



Unforced Errors: Tennis Serve Data Tells Us Little About Loss Aversion

Michał Krawczyk¹

[LINK TO ABSTRACT](#)

In two recent papers, Nejat Anbarci and collaborators (Anbarci, Arin, Okten, and Zenker 2017; Anbarci, Arin, Kuhlenkasper, and Zenker 2018)—I refer to the papers henceforth as A17 and A18—argue that male tennis players’ behavior is consistent with aversion to losses rather than maximization of expected utility. To prove this point they sketch simple theoretical models. They propose that when a player is behind in score more effort will be exerted than when ahead if and only if losses loom larger than gains. They then find that serve speeds of male participants in one tournament are indeed higher (which arguably requires more effort) when the server is behind, for example is losing 0–1 in sets.

However interesting the case may be, I believe the analysis is flawed for three main reasons: First, they implicitly assume that under expected utility, starting from a tied score, losing a game or a set or a match *should* be as bad as winning it is good. That would mean that if we find that such losses in fact loom larger than such gains, this may be seen as loss aversion. However, as I show in the next (second) section, this assumption is clearly incorrect, when applied to losing or winning games at any rate. Also, they interpret greater serve speed in terms of greater effort. In the

1. University of Warsaw, 00-927 Warsaw, Poland. I gratefully acknowledge the advice I received from Adam Romer, Editor-in-Chief of Poland’s leading tennis magazine *Tenisklub*, which was particularly instrumental in the development of the argument provided in the “Stronger serve...” section. I have also greatly benefitted from comments and suggestions made by M. Daniele Paserman, Joanna Tyrowicz, and Łukasz Woźny. Dominika Zychowicz provided research assistance. Still, all the errors are mine, and the above-mentioned individuals do not necessarily endorse the opinions expressed here.

third section I argue that, counterintuitively, a stronger serve ultimately means *less* total effort. Lastly, in their theoretical analysis, they disregard behavior of the returning player (a problem discussed here in the fourth section) and make other important mistakes (the fifth section). Overall, I do not believe that Anbarci and coauthors have shown that the notion of loss aversion is useful in understanding tennis servers' behavior.

Server's lost games *should* loom larger than won games

Consider a single game. The same player keeps serving throughout the game, until one player wins four points while the other has at most two or until one player wins $k > 4$ points while the other has only $k - 2$ points. The confusing convention apparently stemming from medieval France is that, upon winning the first point, the score is given as "15:0" rather than "1:0," followed by "30:0" and "40:0" (of course assuming the server keeps winning). The score of 40:40, and any subsequent tie, is called *deuce*; one point above that is called *advantage*. Suppose now a game is currently tied at 30:30 or deuce, which are strategically equivalent (A17 and A18 call this state t , for tied). Eventually the game may be won (state w) or lost (state l) by the server. A17 normalize their value function such that $V(w) = 1$ and $V(l) = 0$. They then claim that if $V(t) = -\lambda$ is lower than -1 , the player exhibits loss aversion (because losing is more painful than winning is enjoyable). As I show below, however, such a valuation is fully consistent with the standard expected utility model. Because eventually the game will be won (value 1) or lost (value $-\lambda$), then the value of t depends on the server's probability of ultimately winning the game, conditional on the game being tied now, say, p :

$$V(t) = p + (1 - p)(-\lambda)$$

Now, p is *not* 50 percent in tennis. It is considerably higher because the server has the upper hand—it is easier to win a point when serving. Therefore setting $V(t) = 0$ immediately implies that $\lambda > 1$. For example, for $p = 2/3$, we have $\lambda = 2$.

I do not have statistics for the fraction of games that are ultimately won for each current score. One may, however, easily do the following exercise.² First, using data from TennisAbstract.com or elsewhere, calculate the overall fraction of games won by the server. It turns out that on average, it is about 75 percent

2. A more sophisticated way of calculating each point's *importance*, i.e., how it changes the probability of ultimately winning the match, is used by Paserman (2010).

for males and 63 percent for females (whose serves tend to have lower velocity).³ Second, basing on these statistics and assuming, as a handy approximation, that the server's probability of winning each point is identical and independent, one may calculate the probability of ultimate success for each current score. For example, starting from 30:30, the server will win the game with probability 71 percent in male tennis and probability 60 percent in female tennis (note that these will always be subtly different from those for other tied scores, 0:0 for example, but A17 seem to treat all tied scores jointly, which would be a mistake). It is only logical then, that, starting from a 71 percent chance of success, a loss looms much larger than a gain. This consideration not only explains the effect that A17 wrongly call loss aversion but also predicts that it be stronger for males—as they indeed find.

Analogous analysis can account for, as A17 put it, the “diminishing sensitivity” in score difference such that incremental gains in $(T_s - T_r)$ above the reference point, that is, the tied score, result in progressively smaller utility improvements” (A17, 3551). For example, in a typical male tennis match the increase in probability of eventually winning the game when one moves, say, from 30:0 to 40:0 is very low—less than five percentage points—because the chance of winning is very high already at 30:0. Again, contrary to what the authors seem to imply, this has nothing to do with risk preference and all to do with the fact that (by far) not all the points in a tennis match are equally important in terms of the ultimate success. Maximization of expected value is a reasonable benchmark for decisions with uncertain outcomes expressed in money, because (for a price taker) every dollar can buy the same basket of goods. In tennis, the goal is to win the match, not to maximize the number or the fraction of points won,⁴ so clearly not all points are equally important: there is no reason to purport, for example, that the utility difference between 15:0 and 30:0 is the same as the utility difference between 30:0 and 40:0. It seems therefore much more fruitful to analyze players' choices not in terms of utilities of intermediate goals such as winning or losing a point, but in terms of probabilities of the ultimate success (winning the match). Such a strategy was employed by M. Daniele Paserman (2010).

Similar reasoning explaining apparent “loss aversion” and “diminishing sensitivity” applies to the score in games and sets. To illustrate the former effect, note that being the first to serve in a set has been reported to raise the probability of winning this set to 54 percent in males and 53 percent in females (Jensen 2014). Thus if each player has won the same number of games within a set so far (e.g.,

3. The numbers depend on the surface but these overall averages should be about right for the hard courts of the Dubai tournament that Anbarci and coauthors analyze.

4. For example, between 1990 and 2014, Roger Federer lost 24 matches in which he scored more points than the opponent (Rodenberg 2014).

zero), losing the set should rationally loom slightly larger to the currently serving player (i.e., the one who served on the very first game of the set) than winning it, because $54\% > 100\% - 54\%$. To illustrate “diminishing sensitivity,” consider a player currently losing 0–4 in games. The chances to win this set are already extremely slim, so a perfectly rational expected utility maximizer will care very little about losing another game *within the same set*. By contrast, losing a game at 2–2 is painful.

Stronger serve means *less*, not more, total effort

It would seem an obvious manifestation of the laws of physics that sending the ball with a greater speed generally requires more energy. However, the rally does not usually stop there and it seems prudent to assume that experienced, well-paid professionals have a planning horizon which is longer than one or two seconds. Now, there are two main consequences of a stronger serve. First, it is less precise, so there is a greater risk that it fails to go in. Second, it leaves the opponent less time to react, so that she is less likely to successfully return (O’Donoghue and Ballantyne 2004), which is of course why tennis players hit the ball so hard. Because of both of these effects, the rally is expected to be considerably shorter after a stronger serve—in all probability *reducing* the server’s total effort. I was not able to find direct verification of this claim in existing literature,⁵ but there is a lot of corroborative evidence. For example, playing on grass—the fastest surface—results in shorter rallies, both in terms of the number of strokes and duration (Hughes and Clarke 1995). Likewise, males, who serve more strongly on average, generally have shorter rallies than females (O’Donoghue and Liddle 1998).

I did some back-of-the-envelope calculations based on data available at TennisAbstract.com for two matches, one female and one male, played in the 2013 Dubai tournament that A17 and A18 focused on, namely Samantha Stosur vs. Ekaterina Makarova and Roger Federer vs. Malek Jaziri.⁶ The website provides the fraction of rallies won by the server conditional on the rally being at least one, two, etc. strokes long, separately for the first and second serve. From this data, one can calculate the average number of strokes after the first, stronger, serve vs. after the second, weaker, serve. While player-specific numbers differ a lot, for all four players the second serve results in longer rallies. On average, the difference in the

5. Although Paserman in a personal communication said he remembered that indeed there was clearly such a pattern in the data used in his 2010 paper.

6. This data is provided in a way that makes processing it quite cumbersome but in principle the same exercise could be repeated for many games.

number of times the server had to hit the ball after the second serve and after the first serve equalled 0.85.

Now, W. Ben Kibler (2009), basing on the work of Richard Schönborn (2000), states that the effort involved in a tennis serve is comparable to that of one forceful groundstroke. If by hitting the ball somewhat harder on the serve one can considerably reduce subsequent number of groundstrokes, this appears clearly beneficial from the viewpoint of overall effort, precisely contrary to the authors' operationalization of effort.⁷ In particular, if players were to hit the ball on the second serve as strongly as they do on the first serve, they would save on effort.

As a side note, modern-day professionals are superbly fit athletes who can generally endure two or three sets of top performance. Thus in most matches they do not try to economize on effort but simply give their best in every point.⁸ The observed non-trivial within-match within-player variance of the (first) serve speed may be best explained in terms of random variations of performance and the use of mixed strategies (the receiver is in a more difficult situation when she cannot be sure if the next serve will be as strong as possible vs. somewhat weaker but heavily rotated, etc.).

The receiver is also there

Of course, a stronger serve may also affect the expected effort on part of the receiver. As for me, for example, I would probably not manage to exert any effort whatsoever if the opponent served at 150 mph (except for the anxious heartbeat that is). Reducing the opponent's effort may well be a bad thing for the server, but the authors do not even try to consider the opponent's effort.

Most importantly, the authors implicitly assume that the receiver's effort does not depend on current score. This is a very questionable assumption. *A priori*, there is not a clear reason why she should respond to it any less than the server does. Moreover, as it is easy to see, the cross elasticity of server's probability of winning (A17, 3550, eq. 1; A18, 4, eq. 2) with respect to the efforts of the two players does not vanish, so the server *should* adjust her effort in part in response to the anticipated change in the receiver's effort. In other words, disregarding the changes in receiver's effort most probably leads to incorrect conclusions concerning server's optimal effort. Unfortunately, the authors do not provide any relevant game-

7. It should be noted that although Anbarci and coauthors generally interpret greater serve speed in terms of greater effort, they at times seem to cavalierly switch to interpreting it as greater risk acceptance instead, for example in their Result 3 of A17 (p. 3552).

8. Exceptions will include very long tournaments, very long matches, low-stakes matches, matches with a clear dominator, and matches involving older or injured players.

theoretic analysis and do not have any data on receiver's effort. Again, the analysis of Paserman (2010) seems superior here. As a side note, Paserman also uses a larger sample and obtains different results, a fact upon which A17 and A18, surprisingly, do not comment in any way.

Note that all of the considerations of the second, third, and fourth sections here are strongly discipline-specific. A17 and A18 seem to try to partly build credibility for their approach on its similarity to that of Devin G. Pope and Maurice E. Schweitzer (2011), a well-known study published in a prestigious journal, by citing it 15 times in total and emulating its catchy title. Now, Pope and Schweitzer (2011) analyze *golf*; the only similarity with tennis seems to be that it involves millionaires trying to hit a ball of some sort. There are at least three important features of golf that make it quite different from tennis, in a manner relevant for our analysis: there is no interaction with other players; only the total score (number of strokes) matters, so natural reference points at individual holes (i.e., "pars") are purely symbolic; and it is often the case that a player needs to exert a lot of (mental) effort to reach the hole in one putt and no effort at all to make it in two. All of these render the approach of Pope and Schweitzer (2011) one that makes perfect sense for golf but useless for tennis.

Remaining issues and conclusion

It should also be noted that the reading of both A17 and A18 is obstructed by various mistakes and omissions in their theoretical analysis. For example, the first form of the right-hand side of equations 3a and 3b in A17 (p. 3550) is obviously incorrect as the probabilities of winning and losing do not add up to one. It is also not clear whether W stands for (1) any advantageous score, e.g., 30:15, or (2) only the case in which winning the game is just one point away, e.g., 40:15, or (3) only the case in which both winning and tie are just one point away, precisely 40:30 or the strategically equivalent advantage. I sent the authors an email asking for clarification and was informed that interpretation (2) was correct. When I pointed out that this is inconsistent with their equations 3a and 3b, I was told that interpretation (3) was correct. Yet, their Result 1 seems to use interpretation (1)! Similar interpretational obscurity may trouble the readers of A18.

Both in A17 (p. 3551) and A18 (p. 5) the authors also define $\Delta_s = T_s - T_r$ and $-\Delta_r = T_r - T_s$. It must thus be the case that $\Delta_s = \Delta_r$, so it is not clear why they need to define both of them. The authors also claim (in A17) that

$$\Delta_s = -\Delta_r = \Delta \text{ if and only if } T_s - T_r = -(T_r - T_s).$$

The right-hand side of the equivalence is trivially always satisfied (so it is not clear why this condition is mentioned at all). The left-hand side is meaningless because Δ is not defined. Possibly this is exactly where the authors want to define Δ (that's what they do in A18, where they also remove the minus sign in $\Delta_s = -\Delta_r$) but then it would not be clear how it is defined if $\Delta_s \neq -\Delta_r$. Moreover, $\Delta_s = -\Delta_r$ will in general not be true, while the second part of the equivalence is always true, as mentioned before, so the equivalence is not only useless but also incorrect.

The new theoretical elements in A18 leave the reader even more confused. In equation 6 (A18, p. 5) expected utility “[w]hen it is a game (or set or match) point favoring the server,” is defined as

$$P_s V(W) + (1 - P_s) V(l) - c,$$

whereby P_s is the probability of winning a point, undefined $V(W)$ probably denotes utility after the game (or set or match) has been won and $V(l)$ means the utility when the game (or set or match) is currently tied, and c stands for the cost of effort. This seems misconceived, because the game (or set or match) is usually neither over nor tied after just one additional point in tennis. In equation 9 (A18, p. 6) the authors seem to plug in $(-\lambda)$ for $V(l)$, although $V(l)$ is defined to be equal to something else in equation 4 (A18, p. 5). The authors also might have inadvertently skipped Δ_r ; this guess is corroborated by the fact that it reappears when first derivatives are taken in equation 10 (A18, p. 6). Analogous mistakes seem to have been made in equations 6, 7, and 8. The second epsilon in equation 6 as well as second and third in equation 8 should be preceded with a minus sign. It is not clear if it is on purpose that the l is not capitalized whereas W in an analogous position is capitalized (A18, 5–6).

Whereas in preparation of the Propositions, the authors speak in terms of effort when serving, Propositions 1 and 3 concern server's “risk averse serves” (A18, 6–8). It is not clear if “more risk averse serves” are supposed to be the same thing as serves involving less effort. No proof is provided to any of the Propositions anyway. The authors ask the reader to “[o]bserve that when $\Delta_s = \Delta_r \equiv \Delta$, the first-order conditions still imply...” (A18, 6), which is confusing because there is no Δ_r term in the first-order conditions anyway. And on and on.

Inappropriate treatment of the distinction between the first and the second serve is also a major problem. This distinction, empirically, makes by far the largest impact on within-player variation in serve speed; note, e.g., the huge estimate for the second-serve dummy in the regressions. Yet, astonishingly, it plays no role at all in their theoretical analysis (unlike in other economic analysis of tennis, e.g., Klaassen and Magnus 2009)! If conservation of effort was the main driver behind players' preferred serve speed, then one would expect the speed difference between

the first and the second serve to be drastically reduced when the possibility of winning or losing the match draws near: there is little reason to economize on effort at this point. Yet I was not able to find any hint that that this is indeed observed. More generally, the first and second serve are so different that clearly interactions with other variables should be considered.

The empirical analysis is also troubled by omission of crucial variables. For example, the rules of tennis stipulate that after an even number of rallies within a game, in particular when the score is tied, the server must be placed to the right of the center mark and must hit the “deuce court.” The opposite is true after an odd number of rallies, in particular when the score is 40:30, 30:40, or advantage: the “ad court” must be targeted. This has important consequences for some players at least, Roger Federer included (Hodgkinson 2016). The side of the court would be important even for players with completely symmetric serving skills, as, other things being equal, it is generally beneficial to target opponent’s backhand; players do so particularly often in the second serve (Rioult et al. 2015). Obviously, targeting the backhand requires playing down the T (near the center of the court) at deuce and serving wide (near the side line) at advantage, a distinction that can also mean different optimal serve speed. Of course, the situation is reversed if the receiver is left-handed, yet another crucial factor neglected by A17 and A18.

Likewise, they could have easily controlled for the number of rallies since the beginning of the match, as a proxy for the time lapsed and for fatigue. Clearly, this variable can be expected to affect the speed of serve *and* to be (strongly) correlated with the set score dummies (D2, D3, and D4). The estimates on these variables in the current models are thus biased. Without controls for time and fatigue, by the way, the most telling is the comparison between the scores 0–1 and 1–0 in sets, not between 0–1 (or 1–0) and 0–0. It would thus seem more important for the interpretation of the data that the coefficients on D2 and D3 are apparently not significantly different from each other, not that D2 or D3 is significantly different from zero (i.e., from the baseline of 0–0 in sets).

Finally, they do not explicitly include a dummy for being the first player to serve in the current set. This is surprising, because in the first set of the match this would represent a useful exogenous randomization device—the winner of a coin toss serves in the first game, so that the other player is more likely to be behind in games during this set (“in the domain of losses”). *Nota bene*: If, in line with the authors’ interpretation of loss aversion, this leads to more effort being taken by the loser of the coin toss, we should expect that she is *more* likely to win the first set. In fact, the opposite is true, as mentioned before (Jensen 2014). By contrast, who is the first server in the *second and subsequent* sets is *not* exogenously randomized—it is the player who was not serving in the last game of the previous set, so more often than not it is the loser of the previous set. Omitting this variable is thus likely to lead

to biased estimates on the set score dummies (D2, D3, and D4), at least in models 1 and 4 of A17 where the difference in games is not controlled for. For example, if D2, the dummy indicating the 0–1 set score, has a positive coefficient, it could be because the player who is the first one to serve in the (second) set tends to serve faster balls.

To summarize, the theoretical model suffers from several deficiencies; the empirical analysis disregards variables which are extremely important for tennis tactics; interpretation of the key variable is highly doubtful; and the results obtained are inconsistent with previous, more thorough analysis.

Overall, Anbarci et al. have not provided evidence that Roger Federer (or anybody else) is loss averse (although he may well be). Would such papers have been accepted if their conclusion was that there was no loss aversion after all? Or is it a case for the “bias bias in behavioral economics” (Gigerenzer 2018)? Is it enough to place Serena Williams and loss aversion in the title and some gender differences in the abstract and conclusion to compensate for deficiencies of the analysis? That would certainly not serve the scientific community well.

References

- Anbarci, Nejat, K. Peren Arin, Torben Kuhlenkasper, and Christina Zenker.** 2018 (A18). Revisiting Loss Aversion: Evidence from Professional Tennis. *Journal of Economic Behavior and Organization* 153: 1–18.
- Anbarci, Nejat, K. Peren Arin, Cagla Okten, and Christina Zenker.** 2017 (A17). Is Roger Federer More Loss Averse than Serena Williams? *Applied Economics* 49(35): 3546–3559.
- Gigerenzer, Gerd.** 2018. The Bias Bias in Behavioral Economics. *Review of Behavioral Economics* (Now Publishers Inc., Boston) 5(3–4): 303–336. [Link](#)
- Hodgkinson, Mark.** 2016. Roger Federer: The Strengths of a Tennis Genius in Numbers. CNN.com, June 29. [Link](#)
- Hughes, M., and S. Clarke.** 1995. Surface Effect on Elite Tennis Strategy. In *Science and Racket Sports*, ed. Thomas Reilly, Mike Hughes, and Adrian Lees, 272–277. London: E & FN Spon.
- Jensen, Kenneth.** 2014. Keys to the Match: In Tennis, Is It an Advantage to Serve First? IBMBigDataHub.com, January 16. [Link](#)
- Kibler, W. Ben.** 2009. The 4000-Watt Tennis Player: Power Development for Tennis. *Medicine and Science in Tennis* (Society for Tennis Medicine and Science) 14(1): 5–8. [Link](#)
- Klaassen, Franc J. G. M., and Jan R. Magnus.** 2009. The Efficiency of Top Agents: An Analysis Through Service Strategy in Tennis. *Journal of Econometrics* 148(1): 72–85.
- O’Donoghue, P., and A. Ballantyne.** 2004. The Impact of Speed of Service in Grand Slam Singles Tennis. In *Science and Racket Sports III*, ed. Adrian Lees, Jean-Francois Kahn,

and Ian W. Maynard, 212–218. Abingdon, UK: Routledge.

O'Donoghue, P., and D. Liddle. 1998. A Notational Analysis of Time Factors of Elite Men's and Ladies' Singles Tennis on Clay and Grass Surfaces. In *Science and Racket Sports II*, ed. Adrian Lees, Ian W. Maynard, Mike Hughes, and Thomas Reilly, 241–246. London: E & FN Spon.

Paserman, M. Daniele. 2010. Gender Differences in Performance in Competitive Environments? Evidence from Professional Tennis Players. *Boston University Department of Economics Working Papers* WP2010-047. Boston University (Boston, Mass.). [Link](#)

Pope, Devin G., and Maurice E. Schweitzer. 2011. Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. *American Economic Review* 101(1): 129–157.

Rioul, François, Sami Mecheri, Bruno Mantel, François Kauffmann, and Nicolas Benguigui. 2015. What Can Hawk-Eye Data Reveal about Serve Performance in Tennis? In *Proceedings of the 2nd Workshop on Machine Learning and Data Mining for Sports Analytics Co-Located with 2015 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, ed. Jesse Davis, Jan Van Haaren, and Albrecht Zimmermann, 36–45. Aachen, Germany: CEUR Workshop Proceedings. [Link](#)

Rodenberg, Ryan. 2014. What Every Pro Tennis Player Does Better Than Roger Federer. *The Atlantic*, January 13. [Link](#)

Schönborn, Richard. 2000. *Advanced Techniques for Competitive Tennis*, trans. Heather Ross. Oxford: Meyer & Meyer Sport.

About the Author



Michał Krawczyk received his Ph.D. in economics from the University of Amsterdam. He is currently associate professor at the University of Warsaw. His research interests encompass several subfields of behavioral economics. Most of his studies involve laboratory and field experiments. He has published i.a. in *American Economic Review*, *Behavioral and Brain Sciences*, *Experimental Economics*, *Journal of Economic Behavior and Organization*, *Journal of Economic Psychology*, *Journal of the European Economic Association*, *Journal of Public Economics*, *MIS Quarterly*, and *PLoS ONE*. His email address is mkrawczyk@wne.uw.edu.pl.

[Anbarci, Arin, and Zenker's reply to this article](#)

[Go to archive of Comments section](#)

[Go to March 2019 issue](#)



Discuss this article at Journaltalk:
<https://journaltalk.net/articles/5984/>