Mixed Strategy Equilibrium: Analysis of Highly Ranked Professional Tennis Players

by Brady Anderson

Submitted in Partial Fulfillment for Graduating with Distinction

Economics Department

Coe College
Cedar Rapids, IA
April 2018

Contents

Abstract

This paper analyzes serve and return strategy for men's professional tennis in the context of a mixed strategy equilibrium. In tennis where there is no pure strategy equilibrium, economic theory suggests a mixed strategy equilibrium where players randomly mix strategies such that the expected payoff of playing each strategy is equalized. Using data from over 500 matches from 10 of the world's top tennis players, I test whether the expected points won serving to the forehand equal the expected points won serving to the backhand. Likewise, I test whether the mixed strategy equilibrium holds in return strategy such that the expected points won when returning to the forehand equal the expected points when returning to the backhand. I find that the players are largely optimizing their strategies while serving, but are not as optimal when returning. This discrepancy can be explained by stronger adherence to the theory when the player is given more time via the 2^{nd} serve return to carry out strategies. Furthermore, when testing for serial correlation, players were largely found to switch too often between strategies to be consistent with randomness. Despite this, the results found offer strong evidence supporting the economic theory of a mixed strategy equilibrium when applied to professional men's tennis.

Intro

Nearly 70% of professional men's tennis points are between zero and four shots (O'Shannessy, 2017). The vast majority of points are short, indicating these first few shots are crucial to success at the top level. Furthermore, in the past 20 years, the number one player in the world wins just 55% of his total points played during that season (O'Shannessy, 2017). This small point margin results in the best player in the world typically winning 90% or more of their matches in any given season. Consequently, it is clear the marginal benefit of an additional point in a professional men's tennis match is extremely high: very small discrepancies determine the very best players.

Short points and narrow margins signify that modeling of both the serve and return placement are the most crucial shots to analyze when testing for optimization among players. A player's intelligent approach over these high leverage shots is the most fundamental factor in developing optimal strategies for a tennis match. Therefore, professional players should be making the most optimal decisions for these first two shots under the framework of game theory.

Game theory is used to model the strategic interaction between players of a given game. Tennis is an ideal sport to model using game theory as it is focused on the interaction between two individual players, simplifying analysis. However, traditional game theory suggests a pure strategy that each player should always engage in. For a basic example, say a tennis player has two options: hit to the forehand or hit to the backhand of the opponent. A pure strategy would be one in which the player maximizes his winning rates by always choosing one of the two options. However, in tennis and other complex games players do not have a pure strategy. If a server in tennis always hits to one side over the other, the opponent would exploit the systematic behavior of the server to their advantage. Thus, tennis players must randomly alternate between their strategies in order to optimize their chances of winning. This is called a mixed strategy, which focuses on formally modeling the probability-based decision-making process in the interaction between two or more players. Players must randomly mix across available strategies in a manner that makes their behavior unpredictable to their opponent.

In tennis, players can mix their shot placement to the forehand or the backhand when serving and returning. Mixed strategy game theory suggests professional players will do so in a manner which maximizes their chances of winning a match. A player reaches a mixed strategy equilibrium when he produces an equivalence in the effective outcomes (winning rates) across all strategies in such a way as to have no incentive to change behavior. If the theory fails to hold up, this suggests the player is winning a much greater share of points when engaging in a certain strategy, and should shift strategies accordingly until the optimal equilibrium is reached. For instance, a player who is winning 90% of their serving points directed to the forehand but only 10% directed to the backhand should, given an equal allocation of points played between the two, adjust by serving more often to the forehand wing.

Because of the ability and knowledge of both their own and their opponents' games, a significant discrepancy in winning rates across strategies should be incongruent with expectations at the professional level. Previous literature in card games have found that novice players tend to behave less optimally according to a mixed strategy equilibrium theory than do professional players (O'Neill, 1987). Similarly, the very best tennis players are expected to be cognizant and optimal in adhering to mixed strategy equilibrium, because if they are not, it suggests the theory is not robust when applied to professional tennis.

The present paper focuses on testing for evidence of the economic theory of a mixed strategy equilibrium in tennis, and is largely predicated on the previous work by Walker and Wooders (2001). Walker and Wooders analyzed 10 professional men's tennis matches and measured the expected payoff of two different serving locations on first serves for each player in a match-by-match basis. They found, overwhelmingly, that players tended to adhere to the predictions of the mixed strategy equilibrium theory: the expected points won across strategies were equalized. This paper updates and extends this study by analyzing the strength of the mixed strategy equilibrium theory when applied to both serving and returning performances of 10 of the top professional men's tennis players over several matches. When serving, the best professional players were largely found to behave according to theory: players were equally proficient in winning rates when serving to either the forehand or backhand on both first and second serves. When returning, players were found to be optimal on second serve returns but not optimal when returning first serves. Additionally, there was strong evidence of serial correlation in the decision-making process when serving, as players were found to be over-mixing their strategies in a way inconsistent with randomness.

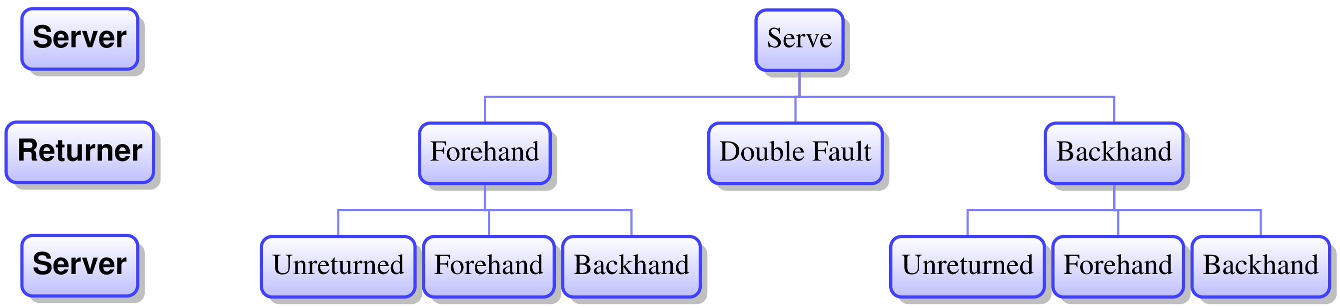
Game theory, specifically the branch involving mixed strategy equilibria, has been analyzed in several different domains. For instance, mixed strategy equilibrium theory was applied in the investment domain to determine the strategies of large firms and smaller competitors in the context of securing patent rights (Amaldoss & Jain, 2002). Evidence of oligopoly pricing in oil markets further supports the theory of a mixed strategy equilibrium (Wang, 2009). Additionally, Aragones and Palfrey (2002) developed a model based on the mixing of policy adoptions between two candidates running for a political office (Aragones & Palfrey, 2002). In the sporting world, pitching and rushing/passing outcomes in baseball and football respectively showed no evidence of mixed strategy equilibrium theory (Kovash & Levitt, 2009). Being an individual sport, it is natural to determine if tennis is an appropriate strategic environment in which player's closely mimic the results suggested by mixed strategy equilibrium theory.

The structure of the paper is as follows. First, Section II explains the underlying model that guides the research. Section III discusses the data sources used and explains some of the limitations and simplifying assumptions made. Following in Section IV is the presentation and discussion of relevant results for both serving and returning outcomes on a player-by-player basis. Section V assesses the level of randomness in each player's decision-making process on serve by checking for serial independence in serving locations. Lastly, Section VI highlights the key findings of the research as a whole and the implications it provides for further study into mixed strategy equilibrium theory.

Model

This paper will use a similar model for both the serving and returning performances of each player. Employing a method similar to Walkers and Wooders, the court is split into the deuce and ad court. I extend the analysis to examine both the first and the second serves in each service box, totaling four different 'point games' for each player while serving and returning. point games are stratified by serve type and court side as the key indicators. The assumption is of no further variability in the point games besides the type of serve and the court to which the serve is being placed. Figure 1 illustrates the overall structure of each sequential point game. The server hits the serve to the forehand or backhand of the returner (or fails to hit the serve in to one of these sides). Likewise, the returner then hits the return back to either the server's forehand or backhand, or misses the return to allow just two possible optimal strategies for each player.

Figure 1: Tennis Game Structure



For example, consider the hypothetical one-shot zero-sum tennis game represented in Table 1 (Dixit & Skeath, 2015). Dixit and Skeath illustrate a mixed strategy game with various payoffs in 100 points for both the server and returner in a certain point game. The Server hits the serve to the forehand or backhand with payoffs represented by the first number in each cell. Likewise the returner leans to the forehand or backhand wing with payoffs represented by the second number in each cell. Theory would expect the server to have a higher payoff when hitting a serve to the opposite side the returner is leaning to, reflected in the table.

Table 1: Theoretical Tennis Point Payoffs

		Returner Leaning To	
		Forehand	Backhand
Server Serving To	Forehand	50,50	80,20
	Backhand	90,10	20,80

In order for the server to be behaving optimally, the payoffs across the mixture of strategies needs to be equivalent. Let's say that the Server decides to carry out a mixture with $p = 0.75$ to the forehand and $(1 - p) = 0.25$ to the backhand, where p represents the proportion of shots hit by the Server to the forehand side. The Server payoffs for this mixture are thus the weighted payoffs based on the Returner's strategy.

$$Forehand = 0.75 \cdot 50 + 0.25 \cdot 90 = \mathbf{60} \tag{1}$$

$$Backhand = 0.75 \cdot 80 + 0.25 \cdot 20 = 65 \tag{2}$$

Equation (1) represents the Server's payoff if the Returner always leans forehand, and (2) reflects the Server's payoff if the Returner always leans backhand. It is clear that the Returner, as a rational optimizer, will always lean to the forehand side in this example as this minimizes the Server's payoff. As a Server, the optimal way to prevent exploitation from the Returner is to carry out a strategy mixture in which the Returner payoffs are equivalent. Setting the returner's weighted payoffs equal:

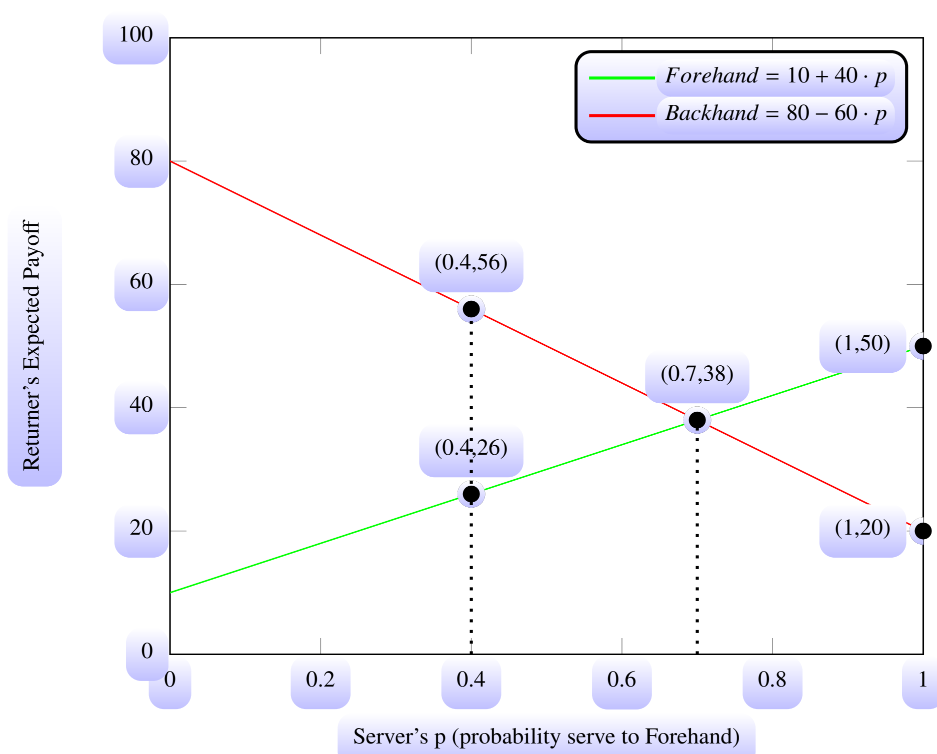
$$Payoff(Forehand) = Payoff(Backhand) \tag{3}$$

$$50p + 10(1-p) = 20p + 80(1-p) \tag{4}$$

$$10 + 40p = 80 - 60p \tag{5}$$

$$\mathbf{p=0.7} \tag{6}$$

Figure 2: Payoffs As Function of Server's p-mix Example



The optimal mixture for the Server in our hypothetical example is hitting 70% of serves to the forehand and 30% to the backhand. Figure 2 illustrates this concept. As the Server hits a greater proportion of serves to the forehand, the expected payoff for the Returner increases when they lean to the forehand, and decreases when they lean to the backhand. The Server will optimally carry out a mixture in which the Returner's expected payoff is minimized. This occurs when the payoffs of the Returner's strategies are equalized, otherwise there are incentives for exploitation by the Returner to increase their expected payoff, noted by the higher of the two strategies in non-optimal p example values shown in Figure 2. Therefore, in each of the point games for both the server and returner, testing for an equilibration in the winning rates between the two strategies yields the level of conformity between actual player performances and those predicted under the mixed strategy equilibrium theory.

In testing actual player data, I solve for evidence of a mixed strategy equilibrium by testing for an equality of winning probabilities between serving to the returners forehand and backhand for each player in each of the four point games. Following is the same procedure but for the returner. Likewise, I test for the equality in winning probabilities when the returner hit the return to the forehand and backhand of the server in each of the four point games.

In order to simplify the model, each point is assumed to be drawn from the same distribution regardless of the score and outcomes of previous points. Previous literature on the subject found that tennis points are not independently and identically distributed (i.i.d.) across a match, however the deviance was small enough that relaxing the assumption still provides a good approximation of professional point play (Klaassen & Magnus, 2001). Additionally, previous literature on tennis matches has found that the only Nash equilibria in point games is for a player to engage in the optimal mixed strategy equilibrium mix on each point (Walker, Wooders, & Amir, 2011). As a result, the model I present tests a binomial process that assumes tennis points are i.i.d. across all serves and returns in each match.

The assumption is made that the skill level of all opponents are equal for the returner, allowing for aggregation of the totality of points played over several matches. This assumption is admittedly incomplete in faithfully testing the theory, as many players tend to favor different shots and strategies over another. Accounting for the heterogeneity of players would provide additional insights to each specific player, and has been tested in a similar domain in penalty kicks in soccer (Chiappori, Levitt, & Groseclose, 2002). However, to avoid unnecessary complexity and to gain a baseline understanding of the validity of the theory in general, these concerns are relaxed. Any error introduced through this method is assumed to be just as likely to aggregate a false significant discrepancy between winning probabilities of each strategy as it is to indicate a false insignificant relationship by canceling out individual matches in which there may have been significant discrepancies in winning rates. Furthermore, because the data primarily examines the top players, the skill sets of each player are similar enough to mitigate this concern.

In testing for mixed strategy equilibria, Pearson's chi-squared goodness-of-fit test for equality of two distributions is used. This method tests whether or not the winning probability parameters between serving/returning to the forehand and backhand are equalized. Because it is testing a two by two matrix, it has 1 d.f. since the data is estimated and the true proportions of winning probabilities are unknown. Formally, the hypothesis investigated is represented as:

$$H_0 : P^i_F = P^i_B$$

$$H_1 : P^i_F \neq P^i_B$$

$$\text{for each player, i, } i \in \{1, 10\}$$

where the null, H_0, is the probability of winning a point when the shot (either a serve or return) is hit the forehand, P_F, is equal to the probability of winning a point when the shot (either a serve or return) is hit to the backhand, P_B. If the distributions of outcomes between the two serving or returning strategies are significantly different, this signals a lack of optimization in accordance to the theory of mixed strategy equilibrium. The player in this instance would not be behaving rationally, and should adjust his strategy accordingly.

For example, consider the point-game of the first sere hit in the deuce court. Roger Federer's mix of serve direction between forehand and backhand is depicted in Table 2 along with his win and loss frequency across each strategy. According to the mixed strategy equilibrium theory, Federer is expected to win an equal percentage of points won when serving

Table 2: Roger Federer Service Points

Deuce 1^{st} *Serve*	Points Won	Points Lost	Total
Forehand	1174	376	1550 (51.20% of 3027)
Backhand	1192	285	1477 (48.80% of 3027)
Total	2366	661	3027

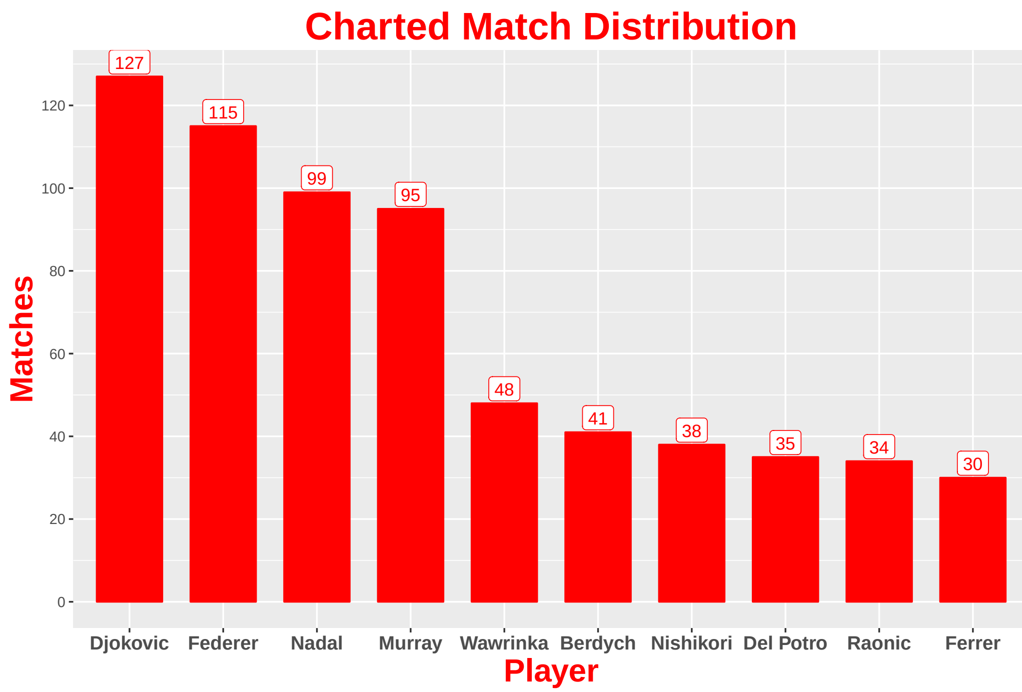
to the forehand as he is to the backhand side. The expected frequencies are compared to the observed frequencies, and the Pearson chi-squared test is used to compare the likelihood of obtaining the observed results if the distributions were truly equally proficient in winning rates. In Federer's case, the expected amount of points won and lost to both forehand and backhand are calculated by taking the marginal distribution of each strategy times the total points won and lost. This yields expected results of: $0.5120 \cdot 2366 \approx 1212$ points won to the forehand, $0.5120 \cdot 661 \approx 338$ points lost to the forehand, $0.4880 \cdot 2366 \approx 1154$ points won to the backhand, and $0.4880 \cdot 661 \approx 323$ points lost to the backhand. The difference between the expected and the observed points won is 38 too few of points won to the forehand side and 38 too few won to the backhand side. This result is just one of two of the 40 service point-games that yielded a statistically significant deviation from the mixed strategy equilibrium theory.

Data

Data was acquired through the public crowdsourced repository, *The Match Charting Project* (Sackmann, 2016). The open-source raw data set contains all necessary details regarding the placement and shot type of both the serve and return across thousands of points for each player. Figure 3 illustrates the distribution of matches per player. All available data was used for players against right-handed opponents in the five year period from 2011-2016.

Due to the nature of the data, the players selected were biased in favor of the top players who had the most charted matches. On the one hand, this was done to provide a plethora of data to perform more robust statistical testing. However, this also limits the scope of the conclusions that can be drawn due to lack of variability in the skill level of the professional players. Additionally, because of the bias towards the very best players in the data, many of

Figure 3: 2011-2016 Match Distribution by Player



the matches are against one-another as well. Most of the matches, therefore, are of later rounds in the tournaments when the higher-ranked players play one another. Including the testing of lower ranked players may shed further light on whether or not the theory holds more generally across varying levels of skill on the men's professional tour.

Limitations to the data include the data voids and other various minor discrepancies. There were some points during a given match that were not accounted for because the charter either missed the point or the video of the point itself was unavailable. This flaw is negligible however, as the majority of the points were captured and it can be assumed there is no pattern in the specific juncture in each match in which these omissions occur. Additionally, the data collected was only of servers against right-handed returners. Nonetheless, this should have little effect on the overall results as the majority of professional players are right-handed. However, accounting for all players is a topic of further research which may further validate the findings presented.

Results

Serving

Table 3 lays out the framework and results for each player when serving. For example, the first row under Federer is meant to be read as such: Federer hits the serve on the Deuce court, the serve is struck to the forehand side of the returner, the first and second serve winning rates are given and the respective p-value follows for each when comparing the distribution against hitting the same serve and court type but to the backhand side of the returner. The p-value measures the likelihood of the winning probability distributions being equal between the two strategies of hitting to the forehand or the backhand in each point game. A critical measure of $p \leq 0.05$ is used to denote significant deviance from the mixed strategy equilibrium theory.

Table 3: Testing Equality In Win Probabilities Of Serving Strategies

Player	Court	Direction	1st Serve Win Rate	p-value	2nd Serve Win Rate	p-value
Federer	Deuce	Forehand	0.757	0.018*	0.544	0.523
		Backhand	0.807		0.564	
	Ad	Forehand	0.764	0.286	0.569	0.670
		Backhand	0.787		0.583	
Djokovic	Deuce	Forehand	0.710	0.526	0.549	0.166
		Backhand	0.723		0.593	
	Ad	Forehand	0.689	0.259	0.556	0.710
		Backhand	0.714		0.568	
Nadal	Deuce	Forehand	0.703	0.950	0.555	0.948
		Backhand	0.702		0.557	
	Ad	Forehand	0.717	0.870	0.541	0.365
		Backhand	0.720		0.588	
Murray	Deuce	Forehand	0.707	0.006*	0.538	0.280
		Backhand	0.778		0.502	
	Ad	Forehand	0.698	0.080	0.536	0.248
		Backhand	0.751		0.497	
Wawrinka	Deuce	Forehand	0.739	0.901	0.578	0.510
		Backhand	0.743		0.550	
	Ad	Forehand	0.688	0.397	0.457	0.062
		Backhand	0.719		0.545	
Nishikori	Deuce	Forehand	0.725	0.521	0.527	0.701
		Backhand	0.697		0.507	
	Ad	Forehand	0.677	0.879	0.503	0.646
		Backhand	0.683		0.527	
Del Potro	Deuce	Forehand	0.694	0.069	0.534	0.668
		Backhand	0.762		0.509	
	Ad	Forehand	0.737	0.718	0.634	0.267
		Backhand	0.722		0.563	
Ferrer	Deuce	Forehand	0.674	0.521	0.579	0.206
		Backhand	0.642		0.518	
	Ad	Forehand	0.633	0.864	0.548	0.595
		Backhand	0.642		0.510	
Raonic	Deuce	Forehand	0.807	0.901	0.548	0.569
		Backhand	0.802		0.515	
	Ad	Forehand	0.752	0.274	0.533	0.770
		Backhand	0.796		0.551	
Berdych	Deuce	Forehand	0.776	0.246	0.535	0.418
		Backhand	0.727		0.491	
	Ad	Forehand	0.795	0.305	0.426	0.105
		Backhand	0.745		0.521	

The top professional players are behaving as the economic theory predicts, as there are only two out of the 40 instances in which there is a significant deviance from equalized winning probabilities. These both occur on the first serve side and are highlighted accordingly, as Federer and Murray are both not optimizing their first serve strategies when serving to the Deuce court. Federer and Murray are both winning a significantly greater percentage of points when serving to the backhand wing of the returner in these point games, suggesting each should adjust accordingly and direct more of their serves toward the backhand against right-handed opponents. Conversely, a returner knowing this information should try to exploit this discrepancy and adjust his strategy by leaning to their backhand wing to defend against the more successful strategy when playing either Federer or Murray.

Returning

Looking at returning strategies, much less optimal performances are seen. Table 4 depicts each players mixed strategy when returning both first and second serves with a forehand or backhand in each court to the server's forehand or backhand. Because of the lack of data, only points in which the returner made the return are included in the analysis, which limits the model slightly. However, theory still expects players to be equally viable when making the return to both sides.

In six of the 40 point game experiments the null hypothesis, H_0 is rejected, five of which occur when returning the first serve. The top players are much less optimal when returning the first serve than when returning the second serve, as each player wins a significantly greater percentage of points when they return the first serve back to their opponent's backhand wing. This is, in fact, expected in men's professional tennis, and the strong evidence of equalized winning probabilities when returning second serves further strengthen the mixed strategy equilibrium theory.

Professional men's tennis is dominated by the forehand wing, as nearly 70% of all groundstroke winners come from the forehand side (O'Shannessy, 2017). Generally, players need to avoid hitting to the forehand unless it is a strong and direct shot that does not allow their opponent to step in and dictate off of that wing. First serves are much faster than second serves and generally to much more precise targets, resulting in returners having much less time and

Table 4: Testing Equality In Winning Probabilities Of Return Strategies

Player	Serve	Return	1st Serve Win Rate	p-value	2nd Serve Win Rate	p-value
Federer	Forehand	Forehand	0.472	0.005*	0.566	0.456
		Backhand	0.589		0.606	
	Backhand	Forehand	0.429	0.069	0.491	0.056
		Backhand	0.497		0.548	
Djokovic	Forehand	Forehand	0.504	0.723	0.573	0.513
		Backhand	0.516		0.602	
	Backhand	Forehand	0.455	0.001*	0.567	0.366
		Backhand	0.542		0.593	
Nadal	Forehand	Forehand	0.476	0.205	0.568	0.516
		Backhand	0.553		0.598	
	Backhand	Forehand	0.451	0.861	0.555	0.294
		Backhand	0.462		0.617	
Murray	Forehand	Forehand	0.457	0.550	0.474	0.128
		Backhand	0.483		0.551	
	Backhand	Forehand	0.382	0.007*	0.503	0.035*
		Backhand	0.497		0.574	
Wawrinka	Forehand	Forehand	0.429	0.276	0.533	0.893
		Backhand	0.518		0.514	
	Backhand	Forehand	0.358	0.072	0.502	0.879
		Backhand	0.467		0.509	
Nishikori	Forehand	Forehand	0.496	0.643	0.532	0.239
		Backhand	0.457		0.626	
	Backhand	Forehand	0.339	0.004*	0.551	0.602
		Backhand	0.532		0.581	
Del Potro	Forehand	Forehand	0.461	0.617	0.511	0.594
		Backhand	0.503		0.446	
	Backhand	Forehand	0.374	0.353	0.482	0.793
		Backhand	0.449		0.495	
Ferrer	Forehand	Forehand	0.423	0.297	0.600	1.000
		Backhand	0.510		0.600	
	Backhand	Forehand	0.320	0.334	0.503	0.910
		Backhand	0.410		0.510	
Raonic	Forehand	Forehand	0.435	0.920	0.524	0.859
		Backhand	0.445		0.505	
	Backhand	Forehand	0.385	0.978	0.422	0.135
		Backhand	0.382		0.522	
Berdych	Forehand	Forehand	0.431	0.989	0.532	0.483
		Backhand	0.432		0.623	
	Backhand	Forehand	0.329	0.050*	0.538	0.783
		Backhand	0.478		0.555	

ability to strategize and direct their return to an ideal location: deep and more often towards the backhand wing. Servers know this, and often are looking to run around to get a forehand and take control of the point from the first ball after the return, resulting in them hitting a forehand at any possible opportunity if the return is a weak shot. Returners are unable to hit the return as precisely and powerfully on first serves, resulting in many of them being hit as forehands by the server even if they were directed towards the backhand side of the servers court. Thus, if a returner is able to get it to the backhand, it most likely means the return was of high enough quality the server did not have time to get a forehand out of it.

The second serve return, however, does allow for more time to carry out an optimal mix of strategies. The serves are much slower, and generally more towards a safer target in the box. Consequently, the returner can then hit a more aggressive and precise shot towards either wing more often, and it shows in the equalized distribution of winning probabilities. Players are much more optimal when they have a greater ability to carry out their ideal mix of returns, as in just one of the 20 second serve point games, or 5% as expected, are at the significant p-value threshold or lower. Thus, the reversion back to the predictions of mixed strategy theory further strengthen the model's representational ability in the context of this professional men's tennis.

Serial Independence

In order for each agent to be optimizing their serving and returning strategies they not only must win an equal share of points but also must be randomly mixing across choices. Randomness prevents systematic exploitation by the opponent, defeating the benefits of uncertainty. Serial independence, or the concept that a player can not be "too predictable" in how they decide to switch up their decision-making process, is also necessary in order to conclude players are optimally strategizing when serving and returning.

Therefore, it was individually tested for each of the 10 players in the data set the hypothesis whether the player's choices were serially independent, when serving across both first and second serves. For the server, let $fs^i = (fs^i_1, fs^i_2, ...fs^i_{n_F+n_B})$ be the sequence of first serve placements of player i, in chronological order, where $fs^i \in \{F, B\}$ up to the total number of points $n_F + n_B$. For example, if player '1' (say Roger Federer) hit the first three first serves to the backhand and the fourth to the forehand, the sequence would be symbolized as $fs^1 = (B, B, B, F)$ with $n_F + n_B = 4$. Likewise, let $ss^i = (ss^i_1, ss^i_2, ...ss^i_{n_F+n_B})$ be the sequence of second serve placements of player i, in chronological order, where $ss^i \in \{F, B\}$. For returners, focus is solely on second serve returns due to the greater ability to enact an optimal strategy as discussed previously. Let $sr^i = (sr^i_1, sr^i_2, ...sr^i_{n_F+n_B})$ be the sequence of second serve return placements of player i, in chronological order, where $sr^i \in \{F, B\}$.

The test for serial independence of each player's choices is based on the amount of runs, r_i, each sequence contains. A run is defined as a maximal string of consecutive identical symbols, either all F's or all B's. In the above example, $fs^1 = (B, B, B, F)$ represents two 'runs'. Using a similar procedure as that of Walker and Wooders (2001), the null hypothesis of serial independence is rejected when there are either "too many" runs or "too few" runs. Too many runs suggest a negative correlation in the player's choice of shot direction, or the player is changing direction too often in a predictable manner inconsistent with randomness. Too few runs suggest a positive correlation in a player's choice of shot direction, or there is a greater likelihood of the player choosing the previous shot direction than would be if the player were mixing randomly.

The Wald-Wolfowitz runs test is employed to determine the level of randomness of each player's serving and returning strategies (Wald & Wolfowitz, 1940). The test for randomness follows the following two-sided procedure:

H_0: The symbols F and B are randomly mixed in the sequence

H_1: The symbols F and B are not randomly mixed in the sequence

where 'F' represents a serve/return directed to the opponent's forehand and 'B' represents a serve directed to the opponent's backhand, for each point game of a given player. The null hypothesis assumes the player independently chooses one of the two strategies from the same distribution in each element in a sequence. The critical value of note will be a p-value = 0.05, meaning to be deemed non-random in serve placement the player must be within 0.025 of the left tail of the distribution (not enough runs to be random) or 0.025 too many runs on the right tail of the normal distribution (too many runs to be random).

The amount of runs were standardized into Z-scores for a normal distribution using the following relationship: $Z = \frac{R-\bar{R}}{s_R}$, where R is the number of observed runs, $\bar{R}$ is the number of expected runs and s_R is the standard deviation. The expected number of runs, $\bar{R}$, and the variance, s_R^2, are thus based on the following equations:

$$\bar{R} = \frac{2n_1n_2}{n_1 + n_2} + 1 \tag{7}$$

$$s_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} \tag{8}$$

Where n_1 and n_2 represent the total number of serves hit to the forehand and backhand for each player. Results are aggregated over the total number of matches played by summing the total number of runs in each individual match for each player.

Table 5 illustrates the results of each player when serving in the four various point games while serving. Values with asterisks indicate a significantly non-random result, either due to too many runs or too few runs. Too many runs is designated as a (−) in the Run Correlation column while too few is given a (+). Statistical randomness suggests there should be 5% or just two out of the 40 point games that do not adhere to the null hypothesis of randomly mixing serves to the forehand and backhand. The results found that 18 of the 40 were not serially independent, as there were either too many or two few of runs in nearly half of the point games to be consistent with the randomness hypothesis. Therefore, I reject the null hypothesis that players are randomly mixing their serving locations to the forehand and backhand. The top players appear to be mixing up their strategies too often, evidenced by the significant positive run correlation seen in 15 out of the 40 point games. This is consistent with experimental evidence regarding the behavior of humans in strategic situations when trying to mix randomly (Wagenaar, 1972).

There are a number of potential shortfalls in this analysis, one of which is the potential overrepresentation of runs on second serves. Because the second serve is the weaker of the two shots, the serve tends to be more predictable by nature, and a receiving opponent will thus try to 'run-around' and create a forehand knowing the location the serve will be hit. In this case, the results from second serves may be overstating the positive run correlation and too much switching found in these results. Nevertheless, players still showed a strong tendency to be too predictable in over-mixing their serving strategy on first serves in six of the 20 point games.

Another limitation is the aggregation of all matches in determining the total amount of runs. This assumes homogeneity in the ability of the returner on the backhand and forehand wing, and simplifies the conditions of court type and tournament round to be irrelevant factors to a given strategy. While taken to even out in the representation of runs, these factors may need to be examined to uncover more information on player mixing strategies. Careful examination of how a player's winning percentages and individual opponents on a given day affect a serving strategy of each player, if at all, would help clarify this point. However, this baseline method provides evidence to comfortably conclude that players are not behaving in line with

Table 5: Runs Test Player Serving Data

Player	Serve	Court	Forehand	Backhand	Runs, r_i	Z-score	p-value	Run Correlation
Federer	1st	Deuce	1550	1477	1407	-3.91	0.000*	(+)
Federer	1st	Ad	1175	1453	1252	-1.91	0.028	
Federer	2nd	Deuce	722	959	886	3.05	0.001*	(–)
Federer	2nd	Ad	642	976	827	2.67	0.004*	(–)
Djokovic	1st	Deuce	1693	1707	1432	-9.23	0.000*	(+)
Djokovic	1st	Ad	1433	1625	1436	-3.19	0.001*	(+)
Djokovic	2nd	Deuce	867	981	975	2.50	0.006*	(–)
Djokovic	2nd	Ad	715	954	864	2.28	0.011*	(–)
Murray	1st	Deuce	1330	1131	1209	-0.587	0.279	
Murray	1st	Ad	1168	919	1044	0.638	0.261	
Murray	2nd	Deuce	580	937	738	1.115	0.132	
Murray	2nd	Ad	497	1007	653	-0.789	0.215	
Del Potro	1st	Deuce	304	609	420	1.002	0.158	
Del Potro	1st	Ad	415	431	444	1.387	0.083	
Del Potro	2nd	Deuce	141	365	208	0.397	0.346	
Del Potro	2nd	Ad	119	315	197	2.810	0.002*	(–)
Nadal	1st	Deuce	821	1792	1212	3.856	0.000*	(–)
Nadal	1st	Ad	869	1603	1138	0.440	0.330	
Nadal	2nd	Deuce	662	548	584	-0.965	0.167	
Nadal	2nd	Ad	379	596	497	2.201	0.014*	(–)
Nishikori	1st	Deuce	490	415	483	2.184	0.014*	(–)
Nishikori	1st	Ad	358	429	411	1.417	0.078	
Nishikori	2nd	Deuce	265	315	302	1.102	0.135	
Nishikori	2nd	Ad	224	349	301	2.383	0.009*	(–)
Raonic	1st	Deuce	452	450	465	0.866	0.193	
Raonic	1st	Ad	342	440	411	1.828	0.034	
Raonic	2nd	Deuce	217	308	262	0.575	0.283	
Raonic	2nd	Ad	139	342	220	2.371	0.009*	(–)
Ferrer	1st	Deuce	344	349	411	4.829	0.000*	(–)
Ferrer	1st	Ad	309	299	350	3.661	0.000*	(–)
Ferrer	2nd	Deuce	185	235	206	-0.201	0.420	
Ferrer	2nd	Ad	112	288	177	1.829	0.034	
Wawrinka	1st	Deuce	618	603	580	-1.799	0.036	
Wawrinka	1st	Ad	482	622	513	-1.905	0.028	
Wawrinka	2nd	Deuce	332	586	451	1.870	0.031	
Wawrinka	2nd	Ad	231	602	355	1.741	0.041	
Berdych	1st	Deuce	443	407	496	4.866	0.000*	(–)
Berdych	1st	Ad	389	300	405	5.060	0.000*	(–)
Berdych	2nd	Deuce	211	365	286	1.580	0.057	
Berdych	2nd	Ad	159	423	253	2.182	0.015*	(–)

the standards of randomness.

Table 6: Runs Test Player Returning Data–Second Serves

Player	Court	Forehand	Backhand	Runs, r_i	Z-score	P-value	Run Correlat
Federer	Deuce	662	803	760	1.756	0.040	
Federer	Ad	638	678	678	1.082	0.140	
Djokovic	Deuce	830	760	857	3.144	0.001*	(–)
Djokovic	Ad	641	920	785	1.487	0.069	
Murray	Deuce	455	643	589	3.427	0.000*	(–)
Murray	Ad	350	678	500	2.594	0.005*	(–)
Del Potro	Deuce	263	197	221	-0.502	0.308	
Del Potro	Ad	166	288	185	-2.695	0.004*	(+)
Nadal	Deuce	809	287	436	0.884	0.188	
Nadal	Ad	571	534	555	0.128	0.449	
Nishikori	Deuce	144	260	200	1.483	0.069	
Nishikori	Ad	156	214	196	1.553	0.060	
Raonic	Deuce	217	137	179	1.126	0.130	
Raonic	Ad	149	182	161	-0.429	0.334	
Ferrer	Deuce	85	188	118	-0.010	0.496	
Ferrer	Ad	108	213	140	-0.542	0.294	
Wawrinka	Deuce	335	249	293	0.536	0.296	
Wawrinka	Ad	226	278	255	0.422	0.337	
Berdych	Deuce	169	175	182	1.924	0.027	
Berdych	Ad	241	127	184	0.978	0.164	

Table 6 illustrates the same runs test procedure when examining the level of randomness of each player's second serve return placement. The first serve return is omitted as player's were not optimally mixing in this set of points as shown previously. On second serve returns there are four statistically significant deviations from random mixing, with three of the four mixing too often to be considered random. As with the return analysis regarding a mixed strategy equilibrium in winning percentages, the analysis here is limited to just returns hit in play to a specific wing. Court side (Deuce/Ad) was analyzed rather than return type (Forehand/Backhand) due to greater consistency in alternating between the two court sides during match play. There is limited data that may tell a different story if holistic, however expectations for strategic randomization should still hold. While there is deviance, randomization of return shot placement is largely the trend, more so than in the serving performances of the players. A possible explanation of the discrepancy includes the lesser control over shot direction for the returner, as the second serve return is still a reactive shot relative to the serve. This may result in additional randomness inadvertently that is not present when, as in the case with the serving outcomes, a player may prefer to mix too often in attempts of deceiving the opponent.

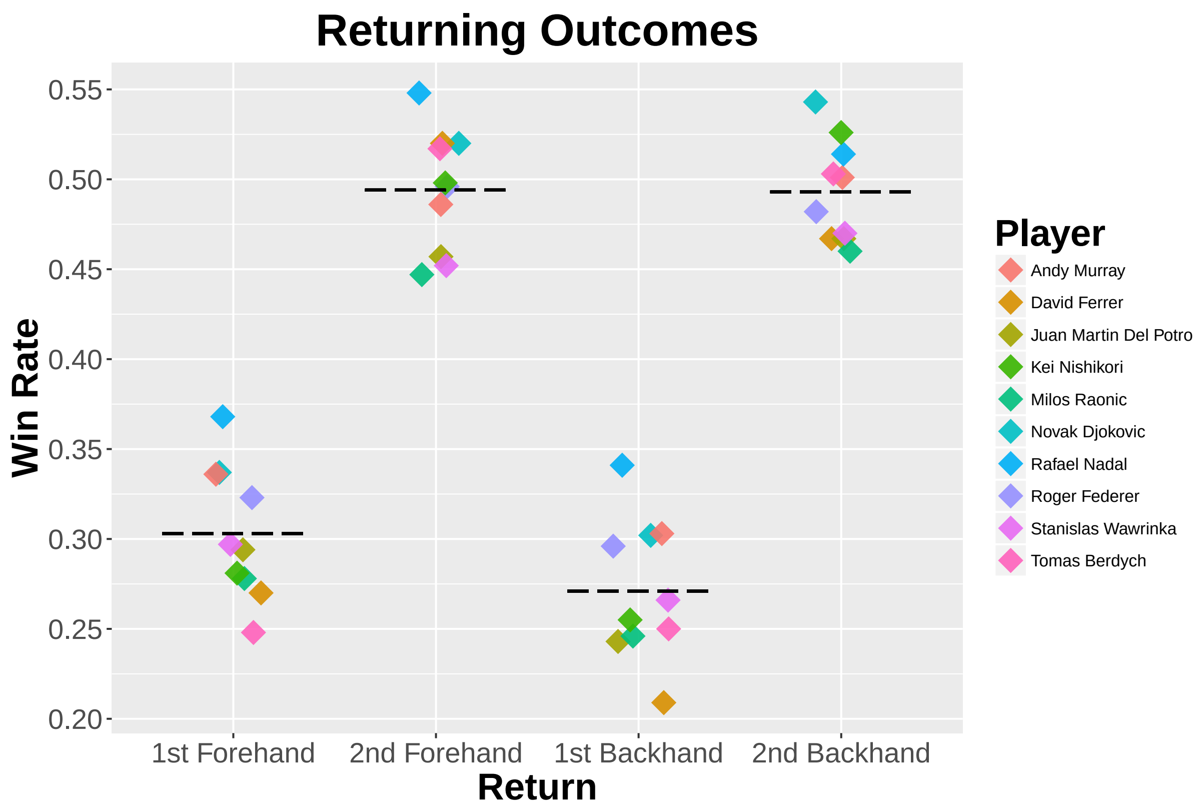
Concluding Remarks

Accounting for Heterogeneity

As discussed earlier, the assumption that each player is homogenous in skillset is invalid. Not all players have the same skillsets and affinities when serving or returning serves, as many prefer one shot over another. Figure 4 illuminates this discrepancy, as several players show different success rates when returning based on serve type and return shot. This graphic differs from the test of mixed strategy equilibrium on return placement, as these numbers account for all types of points, both returned and unreturned. For instance, David Ferrer wins around 52% of his points when returning the second serve with a forehand, above the group average (the black dashed line). Conversely, Ferrer is winning just 47% of his returns when hitting a backhand on the second serve, notably below the mean of the group. When playing against Ferrer, one is more successful if they can keep him from hitting a forehand return, whereas against someone like Kei Nishikori, an optimal player may serve more often to his forehand wing given his relative inefficiency in that department.

Accounting for the diverse skillsets between players requires a plethora of match data for each opponent of each player to have confidence in the results. Aside from a few match-ups between the top players, testing the theory under the heterogeneity framework would have further limited the data available for analysis. Thus further research, either through a match-by-match analysis or by an aggregation of matches between a specific match-up, would shed greater light on the suitability of the theory. If the results still stand firm to the mixed strategy

Figure 4: Player Win Rates by Return Type



equilibrium hypothesis, the economic theory would further cement its explanatory power in the game of professional men's tennis.

Conclusions and Further Study

Professional men's tennis players are largely optimizing their strategic decisions in accordance to the mixed strategy equilibrium theory. Upon examination of 40 point games for both serving and returning performances of the 10 of the top players, the results showed the equilibration between the winning probabilities across the two different strategies of hitting to the forehand or backhand wing. When serving, just two of the forty point games were significant at the 5% level, or exactly as statistically projected.

On the other hand, returners did not behave as in line with the theory. Six of the forty point games yielded significant results, suggesting the players were not optimizing their return strategies. However, upon closer inspection, this discrepancy is likely explained by the relative lack of time to carry out their strategies on the first serve return. Five of the significant deviations on return were observed on the first serve return, compared to just one on the second serve return. When given more time and an easier shot to return, players behavior showed much stronger conformance with the mixed strategy equilibrium theory.

This aligns with intuition, as the greater the skill a player possesses, the more likely they should be to behave optimally. Certainly the players do not crunch these numbers mid-match however, which shows just how powerful this theory is when applied to tennis. The top professional players are constantly adjusting strategies throughout a match and season across a wide set of opponents, formally known as "learning", in a way that equilibrates their winning probabilities in this simple model approach. Despite the nearly infinite shot and point structures, a simple model of the match exhibits strong evidence of a mixed strategy equilibrium.

However, to be perfectly adhering to the mixed strategy equilibrium theory, players must be randomizing their strategies. When assessing serving location runs, players appear to lack the randomness necessary that an optimal mixed strategy equilibrium theory requires. In 18 of the 40 serving point games players were significantly non-random in their serving location. The majority of these instances were those in which the player was "too random" in their behavior as they switched from serving to the forehand to serving to the backhand more often than is consistent with a randomly selected strategy. The results of this paper are in agreement with the findings of Walker and Wooders (2001) while serving. When returning second serves, players were surprisingly more random in their shot direction than when serving, suggesting that more time to prepare a designated strategy may influence the predictability of a player's strategic behavior.

Further unpacking the lack of randomness could be done by assessing the "learning" by a server by examining how winning percentages change throughout each individual match when serving to the respective forehand or backhand. Perhaps players are purposefully more random because they are overthinking and trying to be 'too random' or, as shown, are equally viable in winning points across both strategies regardless of predictability. These results signal opponents are either not picking up on the server's predictable mixing or server's could be winning an even greater percentage of points on both sides with more random mixing.

As previously mentioned, an avenue of additional research is analyzing more precise locations of shots by each player, in comparison to just the type of shot presented in this paper. Because this research focused on the type of shot each hit rather than the location each shot was hit, there may be minor error in calculating the winning probabilities. For instance, several of the top players tend to "run-around" their backhands and hit forehand shots on the second serve, even when the server directed the shot at the returner's backhand wing. These are assumed away for convenience and simplicity in this paper, but may shed more faithful light to the optimization of choices across strategies. Furthermore, no distinctions between what type of shot is hit back is made, although more robust data could provide insights to this interesting question.

Professional men's tennis is a complex game. Players have mere fractions of a second to process their opponents shot selection and accordingly respond with the best possible strategy based on a plethora of conditions. These include, but are not limited to: shot depth, spin rate, spin type, pace, and shot location. Modeling all of these would not only be unfeasible, but uninformative.

Simplifying the approach to focus solely on shot location to either the forehand and backhand wing of the serve and return, these findings further validate those presented by Walker and Wooders (2001). Professional men's tennis players behave according to the economic theory of mixed strategy equilibrium when serving and returning, optimizing their performance. This research demonstrates strong evidence of the ability to model strategic human interaction in the domain of professional men's tennis.

References

Amaldoss, W., & Jain, S. (2002). David vs. Goliath: An analysis of asymmetric mixed-strategy games and experimental evidence. *Management Science*, *48*(8), 972–991.

Aragones, E., & Palfrey, T. R. (2002). Mixed equilibrium in a Downsian model with a favored candidate. *Journal of Economic Theory*, *103*(1), 131–161.

Chiappori, P.-A., Levitt, S., & Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 1138–1151.

Dixit, A. K., & Skeath, S. (2015). *Games of strategy: Fourth international student edition*. WW Norton & Company.

Klaassen, F. J., & Magnus, J. R. (2001). Are points in tennis independent and identically distributed? evidence from a dynamic binary panel data model. *Journal of the American Statistical Association*, *96*(454), 500–509.

Kovash, K., & Levitt, S. D. (2009). *Professionals do not play minimax: evidence from major league baseball and the national football league* (Tech. Rep.). National Bureau of Economic Research.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences*, *84*(7), 2106–2109.

O'Shannessy, C. (2017). The First 4 Shots. Retrieved August 29th, from `https://www.braingametennis.com/`

Sackmann, J. (2016, January). Match Charting Project. Retrieved July 28th, from `http://www.tennisabstract.com/blog/2013/11/26/the-match-charting-project/` (From Heavy Topspin Site)

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, *77*(1), 65.

Wald, A., & Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, *11*(2), 147–162.

Walker, M., & Wooders, J. (2001). Minimax Play at Wimbledon. *The American Economic Review*, *91*(5), 1521–1538.

Walker, M., Wooders, J., & Amir, R. (2011). Equilibrium play in matches: Binary markov games. *Games and Economic Behavior*, *71*(2), 487–502.

Wang, Z. (2009). (Mixed) strategy in oligopoly pricing: evidence from gasoline price cycles before and under a timing regulation. *Journal of Political Economy*, *117*(6), 987–1030.