051$	(.136)	$-.143$	(.142)
	Observations, log-likelihood	2372, -2324.8		2333, -2265.06	
	AIC, BIC	4871.6, 4940.8		4554.1, 4623.2	

Estimates of $\pi(in|x, m, d)$, Djokovic and Federer

Parameter		Djokovic $\rightarrow$ Federer		Federer $\rightarrow$ Djokovic	
θ_{in}		Estimate	Std Error	Estimate	Std Error
1	1st serve, deuce court, $d = l$	.465	(.147)	.486	(.140)
2	1st serve, deuce court, $d = b$	.945	(.213)	.864	(.230)
3	1st serve, deuce court, $d = r$	.744	(.134)	.614	(.115)
4	1st serve, deuce court, $d = d_{-2}$	$-.093$	(.156)	$-.113$	(.145)
5	2nd serve, deuce court, $d = l$	3.468	(.572)	2.277	(.403)
6	2nd serve, deuce court, $d = b$	2.137	(.288)	3.249	(.459)
7	2nd serve, deuce court, $d = r$	2.150	(.445)	1.898	(.360)
8	2nd serve, deuce court, $d = d_{-2}$	.277	(.513)	.118	(.440)
9	1st serve, ad court, $d = l$	.604	(.159)	.122	(.148)
10	1st serve, ad court, $d = b$	.928	(.173)	.530	(.232)
11	1st serve, ad court, $d = r$	.813	(.142)	.422	(.128)
12	1st serve, ad court, $d = d_{-2}$	$-.138$	(.165)	.164	(.154)
13	2nd serve, ad court, $d = l$	2.214	(.406)	2.652	(.362)
14	2nd serve, ad court, $d = b$	1.917	(.352)	3.820	(.719)
15	2nd serve, ad court, $d = r$	1.430	(.374)	2.033	(.387)
16	2nd serve, ad court, $d = d_{-2}$	.294	(.452)	.329	(.499)
Observations, log-likelihood AIC, BIC		2333, -2403.9 4871.9, 5056.6		2372, -2324.8 4625.4, 4809.6	

Estimates of $\pi(win|x, m, d)$, Djokovic and Federer

Parameter		Djokovic $\rightarrow$ Federer		Federer $\rightarrow$ Djokovic	
θ_{win}		Estimate	Std Error	Estimate	Std Error
1	1st serve, deuce court, $d = l$	.641	(.182)	1.092	(.208)
2	1st serve, deuce court, $d = b$	.470	(.196)	.760	(.250)
3	1st serve, deuce court, $d = r$	.439	(.143)	.795	(.154)
4	1st serve, deuce court, $d = d_{-2}$	.456	(.188)	.314	(.210)
5	2nd serve, deuce court, $d = l$	.975	(.235)	.148	(.247)
6	2nd serve, deuce court, $d = b$	.593	(.195)	.073	(.182)
7	2nd serve, deuce court, $d = r$	.650	(.290)	.063	(.285)
8	2nd serve, deuce court, $d = d_{-2}$	$-.537$	(.273)	.038	(.270)
9	1st serve, ad court, $d = l$	.878	(.185)	1.397	(.232)
10	1st serve, ad court, $d = b$	.614	(.223)	1.055	(.305)
11	1st serve, ad court, $d = r$	.728	(.171)	.933	(.167)
12	1st serve, ad court, $d = d_{-2}$	$-.182$	(.194)	$-.392$	(.216)
13	2nd serve, ad court, $d = l$	$-.090$	(.221)	.292	(.221)
14	2nd serve, ad court, $d = b$	$-.282$	(.222)	$-.022$	(.222)
15	2nd serve, ad court, $d = r$	.530	(.295)	.142	(.263)
16	2nd serve, ad court, $d = d_{-2}$	.489	(.281)	.206	(.259)
Observations, log-likelihood AIC, BIC		2333, -2403.9 4871.9, 5056.6		2372, -2324.8 4625.4, 4809.6	

Summary of Reduced-Form Estimation Results

- Our preferred specification (i.e. smallest AIC) has 44 parameters: 12 θ_P and 16 each for $(\theta_{in}, \theta_{win})$.
- It balances the tradeoff described above by providing an accurate model of the entire service game for individual server-receiver pairs while avoiding overfitting.
- Conditional on a first serve going in, Federer has a higher probability of winning the subsequent rally than does Djokovic when he serves to Federer.
- However we find that Djokovic has a lower probability of faulting relative to Federer. (An exception is for 2nd serves to the deuce court where Federer has a uniformly lower probability of faulting for all three serve directions.)
- Thus, our estimates reflect an intuitive trade-off: a faster serve has a higher probability of faulting, but conditional on it going in, the receiver is at a disadvantage.

Stationarity Tests: Outline

- We assume the POPs $(\pi(in), \pi(win))$ are stationary.
- This assumption implies that our reduced-form model has stable parameters across service games.
- To test this, we first divide the service games between a server-receiver pair into multiple sub-groupings.
- We then estimate two models: the “unrestricted” model allows the parameters to differ across sub-groups, while the “restricted” model assumes the parameters are the same across subgroups.
- Our stationarity test statistic compares the difference in the log-likelihoods of the unrestricted and restricted models.
- We analyzed two types of sub-groupings: (1) early vs. late in a match and (2) across different calendar years.

Testing for Stationary POPs: Early vs. Late in Matches

Server → receiver	Muscle Memory			No Muscle Memory		
	Restricted	Unrestricted	LR test (df)	Restricted	Unrestricted	LR test (df)
	LL, AIC	LL, AIC	P-value	LL, AIC	LL, AIC	P-value
Roger Federer →	-1934.3	-1919.8	29.0 (32)	-1940.1	-1928.0	24.3 (24)
Rafael Nadal	**3932.6**	3967.6	.620	**3928.2**	3952.0	.445
Rafael Nadal →	-1880.9	-1860.3	41.3 (32)	-1883.2	-1867.5	31.4 (24)
Roger Federer	**3825.9**	3848.6	.125	**3814.5**	3831.1	.142
Roger Federer →	-2280.7	-2265.7	30.0 (32)	-2284.7	-2272.6	24.2 (24)
Novak Djokovic	**4625.4**	4659.4	.570	**4617.5**	4641.3	.448
Novak Djokovic →	-2403.9	-2389.0	29.9 (32)	-2411.7	-2399.0	25.2 (24)
Roger Federer	**4871.9**	4906.0	.572	**4871.3**	4894.1	.393
Rafael Nadal →	-1414.2	-1397.0	34.4 (32)	-1415.8	-1404.7	22.2 (24)
Novak Djokovic	**2892.4**	2922.0	.355	**2879.6**	2905.4	.565
Novak Djokovic →	-1302.1	-1289.7	24.7 (32)	-1304.5	-1294.3	20.3 (24)
Rafael Nadal	**2668.1**	2707.5	.819	**2656.9**	2684.6	.681
Novak Djokovic →	-1183.2	-1166.0	34.4 (32)	-1188.7	-1179.3	19.0 (24)
Andy Murray	**2430.3**	2459.9	.355	**2425.5**	2454.5	.753
Andy Murray →	-1280.1	-1263.7	32.7 (32)	-1287.9	-1277.1	21.6 (24)
Novak Djokovic	**2624.1**	2655.4	.431	**2623.9**	2650.2	.601
Pete Sampras →	-1117.9	-1103.5	28.6 (32)	-1124.1	-1111.9	24.4 (24)
Andre Agassi	**2299.7**	2335.1	.638	**2296.2**	2319.7	.437
Andre Agassi →	-1031.1	-1005.3	51.6 (32)	-1032.6	-1017.4	30.3 (24)
Pete Sampras	**2126.2**	2183.6	.015*	**2113.2**	2130.9	.174

Testing for Stationary POPs: by Calendar Year

Server → receiver	Muscle Memory			No Muscle Memory		
	Restricted	Unrestricted	LR test (df)	Restricted	Unrestricted	LR test (df)
	LL, AIC	LL, AIC	P-value	LL, AIC	LL, AIC	P-value
Roger Federer → Rafael Nadal	-1934.3 **3932.6**	-1901.9 3995.7	64.9 (64) .447	-1940.1 **3928.2**	-1916.9 3977.7	46.5 (48) .533
Rafael Nadal → Roger Federer	-1880.9 **3825.9**	-1843.1 3878.2	75.7 (64) .150	-1883.2 **3814.5**	-1853.9 3851.9	58.6 (48) .140
Roger Federer → Novak Djokovic	-2280.7 **4625.4**	-2242.0 4676.0	77.4 (64) .122	-2284.7 **4617.5**	-2256.9 4657.7	55.8 (48) .206
Novak Djokovic → Roger Federer	-2403.9 **4871.9**	-2363.9 4919.7	80.1 (64) .084	-2411.7 **4871.3**	-2383.3 4910.7	56.7 (48) .183
Rafael Nadal → Novak Djokovic	-1414.2 **2892.4**	-1402.0 2932.1	24.3 (32) .832	-1415.8 **2879.6**	-1408.5 2913.0	14.6 (24) .932
Novak Djokovic → Rafael Nadal	-1302.1 **2668.1**	-1280.7 2689.4	42.7 (32) .098	-1304.5 **2656.9**	-1285.9 2667.9	37.1 (24) .043*
Novak Djokovic → Andy Murray	-1183.2 **2430.3**	-1165.6 2459.3	35.0 (32) .326	-1188.7 **2425.5**	-1175.4 2446.8	26.7 (24) .317
Andy Murray → Novak Djokovic	-1280.1 **2624.1**	-1258.5 2645.1	43.0 (32) .092	-1287.9 **2623.9**	-1273.0 2641.9	30.0 (24) .186
Pete Sampras → Andre Agassi	-1117.9 **2299.7**	-1097.4 2322.9	40.9 (32) .135	-1124.1 **2296.2**	-1107.3 2310.5	33.7 (24) .091
Andre Agassi → Pete Sampras	-1031.1 **2126.2**	-1009.1 2146.2	44.0 (32) .077	-1032.6 **2113.2**	-1012.3 2120.6	40.6 (24) .019*

Results of Stationarity Tests

- For our early vs. late sub-groupings, we fail to reject stationarity at the 10% level for all server-receiver pairs in the specification without MM and for 9 of 10 pairs in the MM specification.
- For our calendar year sub-groupings, we fail to reject stationarity at the 5% level for all server-receiver pairs in the specification with MM and for 9 of 10 pairs in the no MM specification.
- For both the early vs. late and calendar year sub-groupings, the AIC selects the stationary (restricted) model over the non-stationary (unrestricted) model.
- We conclude that stationarity is a reasonable assumption for our data, which justifies grouping service games in order to get the most reliable possible estimates of serve direction probabilities and POPs.

Testing Equal Game Win Probabilities at All (x, m)

- We are finally ready to test for equal conditional win probabilities across serve directions!
- The p values for Wald tests that the *service game win probabilities* are equal across all serve directions at all (x, m) round to zero at many decimal places.
- The next slide presents a Wald test that the *point win probabilities* are equal across serve directions at all (x, m).

Testing Equal Point Win Probabilities at All (x, m)

Server $\rightarrow$ Receiver	Wald statistic	Degrees of freedom	P-value
Roger Federer $\rightarrow$ Rafael Nadal	405.4	29	5.9×10^{-68}
Rafael Nadal $\rightarrow$ Roger Federer	243.2	30	2.9×10^{-35}
Roger Federer $\rightarrow$ Novak Djokovic	23.6	30	.75
Novak Djokovic $\rightarrow$ Roger Federer	274.5	27	8.9×10^{-43}
Rafael Nadal $\rightarrow$ Novak Djokovic	83.5	29	3.5×10^{-7}
Novak Djokovic $\rightarrow$ Rafael Nadal	69.6	28	2.1×10^{-5}
Novak Djokovic $\rightarrow$ Andy Murray	52.3	30	0.007
Andy Murray $\rightarrow$ Novak Djokovic	212.0	30	2.7×10^{-29}
Pete Sampras $\rightarrow$ Andre Agassi	146.4	30	2.9×10^{-17}
Andre Agassi $\rightarrow$ Pete Sampras	198.6	30	8.9×10^{-27}

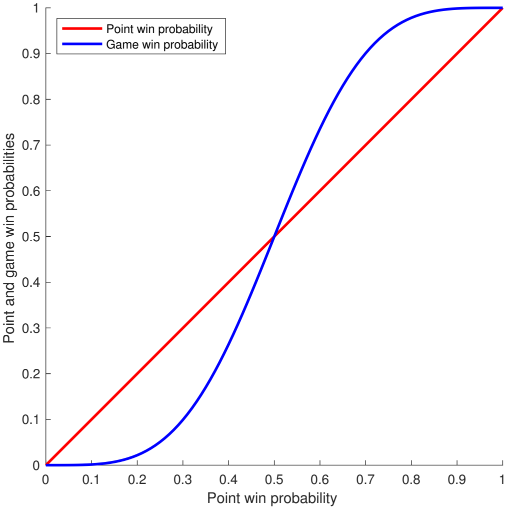
Small Sample Issues with our Wald Tests

- It is known that Wald tests of equality restrictions can lead to spurious rejections of the null when the number of restrictions being tested is large relative to the number of parameters in the model.
- This is potentially a problem in our case, since our reduced-form model of the server's strategy and the POPs has 44 parameters, while we have 596 equality restrictions (across all score-muscle memory states).
- More strongly: we do get over rejection when we apply our Wald test to simulated data for which conditional win probabilities are equal.
- To overcome this issue we perform Wald tests of equal service game win probabilities on subsets of the score-muscle memory state space, so that we are testing far fewer than 44 equality restrictions.

Testing Equal Game Win Probabilities, Some (x, m)

Server → Receiver	4 fixed serve strategies at 3 states, 9 df	RF serve strategy at 4 states, 12 df
Roger Federer → Rafael Nadal	1.4×10^{-11}	.605
Rafael Nadal → Roger Federer	.873	.018
Roger Federer → Novak Djokovic	6.5×10^{-30}	1.6×10^{-68}
Novak Djokovic → Roger Federer	.0009	.220
Rafael Nadal → Novak Djokovic	2.2×10^{-254}	.526
Novak Djokovic → Rafael Nadal	4.0×10^{-91}	4.5×10^{-44}
Novak Djokovic → Andy Murray	1.4×10^{-69}	.018
Andy Murray → Novak Djokovic	9.1×10^{-91}	.00001
Pete Sampras → Andre Agassi	.787	.003
Andre Agassi → Pete Sampras	.764	.667

Point vs. Game Win Probabilities



Dynamic Structural Model of Tennis

- Our flexible agnostic model of tennis decisively rejects the key implication of a mixed strategy Nash equilibrium: namely that the probability of winning the game is the same regardless of serve direction.
- To get more insight into *why* this might be, we estimate three different structural models of serve behavior.
- The three models differ in how far the server “looks ahead”:
  - a *fully-dynamic model* that assumes the server maximizes the probability of winning the *service game;*
  - a *point-myopic model* that assumes the server maximizes the probability of winning each *point;*
  - a *serve-myopic model* that assumes the server maximizes the probability of winning each *serve.*

Recovering Subjective Beliefs of the POPs

- We estimate structural discrete choice models of the mixing probabilities over serve directions $P(d|x, m)$. A key feature of our structural approach is that we relax the assumption of *rational expectations* about the POPs, Π.

- Instead we are able to estimate a tennis server's *subjective beliefs* of the POPs. One explanation of the suboptimality of serve strategies is that tennis pros have subjective beliefs about their own strengths and weaknesses, and those of their opponent, that do not fully square with reality.

Incorporating "Trembles" That Affect Serve Directions

- We assume that at the moment each serve is made, the server's choice of direction reflects *trembles.*
- These *IID* shocks affect their *perception* of the probability of winning when serving to different directions d.
- We assume that these trembles or preference shocks are observed only by the server but not by the receiver or the econometrician. For example, the server may feel more comfortable hitting to a certain direction at some point in the match due to a psychological factor.
- Let $\epsilon(d)$ be the tremble associated with serving to direction d. We assume the trembles are independently distributed across all three possible serve directions $\{l, b, r\}$ and *IID* across successive serves, and have a Type 1 extreme value distribution with location parameter normalized so that $E\{\max_d \epsilon(d)\} = 0$ and scale parameter $\lambda \geq 0$.

Serve Strategies with Trembles

- Let $\sigma_{FD}(x, m, \epsilon)$ be the serve strategy under the fully-dynamic model as a function of the observed state (x, m) and the unobserved trembles $\epsilon = (\epsilon(l), \epsilon(b), \epsilon(r))$.
- The fully-dynamic model presumes that for each (x, m, ϵ) the server chooses the serve direction that maximizes the probability of winning the game, given by

$$\sigma_{FD}(x, m, \epsilon) = \underset{d \in \{l,b,r\}}{argmax}[\lambda\epsilon(d) + V_\lambda(x, m, d)]$$

where the value function $V_\lambda(x, m, d)$ is the analog of the conditional win probability $W_S(x, m, d)$, given by

$$V_\lambda(x, m) = \lambda \log\left(\sum_{d \in \{l,b,r\}} \exp\{V_\lambda(x, m, d)/\lambda\}\right).$$

Structural Serve Direction Probabilities

- The serve direction probabilities implied by the fully-dynamic model, denoted by $P_{FD}(d|x,m)$, is given by

$$P_{FD}(d|x,m) = Pr\left\{d = \sigma_{FD}(x,m,\epsilon)|x,m\right\} = \frac{\exp\{V_\lambda(x,m,d)/\lambda\}}{\sum_{d' \in \{l,b,r\}} \exp\{V_\lambda(x,m,d')/\lambda\}}.$$

- Though $\sigma_{FD}(x,m,\epsilon)$ is a pure strategy from the standpoint of the server, it appears to be a mixed strategy from the standpoint of someone who does not observe ϵ. This allows us to rationalize observed mixed serve strategies without imposing equal win probabilities, i.e. imposing equality of $V_\lambda(x,m,d)$ over serve directions d.

Convergence to Mixed Serve Strategies

- Since the trembles are *IID* across serves, it would appear that this model should also imply conditional independence in serve directions across successive first and second serves. However, that will actually only be true if muscle memory m does not enter $V_\lambda(x, m, d)$.

- The following limit holds uniformly over (x, m, d) as $\lambda \downarrow 0$:

$$W_S(x, m, d) = \lim_{\lambda \downarrow 0} V_\lambda(x, m, d)$$

so the only way for $P_{FD}(d|x, m)$ to converge to a mixed strategy as $\lambda \downarrow 0$ is when conditional win probabilities $W_s(x, m, d)$ obey the equal win probability constraints,

$$W_S(x, m, l) = W_S(x, m, b) = W_S(x, m, r) \quad \forall (x, m).$$

Myopic Structural Models of Serve Behavior

- The point-myopic and serve-myopic models have the same general structure as the fully-dynamic model.
- Thus, the serve strategies, value functions, and choice (mixing) probabilities are given by the same equations as above, except for how the value function V_λ is defined.
- In the serve-myopic model, the server's value is the probability of winning each serve

$$V_\lambda(x, m, d) = \pi(\text{in}|d, x, m)\pi(\text{win}|d, x, m),$$

Point-Myopic vs. Serve-Myopic Models

- The point-myopic server's objective is to win each *point,* but the server does recognize the option value provided by the second serve in the event of a faulted first serve.
- This requires a two period backward induction calculation. If x is a second serve state, then $V_\lambda(x, m, d)$ coincides with the serve-myopic value function given above.
- However, for any first serve, V_λ is given by

$$V_\lambda(x, m, d) = \pi(\text{in}|d, x, m)\pi(\text{win}|d, x, m) + [1 - \pi(\text{in}|d, x, m)]V_\lambda(x + 1, m')$$

  where $m' = (d, d_1)$ and $V_\lambda(x, m)$ is the the expected maximum of the serve-myopic values given above.
- The point-myopic serve strategy coincides with the fully-dynamic serve strategy in the limit as $\lambda \downarrow 0$ when the GMC (winning a point is always better than losing it) holds.

Comments on the Structural Models

- Note that all three structural models have mixed serve probabilities that are implicit functions of the POPs. Thus, the mixed serve directions for the three structural models are entirely determined by the POPs and the single parameter λ controlling the magnitude of the trembles.
- In contrast, the reduced-form model of serve directions is estimated separately from the POPs with flexible parameterization of serve directions.
- The structural models can be viewed as restricted special cases of the most flexible specification of the reduced-form model. This enables us to conduct likelihood-ratio specification tests for the three structural models relative to the unrestricted reduced-form specification.

Structural Estimation Results

Player pair	**Reduced-Form**	**Serve-Myopic**	**Point-Myopic**	**Fully-Dynamic**
Server $\rightarrow$ receiver	LL, N BIC	LL, $\hat{\lambda}$ AIC, LR P-value	LL, $\hat{\lambda}$ AIC, LR P-value	LL, $\hat{\lambda}$ AIC, LR P-value
Roger Federer $\rightarrow$ Rafael Nadal	-3779.1, 2011 7646.1	-3788.2, 7.9×10^{-4} 7642.7, .074	-3783.8, 5.8×10^{-3} **7633.7, .571**	-3817.3, 1.1×10^{-4} 7700.7, 6.9×10^{-12}
Rafael Nadal $\rightarrow$ Roger Federer	-3569.1, 1882 7226.2	-3571.3, 6.1×10^{-3} 7208.6, .957	-3570.6, 2.7×10^{-3} **7207.3, .990**	-3632.4, 2.2×10^{-4} 7330.7, 8.8×10^{-22}
Roger Federer $\rightarrow$ Novak Djokovic	-4545.8, 2333 9179.5	-4551.2, .010 **9168.4, .457**	-4552.2, 4.5×10^{-3} 9170.4, .300	-4576.0, 9.1×10^{-4} 9128, 7.5×10^{-9}
Novak Djokovic $\rightarrow$ Roger Federer	-4827.7, 2372 **9743.5**	-4840.0, .011 9746.0, .010	-4842.0, 1.9×10^{-3} 9750.0, 2.6×10^{-3}	-4844.8, 2.4×10^{-4} 9755.7, 3.3×10^{-4}
Rafael Nadal $\rightarrow$ Novak Djokovic	-2846.8, 1405 5781.7	-2853.8, 1.1×10^{-4} 5773.7, .232	-2853.2, 8.2×10^{-5} **5772.4, .310**	-2864.5, 1.2×10^{-5} 5795.0, 2.2×10^{-4}
Novak Djokovic $\rightarrow$ Rafael Nadal	-2649.5, 1344 5387	-2659.9, .070 5385.9, .035	-2656.1,.097 5378.2, .285	-2656.7, 1.7×10^{-5} **5375.3, .505**
Novak Djokovic $\rightarrow$ Andy Murray	-2384.7, 1201 4857.5	-2396.2, 9.9×10^{-3} **4858.3, .018**	-2396.9, .044 4859.8, .011	-2413.0, 5.5×10^{-4} 4892.0, 4.0×10^{-8}
Andy Murray $\rightarrow$ Novak Djokovic	-2649.5, 1328 5387.0	-2536.4, .014 **5138.9, .309**	-2539.8, 6.9×10^{-3} 5145.7, .051	-2556.3, $9.7 \times 1-^{-5}$ 5350.0, 2.2×10^{-7}
Pete Sampras $\rightarrow$ Andre Agassi	-2203.3, 1181 **4494.6**	-2219.6, .031 4505.3, 5.9×10^{-4}	-2217.7, .037 4501.4 2.5×10^{-3}	-2240.2, 1.9×10^{-6} 4546.4, 2.3×10^{-11}
Andre Agassi $\rightarrow$ Pete Sampras	-1962.9, 1050 4013.8	-1973.0, 1.8×10^{-3} 4011.9, .043	-1970.8, 1.9×10^{-5} **4007.6, .145**	-2004.7, 5.7×10^{-5} 4075.5, 2.8×10^{-13}

Summary of Structural Estimation Results

- The maximum likelihood estimates of the POPs $(\hat{\theta}_{in}, \hat{\theta}_{win})$ are distorted in a manner that results in conditional win probabilities much closer to equality than the ones implied by the reduced-form estimates of the POPs.
- The best model selected by the AIC is generally also that for which there is the least evidence for rejecting it in favor of the reduced-form model via the likelihood-ratio test.
- The AIC selects the point-myopic model as the best model for four of the servers, and it selects the serve-myopic model for three of them. In two cases, Djokovic serving to Federer and Sampras serving to Agassi, it selects the reduced-form model. It selects the fully-dynamic model in only one case, Djokovic serving to Nadal.
- Bottom Line: Servers seem to be correctly solving the point-maximization problem, except they have incorrect beliefs about the POPs.

Calculating Best-Response Serve Strategies

- We conclude by providing a more powerful direct test of Nash equilibrium play in tennis: we construct alternative *deviation* serve strategies that significantly increase a server's probability of winning the service game compared to the mixed strategy they are currently using.
- If the hypothesis of Nash equilibrium is correct, it should be impossible to construct any such deviation strategies. We construct deviation strategies using numerical DP, so they are pure strategies.
- The DP serve strategies exploit the unequal win probabilities captured by our reduced-form estimates of the POPs. At each stage of the game, the DP strategies choose the serve direction that has the maximum conditional win probability.

Direct Test of Nash Equilibrium

- Our test of Nash equilibrium serve strategies appeals to the *one shot deviation principle*.
- We find that there are profitable one shot deviations at many stages of the game.
- While each such deviation yields a modest improvement in the win probability, the cumulative effect of all profitable deviations is often a large improvements in the overall game win probability.
- Caveat: If a server were to switch to the optimal serve strategy we estimate, the receiver may adjust their own strategies, likely mitigating the gain we estimate.

Shortcoming of Direct Test of Nash Equilibrium

- One key shortcoming of our approach to testing the hypothesis of Nash equilibrium: we only have *estimates* of the POPs rather than the *true* POPs.
- Estimation error in the POPs could result in spurious, upward-biased, estimates of the win probability when we use a noisy estimate of the POPs to calculate a best response strategy via DP instead of using the true POPs.
- Basic Idea: Fix a given serve strategy (e.g., RF) and simulate the win probabilities vs. the distribution of the POPs implied by our RF estimates.

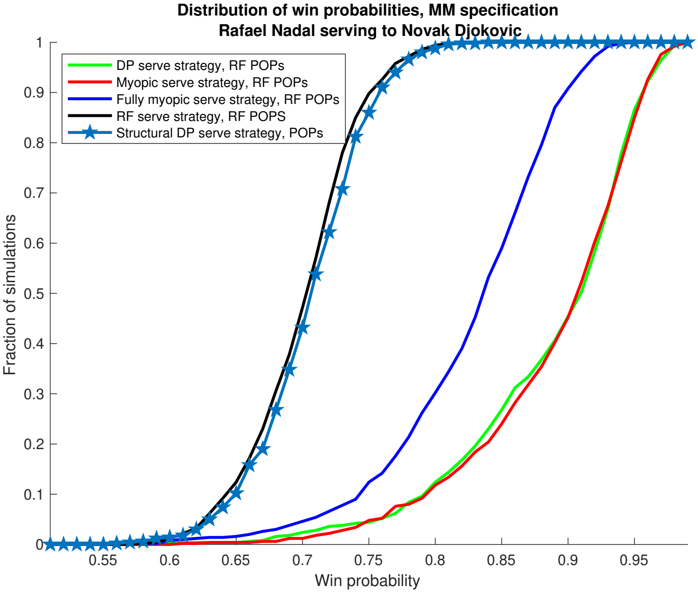
Test of Nash Equilibrium Robust to Estimation Error

- We use the implied asymptotic distribution of our maximum likelihood parameter estimates to calculate an approximate probability distribution for the true POPs based on the data we observe.
- Specifically, the true POPs are asymptotically distributed about the maximum likelihood estimate of the POPs according to the normal asymptotic distribution of the reduced-form POP parameters $(\hat{\theta}_{in}, \hat{\theta}_{win})$.
- Via stochastic simulation, we can draw from the distribution of the true POPs by drawing values of $(\tilde{\theta}_{in}, \tilde{\theta}_{win})$ from an asymptotic normal distribution centered at the MLE $(\hat{\theta}_{in}, \hat{\theta}_{win})$ with a covariance matrix equal to the asymptotic covariance of the MLE.
- This gives us a probability distribution over the true POPs that the server might actually be facing.

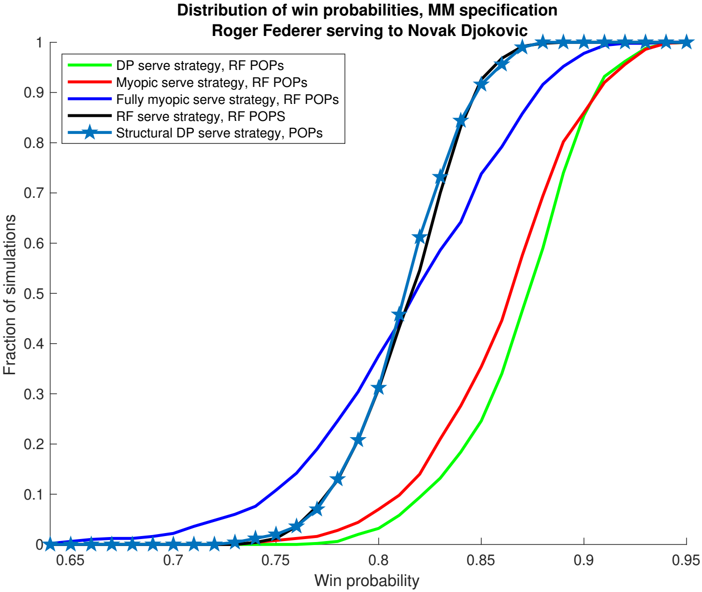
Average Simulated Improvements in Win Probabilities

Player pair	Reduced-Form	Serve-Myopic	Point-Myopic	Fully-Dynamic
Server $\rightarrow$ receiver	$W(1,1)$	$W(1,1)$ P-value	$W(1,1)$ P-value	$W(1,1)$ P-value
Roger Federer $\rightarrow$ Rafael Nadal	.821 (.035)	.850 (.037) 0	.888 (.029) 0	.890 (.028) 0
Rafael Nadal $\rightarrow$ Roger Federer	.798 (.045)	.830 (.052) 0	.870 (.049) 0	.867 (.048) 0
Roger Federer $\rightarrow$ Novak Djokovic	.810 (.028)	.816 (.049) 0	.865 (.036) 0	.869 (.033) 0
Novak Djokovic $\rightarrow$ Roger Federer	.776 (.027)	.843 (.038) 0	.856 (.032) 0	.861 (.034) 0
Rafael Nadal $\rightarrow$ Novak Djokovic	.704 (.042)	.830 (.059) 0	.888 (.071) 0	.886 (.073) 0
Novak Djokovic $\rightarrow$ Rafael Nadal	.838 (.029)	.929 (.023) 0	.908 (.043) 0	.931 (.022) 0
Novak Djokovic $\rightarrow$ Andy Murray	.780 (.040)	.890 (.034) 0	.893 (.031) 0	.893 (.032) 0
Andy Murray $\rightarrow$ Novak Djokovic	.709 (.045)	.833 (.056) 0	.858 (.057) 0	.858 (.062) 0
Pete Sampras $\rightarrow$ Andre Agassi	.857 (.031)	.921 (.052) 0	.940 (.028) 0	.939 (.027) 0
Andre Agassi $\rightarrow$ Pete Sampras	.843 (.035)	.895 (.059) 0	.921 (.041) 0	.920 (.042) 0

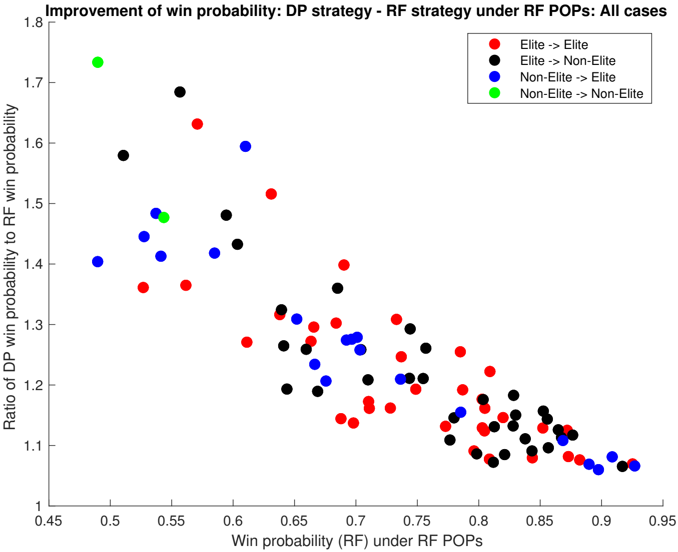
Distributions of Win Probs: Nadal serving to Djokovic



Distributions of Win Probs: Federer serving to Djokovic



Expected Gains Across Many Pairs



Conclusions

- Suppose you are convinced by our analysis that tennis pros don't play minimax. Why should you care? After all tennis is "just a game."
- One reason to care is that Nash equilibrium is a key part of the foundation not only of economic theory, but much of our empirical work as well, especially in industrial organization.
- To the extent that firms face vastly harder DP and equilibrium problems than tennis players face, one might worry that assuming rational Nash behavior when this assumption is not true could seriously distort our empirical work and policy conclusions.