Performing Best when it Matters Most: Evidence from professional tennis

Julio González-Díaz 1  Olivier Gossner 2  Brian Rogers 3

1 Research Group in Economic Analysis
   Universidad de Vigo

2 London School of Economics
   University of London

3 Kellogg School of Management
   Northwestern University

October 16th, 2009
Performing Best when it Matters Most

J. González-Díaz, B. Rogers, and O. Gossner
This paper belongs to a growing literature aiming at incorporating insights from psychology into (behavioral) economics.
This paper belongs to a growing literature aiming at incorporating insights from psychology into (behavioral) economics.

In particular, part of this literature tries to understand how economic agents react to different forms of pressure.
This paper belongs to a growing literature aiming at incorporating insights from psychology into (behavioral) economics.

In particular, part of this literature tries to understand how economic agents react to different forms of pressure.

Note that a perfectly rational agent’s performance should be unaffected by changes in pressure or in the stakes (importance of the situation).
This paper belongs to a growing literature aiming at incorporating insights from psychology into (behavioral) economics.

In particular, part of this literature tries to understand how economic agents react to different forms of pressure.

Note that a perfectly rational agent’s performance should be unaffected by changes in pressure or in the stakes (importance of the situation).

This literature acknowledges the fact that these changes may affect an agent’s ability to play optimally.
Motivation

Questions we want to address
Motivation

Questions we want to address

Is there heterogeneity in the abilities of the agents to perform best when it matters most?
Questions we want to address

1. Is there heterogeneity in the abilities of the agents to perform best when it matters most?
2. Does this heterogeneity have a significant impact on the success of the agents?
Motivation

Questions we want to address

1. Is there heterogeneity in the abilities of the agents to perform best when it matters most?

2. Does this heterogeneity have a significant impact on the success of the agents?

Why?
Questions we want to address

1. Is there heterogeneity in the abilities of the agents to perform best when it matters most?

2. Does this heterogeneity have a significant impact on the success of the agents?

Why?
This heterogeneity should be taken into account when designing contracts and providing incentives.
Related Literature
Related Literature

Related Literature


**Detrimental performance of the agents in the presence of superstars**

(Golf/Tiger Woods)
Related Literature


**Detrimental performance of the agents in the presence of superstars**

(Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”
Related Literature

  
  **Detrimental performance of the agents in the presence of superstars**  
  (Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”
  
  **Women’s performance decreases under pressure and male’s performance remains unchanged**  
  (Tennis)
Related Literature


**Detrimental performance of the agents in the presence of superstars**
(Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** (Tennis)

Heterogeneity?
Related Literature


**Detrimental performance of the agents in the presence of superstars**
(Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** (Tennis)


**Heterogeneity?**
Related Literature


  *Detrimental performance of the agents in the presence of superstars*
  (Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

  *Women’s performance decreases under pressure and male’s performance remains unchanged* (Tennis)


  *The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure* (Soccer/Penalty kicks)
Related Literature


  Detrimental performance of the agents in the presence of superstars
  (Golf/Tiger Woods)

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

  Women’s performance decreases under pressure and male’s performance remains unchanged (Tennis)


  The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure (Soccer/Penalty kicks)

Related Literature

• J. Brown (2009) “Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars”

**Detrimental performance of the agents in the presence of superstars**
(Golf/Tiger Woods)

• D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** (Tennis)

Heterogeneity?

• T. J. Dohmen (2006) “Do professionals choke under pressure?”

**The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure** (Soccer/Penalty kicks)

• Apesteguia and Palacios-Huerta (2009) “Psychological Pressure in Competitive Environments”

**Better performance at high stake situations when you have a better prospect; choking under “negative pressure”** (Soccer/Penalty shoot-outs)
Motivation

Tennis and point importance

Data & Methodology

Results

Related Literature


**Detrimental performance of the agents in the presence of superstars**

*(Golf/Tiger Woods)*

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** *(Tennis)*


**The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure** *(Soccer/Penalty kicks)*


**Better performance at high stake situations when you have a better prospect; choking under “negative pressure”** *(Soccer/Penalty shoot-outs)*

- Ariely *et al* (2009) “Large Stakes and Big Mistakes”
Motivation
Tennis and point importance
Data & Methodology
Results

Related Literature


**Detrimental performance of the agents in the presence of superstars**
*(Golf/Tiger Woods)*

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** *(Tennis)*


**The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure** *(Soccer/Penalty kicks)*


**Better performance at high stake situations when you have a better prospect; choking under “negative pressure”** *(Soccer/Penalty shoot-outs)*

- Ariely *et al* (2009) “Large Stakes and Big Mistakes”

**Detrimental effect of exceptionally high incentives** *(field experiment/India)*
Related Literature


**Detrimental performance of the agents in the presence of superstars**
*(Golf/Tiger Woods)*

- D. Paserman (2008) “Gender Differences in Performance in Competitive Environments”

**Women’s performance decreases under pressure and male’s performance remains unchanged** *(Tennis)*


**The performance of soccer players unaffected by changes in the stakes but they choke under (favorable) social pressure** *(Soccer/Penalty kicks)*


**Better performance at high stake situations when you have a better prospect; choking under “negative pressure”** *(Soccer/Penalty shoot-outs)*

- Ariely *et al* (2009) “Large Stakes and Big Mistakes”

**Detrimental effect of exceptionally high incentives** *(field experiment/India)*
Summary of results

We look at the behavior of professional tennis players
Summary of results

We look at the behavior of professional tennis players

Findings:
Summary of results

We look at the behavior of professional tennis players

Findings:

1. There is heterogeneity in agent’s reactions to changes in the importance of the situation.
Summary of results

We look at the behavior of professional tennis players

Findings:

1. There is heterogeneity in agent’s reactions to changes in the importance of the situation. What changes is the importance of the points of a tennis match
Summary of results

We look at the behavior of professional tennis players

Findings:

1. There is heterogeneity in agent’s reactions to changes in the importance of the situation. **What changes is the importance of the points of a tennis match**

2. This heterogeneity has a significant impact on an agent’s career.
Summary of results

We look at the behavior of professional tennis players

Findings:

1. There is heterogeneity in agent’s reactions to changes in the importance of the situation. *What changes is the importance of the points of a tennis match*

2. This heterogeneity has a significant impact on an agent’s career. *What we measure is the impact of the ability to perform best when it matters the most on the ratings/rankings of elite tennis players*
Real-life scenarios
Real-life scenarios

Financial traders

- Trading decisions must be made quickly and repeatedly
- Some decisions will involve a steeper risk/reward tradeoff
- Similar for some corporate managers
Real-life scenarios

Financial traders
- Trading decisions must be made quickly and repeatedly
- Some decisions will involve a steeper risk/reward tradeoff
- Similar for some corporate managers

Political campaigning
- U.S. presidential candidates campaign for years
- Many decisions to make, performances to give, along the way
- Some (nationally televised debates) have far more impact than others; some states are hugely influential
- Choking in an important performance/debate may mean losing the election
Critical Ability

Two different types of skill
Two different types of skill

**Standard ability:** Some agents may be generally better and making the correct decision
Critical Ability

Two different types of skill

1. **Standard ability:** Some agents may be generally better and making the correct decision

2. **Critical ability:** Some agents may be better than others “under pressure” at crucial moments
Critical Ability

Two different types of skill

1. **Standard ability**: Some agents may be generally better and making the correct decision

2. **Critical ability**: Some agents may be better than others “under pressure” at crucial moments

For the *critical ability*
Critical Ability

Two different types of skill

1. **Standard ability**: Some agents may be generally better and making the correct decision

2. **Critical ability**: Some agents may be better than others “under pressure” at crucial moments

For the *critical ability*

- Is it an issue of resource allocation?
Critical Ability

Two different types of skill

1. **Standard ability:** Some agents may be generally better and making the correct decision

2. **Critical ability:** Some agents may be better than others “under pressure” at crucial moments

For the *critical ability*

- Is it an issue of resource allocation?
- Is it a matter of psychological ability to handle pressure?
Critical Ability

Two different types of skill

1. **Standard ability:** Some agents may be generally better and making the correct decision

2. **Critical ability:** Some agents may be better than others "under pressure" at crucial moments

For the *critical ability*

- Is it an issue of resource allocation?
- Is it a matter of psychological ability to handle pressure?
Critical Ability

Two different types of skill

1. **Standard ability:** Some agents may be generally better and making the correct decision

2. **Critical ability:** Some agents may be better than others “under pressure” at crucial moments

For the *critical ability*

- Is it an issue of resource allocation?
- Is it a matter of psychological ability to handle pressure?
How do we analyze the described questions?
How do we analyze the described questions?

Data from professional tennis matches
How do we analyze the described questions?

Data from professional tennis matches

- Elite, trained, highly motivated agents
How do we analyze the described questions?

Data from professional tennis matches

- Elite, trained, highly motivated agents
- In a tournament, each agent plays many consecutive points
How do we analyze the described questions?

Data from professional tennis matches

- Elite, trained, highly motivated agents
- In a tournament, each agent plays many consecutive points
  - Unambiguous data available on point outcomes
How do we analyze the described questions?

Data from professional tennis matches

- Elite, trained, highly motivated agents
- In a tournament, each agent plays many consecutive points
  - Unambiguous data available on point outcomes
  - Points differ substantially in terms of their significance
How do we analyze the described questions?

Data from professional tennis matches

- Elite, trained, highly motivated agents
- In a tournament, each agent plays many consecutive points
  - Unambiguous data available on point outcomes
  - Points differ substantially in terms of their significance

⇒ High-quality information the context and on players’ performance
Outline

1. Motivation
2. Tennis and point importance
3. Data & Methodology
4. Results
Structure of a tennis match

Scoring structure
Structure of a tennis match

Scoring structure

- Player’s objective is to win the match.
Structure of a tennis match

Scoring structure

- Player’s objective is to win the match.
- These tennis matches consists of best of five sets.
Scoring structure

- Player’s objective is to win the match.
- These tennis matches consists of best of five sets.
- A set is won by winning 6 games (win by two, tie-break).
  - 12-point tie break is first player to win 7 points, win by two.
Structure of a tennis match

Scoring structure

- Player’s objective is to win the match.
- These tennis matches consist of best of five sets.
- A set is won by winning 6 games (win by two, tie-break).
  - 12-point tie break is first player to win 7 points, win by two.
- A game is won by winning 4 points (win by two, no tie-break).
**Structure of a tennis match**

**Scoring structure**

- Player’s objective is to win the match.
- These tennis matches consists of best of five sets.
- A set is won by winning 6 games (win by two, tie-break).
  - 12-point tie break is first player to win 7 points, win by two.
- A game is won by winning 4 points (win by two, no tie-break).
- Players alternate service games, with the first server chosen randomly.
Structure of a tennis match

Scoring structure

- Player’s objective is to win the match.
- These tennis matches consist of best of five sets.
- A set is won by winning 6 games (win by two, tie-break).
  - 12-point tie break is first player to win 7 points, win by two.
- A game is won by winning 4 points (win by two, no tie-break).
- Players alternate service games, with the first server chosen randomly.

Implication
Structure of a tennis match

Scoring structure

- Player’s objective is to win the match.
- These tennis matches consists of best of five sets.
- A set is won by winning 6 games (win by two, tie-break).
  - 12-point tie break is first player to win 7 points, win by two.
- A game is won by winning 4 points (win by two, no tie-break).
- Players alternate service games, with the first server chosen randomly.

Implication

Points are not all equally important
Defining the importance variable: PiM
Defining the importance variable: PiM

Differences in point importances are essential for our analysis
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won *i.i.d.* conditional on the server with probabilities $p_1, p_2$
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won \textit{i.i.d.} conditional on the server with probabilities $p_1, p_2$
Defining the importance variable: PiM

Differences in point importances are essential for our analysis.

Notion of importance

- The state of a match, \( \theta \), is given by the current score and server.
- Assume points are won i.i.d. conditional on the server with probabilities \( p_1, p_2 \).
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won i.i.d. conditional on the server with probabilities $p_1, p_2$
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won i.i.d. conditional on the server with probabilities $p_1, p_2$
- Thus one can compute $P(i \text{ wins match} \mid \theta)$
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won i.i.d. conditional on the server with probabilities $p_1, p_2$
- Thus one can compute $P(i \text{ wins match} \mid \theta)$

The **importance** of the point at state $\theta$ is:

$$P(i \text{ wins match} \mid \theta, i \text{ wins at } \theta) - P(i \text{ wins match} \mid \theta, i \text{ loses at } \theta)$$
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won i.i.d. conditional on the server with probabilities $p_1, p_2$
- Thus one can compute $P(i \text{ wins match} | \theta)$

The importance of the point at state $\theta$ is:

$$P(i \text{ wins match} | \theta, i \text{ wins at } \theta) - P(i \text{ wins match} | \theta, i \text{ loses at } \theta)$$

- Morris (1977), Klaassen and Magnus (2001), and Paserman (2008)
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won i.i.d. conditional on the server with probabilities $p_1, p_2$
- Thus one can compute $P(i \text{ wins match} | \theta)$

The importance of the point at state $\theta$ is:

$P(i \text{ wins match} | \theta, i \text{ wins at } \theta) - P(i \text{ wins match} | \theta, i \text{ loses at } \theta)$

- Morris (1977), Klaassen and Magnus (2001), and Paserman (2008)
Defining the importance variable: PiM

Differences in point importances are essential for our analysis

Notion of importance

- The state of a match, $\theta$, is given by the current score and server
- Assume points are won *i.i.d.* conditional on the server with probabilities $p_1, p_2$
- Thus one can compute $P(i \text{ wins match} \mid \theta)$

The importance of the point at state $\theta$ is:

$$P(i \text{ wins match} \mid \theta, i \text{ wins at } \theta) - P(i \text{ wins match} \mid \theta, i \text{ loses at } \theta)$$

- Morris (1977), Klaassen and Magnus (2001), and Paserman (2008)
Decomposition of the importance variable: PiM
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match.
Motivation

Tennis and point importance

Data & Methodology

Results

Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game

Performing Best when it Matters Most

J. González-Díaz, B. Rogers, and O. Gossner
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game
- **GiS** is the importance of the game in the set
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game
- **GiS** is the importance of the game in the set
- **SiM** is the importance of the set in the match
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game
- **GiS** is the importance of the game in the set
- **SiM** is the importance of the set in the match

**Proposition**

\[ PiM = PiG \cdot GiS \cdot SiM \]
Decomposition of the importance variable: $\Pi_{\text{M}}$

- $\Pi_{\text{M}}$ is the importance of the point in the match
- $\Pi_{\text{G}}$ is the importance of the point in the game
- $\Gamma_{\text{S}}$ is the importance of the game in the set
- $\Sigma_{\text{M}}$ is the importance of the set in the match

**Proposition** (Thanks i.i.d!)

\[ \Pi_{\text{M}} = \Pi_{\text{G}} \cdot \Gamma_{\text{S}} \cdot \Sigma_{\text{M}} \]
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game
- **GiS** is the importance of the game in the set
- **SiM** is the importance of the set in the match

**Proposition** (Thanks i.i.d!)

\[ PiM = PiG \cdot GiS \cdot SiM \]
Decomposition of the importance variable: \( \PiM \)

- \( \PiM \) is the importance of the point in the match
- \( \PiG \) is the importance of the point in the game
- \( \GiS \) is the importance of the game in the set
- \( \SiM \) is the importance of the set in the match

**Proposition (Thanks i.i.d!)**

\[
\PiM = \PiG \cdot \GiS \cdot \SiM
\]

The formulas for \( \PiG \), \( \GiS \), and \( \SiM \) are easy to derive.
Decomposition of the importance variable: PiM

- **PiM** is the importance of the point in the match
- **PiG** is the importance of the point in the game
- **GiS** is the importance of the game in the set
- **SiM** is the importance of the set in the match

**Proposition** (Thanks i.i.d!)

\[ PiM = PiG \cdot GiS \cdot SiM \]

The formulas for **PiG**, **GiS**, and **SiM** are easy to derive (from \( p_1 \) and \( p_2 \))
Why not to use break points as importance variables
Why not to use break points as importance variables

Break Point variable

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Why not to use break points as importance variables

Break Point variable

<table>
<thead>
<tr>
<th>$s_2 \setminus s_1$</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

PiG variable

(Average serving percentage in our data set: $p_1 = 0.63$)

<table>
<thead>
<tr>
<th>$s_2 \setminus s_1$</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.24</td>
<td>.18</td>
<td>.10</td>
<td>.03</td>
</tr>
<tr>
<td>15</td>
<td>.34</td>
<td>.31</td>
<td>.22</td>
<td>.09</td>
</tr>
<tr>
<td>30</td>
<td>.39</td>
<td>.45</td>
<td>.44</td>
<td>.25</td>
</tr>
<tr>
<td>40</td>
<td>.30</td>
<td>.47</td>
<td>.75</td>
<td>.44</td>
</tr>
</tbody>
</table>
Why not to use break points as importance variables

Break Point variable

<table>
<thead>
<tr>
<th>( s_2 \backslash s_1 )</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

PiG variable

(Average serving percentage in our data set: \( p_1 = 0.63 \))

<table>
<thead>
<tr>
<th>( s_2 \backslash s_1 )</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.24</td>
<td>.18</td>
<td>.10</td>
<td>.03</td>
</tr>
<tr>
<td>15</td>
<td>0.34</td>
<td>.31</td>
<td>.22</td>
<td>.09</td>
</tr>
<tr>
<td>30</td>
<td>0.39</td>
<td>.45</td>
<td>.44</td>
<td>.25</td>
</tr>
<tr>
<td>40</td>
<td>0.30</td>
<td>.47</td>
<td>.75</td>
<td>.44</td>
</tr>
</tbody>
</table>

Without using GiS and SiM!!
### Why not to use break points as importance variables

#### Break Point variable

<table>
<thead>
<tr>
<th>$s_2 \setminus s_1$</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

#### PiG variable

(Average serving percentage in our data set: $p_1 = 0.63$)

<table>
<thead>
<tr>
<th>$s_2 \setminus s_1$</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.24</td>
<td>.18</td>
<td>.10</td>
<td>.03</td>
</tr>
<tr>
<td>15</td>
<td>.34</td>
<td>.31</td>
<td>.22</td>
<td>.09</td>
</tr>
<tr>
<td>30</td>
<td>.39</td>
<td>.45</td>
<td>.44</td>
<td>.25</td>
</tr>
<tr>
<td>40</td>
<td>.30</td>
<td>.47</td>
<td>.75</td>
<td>.44</td>
</tr>
</tbody>
</table>
Baseline probabilities: $p_1$ and $p_2$
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that PiM is a reasonable measure of the importance of a point.
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that PiM is a reasonable measure of the importance of a point
- How do we get the variables $p_1$ and $p_2$?
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that PiM is a reasonable measure of the importance of a point.
- How do we get the variables $p_1$ and $p_2$?
- Why do not use 0.63, the average serving percentage in our data set?
Baseline probabilities: \( p_1 \) and \( p_2 \)

- By now you might be somewhat convinced that \textbf{PiM} is a reasonable measure of the importance of a point.
- How do we get the variables \( p_1 \) and \( p_2 \)?
- Why do not use 0.63, the average serving percentage in our data set?
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that **PiM** is a reasonable measure of the importance of a point.
- How do we get the variables $p_1$ and $p_2$?
- Why do not use 0.63, the average serving percentage in our data set?

![Graph showing baseline probabilities](image)
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that PiM is a reasonable measure of the importance of a point.
- How do we get the variables $p_1$ and $p_2$?
- Why do not use 0.63, the average serving percentage in our data set?
Baseline probabilities: $p_1$ and $p_2$

- By now you might be somewhat convinced that PiM is a reasonable measure of the importance of a point.
- How do we get the variables $p_1$ and $p_2$?
- Why do not use 0.63, the average serving percentage in our data set?

**Figure:** Red: 30-0, Turquoise: deuce, Green: 0-30, Purple: 0-40
Computing $p_1$ and $p_2$
Computing $p_1$ and $p_2$

- Main analysis: US Open data set
Computing $p_1$ and $p_2$

- Main analysis: US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
Computing $p_1$ and $p_2$

- Main analysis: US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- Preliminary analysis: Pre-US Open tournaments
Computing $p_1$ and $p_2$

- **Main analysis:** US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis:** Pre-US Open tournaments (We use them to get the probabilities)
Computing $p_1$ and $p_2$

- **Main analysis: US Open data set** (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis: Pre-US Open tournaments** (We use them to get the probabilities)

**Pre-US Open tournaments**
Computing $p_1$ and $p_2$

- Main analysis: US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- Preliminary analysis: Pre-US Open tournaments (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federer-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadal-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadal-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roddick-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roddick-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Computing $p_1$ and $p_2$

- **Main analysis: US Open data set** (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis: Pre-US Open tournaments** (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadal-S</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadal-R</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roddick-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Computing $p_1$ and $p_2$

- **Main analysis:** US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis:** Pre-US Open tournaments (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Computing $p_1$ and $p_2$

- **Main analysis: US Open data set** (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis: Pre-US Open tournaments** (We use them to get the probabilities)

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Computing $p_1$ and $p_2$

- **Main analysis:** US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis:** Pre-US Open tournaments (We use them to get the probabilities)

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- **Standard maximum likelihood techniques:**
Computing $p_1$ and $p_2$

- **Main analysis:** US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- **Preliminary analysis:** Pre-US Open tournaments (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- **Standard maximum likelihood techniques:** We get 6 ratings
Computing $p_1$ and $p_2$

- Main analysis: US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- Preliminary analysis: Pre-US Open tournaments (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th></th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- Standard maximum likelihood techniques: We get 6 ratings
  
  $$p_1 = f(r_{1S} - r_{2R})$$

  ($f$ is given by the likelihood function)
Computing $p_1$ and $p_2$

- Main analysis: US Open data set (For each match, we need $p_1$ and $p_2$ to compute PiM variable)
- Preliminary analysis: Pre-US Open tournaments (We use them to get the probabilities)

### Pre-US Open tournaments

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- Standard maximum likelihood techniques: We get 6 ratings
  
  \[
  p_1 = f(r_{1S} - r_{2R}) \quad p_2 = f(r_{2S} - r_{1R})
  \]
  
  ($f$ is given by the likelihood function)
Defining PiM variable
Defining PiM variable

We finally have all the ingredients to define PiM
Defining PiM variable

- We finally have all the ingredients to define PiM
- Results robust to variations in the way to compute $p_1$ and $p_2$
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

Point I

Point II
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500

Point I

Point II
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = .65$ and $p_S = .72$

Point $I$

Point $II$
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match \( p_A = 0.65 \) and \( p_S = 0.72 \)

Point I

- Sampras is serving at 40-0

Point II
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = 0.65$ and $p_S = 0.72$

Point I

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)

Point II
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = .65$ and $p_S = .72$

Point $I$

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)
- He’s leading 2 sets to none

Point $II$
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = .65$ and $p_S = .72$

Point I

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)
- He’s leading 2 sets to none
- $PiM = 8 \cdot 10^{-4}$
- $380$ in current dollars is at stake

Point II
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = .65$ and $p_S = .72$

Point I

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)
- He’s leading 2 sets to none
- $PiM = 8 \cdot 10^{-4}$
- $380$ in current dollars is at stake

Point II

- Sampras is serving at 30-40
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match $p_A = .65$ and $p_S = .72$

**Point I**

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)
- He’s leading 2 sets to none
- $PiM = 8 \cdot 10^{-4}$
- $380$ in current dollars is at stake

**Point II**

- Sampras is serving at 30-40
- It’s 2-2 in the first set
The importance of a point varies substantially

1995 U.S. Open Finals: Agassi vs. Sampras

- The winner earns $575,000, the loser earns $287,500
- Compute from the match \( p_A = 0.65 \) and \( p_S = 0.72 \)

**Point I**

- Sampras is serving at 40-0
- He’s down 2 games to 3 (so attempting to stay on serve)
- He’s leading 2 sets to none
- \( PiM = 8 \cdot 10^{-4} \)
- $380 in current dollars is at stake

**Point II**

- Sampras is serving at 30-40
- It’s 2-2 in the first set
- \( PiM = 0.13 \)
- $64000 in current dollars is at stake
Wrapping up

The importance variable: PiM
Wrapping up

The importance variable: PiM

- Depends on the players’ relative abilities
The importance variable: PiM

- Depends on the players’ relative abilities
- Closer matches have more points with higher importance
Wrapping up

The importance variable: PiM

- Depends on the players’ relative abilities
- Closer matches have more points with higher importance
- As players’ abilities become different, almost all points converge to zero importance
Main Objectives

The goals

The data set
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a **critical ability**

The data set
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a critical ability

The data set
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a **critical ability**
- We want to see if there is heterogeneity across players for each of these variables

The data set
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a critical ability.
- We want to see if there is heterogeneity across players for each of these variables.
- If so, we want to see how these abilities (especially the critical ability) help explain differences in success.

The data set
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a **critical ability**
- We want see if there is heterogeneity across players for each of these variables
- If so, we want to see how these abilities (especially the critical ability) help explain differences in success

The data set

- Point by point data from 12 U.S. Open tournaments, 1994-2006
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a **critical ability**
- We want to see if there is heterogeneity across players for each of these variables
- If so, we want to see how these abilities (especially the critical ability) help explain differences in success

The data set

- Point by point data from 12 U.S. Open tournaments, 1994-2006
- Focus on men singles matches
Main Objectives

The goals

- We want to identify, for each player, a serving ability, a returning ability, and a critical ability
- We want to see if there is heterogeneity across players for each of these variables
- If so, we want to see how these abilities (especially the critical ability) help explain differences in success

The data set

- Point by point data from 12 U.S. Open tournaments, 1994-2006
- Focus on men singles matches
- 1009 matches; 223140 points
Relative abilities
Relative abilities

We can only observe relative abilities
Relative abilities

We can only observe relative abilities

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Relative abilities

We can only observe relative abilities

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Ricardo-S</th>
<th>Ricardo-R</th>
<th>Julio-S</th>
<th>Julio-R</th>
<th>M.Angel-S</th>
<th>M.Angel-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ricardo-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Ricardo-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Julio-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Julio-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Relative abilities

We can only observe relative abilities

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Ricardo-S</th>
<th>Ricardo-R</th>
<th>Julio-S</th>
<th>Julio-R</th>
<th>M.Angel-S</th>
<th>M.Angel-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ricardo-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Ricardo-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Julio-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Julio-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

We can say: “Federer is better than Nadal”
## Relative abilities

We can only observe relative abilities.

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Ricardo-S</th>
<th>Ricardo-R</th>
<th>Julio-S</th>
<th>Julio-R</th>
<th>M.Angel-S</th>
<th>M.Angel-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ricardo-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Ricardo-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Julio-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Julio-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- We **can** say: “Federer is better than Nadal”
- We **can** say: “Roddick is worse than the average player in the data set”
### Relative abilities

**We can only observe relative abilities**

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Federer-S</th>
<th>Federer-R</th>
<th>Nadal-S</th>
<th>Nadal-R</th>
<th>Roddick-S</th>
<th>Roddick-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federer-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Federer-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nadal-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Nadal-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Roddick-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>won \ lost</th>
<th>Ricardo-S</th>
<th>Ricardo-R</th>
<th>Julio-S</th>
<th>Julio-R</th>
<th>M.Angel-S</th>
<th>M.Angel-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ricardo-S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Ricardo-R</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Julio-S</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Julio-R</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-S</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M.Angel-R</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- We **can** say: “Federer is better than Nadal”
- We **can** say: “Roddick is worse than the average player in the data set”
- We **cannot** say: Federer is very good
The more, the better??

Pooling

Final data set
The more, the better??

Having more data may not be beneficial for the analysis.

Pooling

Final data set
The more, the better??

**Having more data may not be beneficial for the analysis**

Identifiability may be a problem

Pooling

Final data set
The more, the better?!

**Having more data may not be beneficial for the analysis**

Identifiability may be a problem
- Each added player requires to estimate three more variables

Pooling

Final data set
The more, the better??

Having more data may not be beneficial for the analysis

Identifiability may be a problem
- Each added player requires to estimate three more variables
- The more connected the players are, the better

Pooling

Final data set
The more, the better??

**Having more data may not be beneficial for the analysis**

Identifiability may be a problem

- Each added player requires to estimate three more variables
- The more connected the players are, the better

Pooling

Final data set
The more, the better??

**Having more data may not be beneficial for the analysis**

Identifiability may be a problem

- Each added player requires to estimate three more variables
- The more connected the players are, the better
- If a player only plays unequal matches we may not get enough variability in \textbf{PiM} variable to identify his \textbf{critical ability}

**Pooling**

**Final data set**
The more, the better??

**Having more data may not be beneficial for the analysis**

Identifiability may be a problem
- Each added player requires to estimate three more variables
- The more connected the players are, the better
- If a player only plays unequal matches we may not get enough variability in \textbf{PiM} variable to identify his \textbf{critical ability}

**Pooling**
- We pool all US-Open tournaments together (connectedness)

**Final data set**
The more, the better??

Having more data may not be beneficial for the analysis

Identifiability may be a problem

- Each added player requires to estimate three more variables
- The more connected the players are, the better
- If a player only plays unequal matches we may not get enough variability in PiM variable to identify his critical ability

Pooling

- We pool all US-Open tournaments together (connectedness)
- No room for intertemporal effects

Final data set
The more, the better??

Having more data may not be beneficial for the analysis

Identifiability may be a problem

- Each added player requires to estimate three more variables
- The more connected the players are, the better
- If a player only plays unequal matches we may not get enough variability in PiM variable to identify his critical ability

Pooling

- We pool all US-Open tournaments together (connectedness)
- No room for intertemporal effects

Final data set

- We take the maximal subset of our data set in which all the remaining players play, at least, 5 matches
The more, the better??

Having more data may not be beneficial for the analysis

Identifiability may be a problem

- Each added player requires to estimate three more variables
- The more connected the players are, the better
- If a player only plays unequal matches we may not get enough variability in PiM variable to identify his critical ability

Pooling

- We pool all US-Open tournaments together (connectedness)
- No room for intertemporal effects

Final data set

- We take the maximal subset of our data set in which all the remaining players play, at least, 5 matches
- We end up with 94 players and about 110000 points
Towards our first regression
Towards our first regression

Computing importance of points
Towards our first regression

Computing importance of points

- For each match, we already have $p_1$ and $p_2$
Towards our first regression

Computing importance of points

- For each match, we already have $p_1$ and $p_2$
- Compute $\text{PiG}$, $\text{GiS}$, $\text{SiM}$ for each score.
Towards our first regression

Computing importance of points

- For each match, we already have $p_1$ and $p_2$
- Compute $\text{PiG}$, $\text{GiS}$, $\text{SiM}$ for each score.
- Use them to compute $\text{PiM}$ for each score, which is what we focus on.
Towards our first regression
Towards our first regression

What should determine the outcome of a point?
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
- Both players’ critical abilities
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
- Both players’ critical abilities

So we want to estimate coefficients of
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
- Both players’ critical abilities

So we want to estimate coefficients of

- Server dummy variables
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
- Both players’ critical abilities

So we want to estimate coefficients of

- Server dummy variables
- Returner dummy variables
Towards our first regression

What should determine the outcome of a point?

- The server’s serving ability
- The returner’s returning ability
- Both players’ critical abilities

So we want to estimate coefficients of

- Server dummy variables
- Returner dummy variables
- Critical ability slope dummy variables
Regression specification
Regression specification

- Suppose we have a point at score $\theta$. 
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$P(\text{Nadal wins} \mid \theta) =$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi(\ )$$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi(\beta^S_N)$$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi\left( \beta^S_N + \beta^C_N \cdot PiM(\theta) \right)$$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_N^S + \beta_N^C \cdot P_iM(\theta) - \beta_F^R) \]
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi(\beta^S_N + \beta^C_N \cdot PiM(\theta) - \beta^R_F - \beta^C_F \cdot PiM(\theta))$$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} | \theta) = \Phi \left( \beta_0 + \beta_N^S + \beta_N^C \cdot PiM(\theta) - \beta_F^R - \beta_F^C \cdot PiM(\theta) \right)$$
Regression specification

- Suppose we have a point at score \( \theta \).
- Nadal (N) serves
- Federer (F) returns

\[
P(\text{Nadal wins} \mid \theta) = \Phi \left( \beta_0 + \beta_N^S + \beta_N^C \cdot P_iM(\theta) - \beta_F^R - \beta_F^C \cdot P_iM(\theta) \right)
\]

We want to identify the \( \beta \) parameters:
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi \left( \beta_0 + \beta^S_N + \beta^C_N \cdot PiM(\theta) - \beta^R_F - \beta^C_F \cdot PiM(\theta) \right)$$

We want to identify the $\beta$ parameters:

$$P(\text{server wins point } p \mid \theta) = \Phi \left( \beta_0 + \sum_{i=1}^{n} (\beta^S_i \delta^S_i + \beta^R_i \delta^R_i + \beta^C_i \delta^C_i \cdot PiM) \right)$$
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi \left( \beta_0 + \beta_N^S + \beta_N^C \cdot P_i M(\theta) - \beta_F^R - \beta_F^C \cdot P_i M(\theta) \right)$$

We want to identify the $\beta$ parameters:

$$P(\text{server wins point } p \mid \theta) = \Phi \left( \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P_i M) \right)$$

Discrete Response Model (Logit regression with two outcomes)

Dependent variable $y$: 
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta_N^S + \beta_N^C \cdot PiM(\theta) - \beta_F^R - \beta_F^C \cdot PiM(\theta))$$

We want to identify the $\beta$ parameters:

$$P(\text{server wins point } p \mid \theta) = \Phi(\beta_0 + \sum_{i=1}^{n}(\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot PiM))$$

**Discrete Response Model (Logit regression with two outcomes)**

**Dependent variable $y$:**
- $y = 1$ if “server wins point”
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

$$
P(\text{Nadal wins} \mid \theta) = \Phi\left(\beta_0 + \beta_N^S + \beta_N^C \cdot P\text{iM}(\theta) - \beta_F^R - \beta_F^C \cdot P\text{iM}(\theta)\right)
$$

We want to identify the $\beta$ parameters:

$$
P(\text{server wins point } p \mid \theta) = \Phi\left(\beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P\text{iM})\right)
$$

**Discrete Response Model (Logit regression with two outcomes)**

**Dependent variable $y$:**

- $y = 1$ if “server wins point”
- $y = 0$ otherwise
Regression specification

- Suppose we have a point at score $\theta$.
- Nadal (N) serves
- Federer (F) returns

\[
P(\text{Nadal wins } | \theta) = \Phi(\beta_0 + \beta_S^N + \beta_C^N \cdot P_iM(\theta) - \beta_R^F - \beta_C^F \cdot P_iM(\theta))
\]

We want to identify the $\beta$ parameters:

\[
P(\text{server wins point } p | \theta) = \Phi(\beta_0 + \sum_{i=1}^{n} (\beta_S^i \delta_S^i + \beta_R^i \delta_R^i + \beta_C^i \delta_C^i \cdot P_iM))
\]

Discrete Response Model (Logit regression with two outcomes)

Dependent variable $y$:
- $y = 1$ if “server wins point”
- $y = 0$ otherwise

\[
y = \beta_0 + \sum_{i=1}^{n} (\beta_S^i \delta_S^i + \beta_R^i \delta_R^i + \beta_C^i \delta_C^i \cdot P_iM)
\]
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta_N^S + \beta_N^C \cdot \text{PiM}(\theta) - \beta_F^R - \beta_F^C \cdot \text{PiM}(\theta)) \]
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi \left( \beta_0 + \beta^S_N + \beta^C_N \cdot \text{PiM}(\theta) - \beta^R_F - \beta^C_F \cdot \text{PiM}(\theta) \right) \]
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta_N^S + \beta_N^C \cdot \text{PiM}(\theta) - \beta_F^R - \beta_F^C \cdot \text{PiM}(\theta)) \]

- PiM is defined as a positive variable
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta_S^N + \beta_{C}^N \cdot P_iM(\theta) - \beta_{RF}^C - \beta_{CF}^C \cdot P_iM(\theta)) \]

- PiM is defined as a positive variable
- Having a positive critical ability (\( \beta_{C}^N \)) implies winning more points in general
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta^S_N + \beta^C_N \cdot \text{PiM}(\theta) - \beta^R_F - \beta^C_F \cdot \text{PiM}(\theta)) \]

- PiM is defined as a positive variable
- Having a positive critical ability \((\beta^C_N)\) implies winning more points in general \textbf{We do not want this!!}
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi \left( \beta_0 + \beta^S_N + \beta^C_N \cdot \text{PiM}(\theta) - \beta^R_F - \beta^C_F \cdot \text{PiM}(\theta) \right) \]

- PiM is defined as a positive variable
- Having a positive critical ability (\(\beta^C_N\)) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi\left(\beta_0 + \beta^S_N + \beta^C_N \cdot \text{PiM}(\theta) - \beta^R_F - \beta^C_F \cdot \text{PiM}(\theta)\right) \]

- PiM is defined as a positive variable
- Having a positive critical ability \((\beta^C_N)\) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
- We have to demean PiM
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta_S^N + \beta_C^N \cdot PiM(\theta) - \beta_R^F - \beta_C^F \cdot PiM(\theta)) \]

- PiM is defined as a positive variable
- Having a positive critical ability \((\beta_C^N)\) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
- We have to demean PiM
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi\left(\beta_0 + \beta_S^N + \beta_C^N \cdot \PiM(\theta) - \beta_R^F - \beta_C^F \cdot \PiM(\theta)\right) \]

- PiM is defined as a positive variable
- Having a positive critical ability (\(\beta_C^N\)) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
- We have to demean PiM
  \[ \PiM_{\text{Demeaned}}(\theta) = \PiM(\theta) - \text{mean}(\PiM(\theta)) \]
Demeaning PiM

\[ P(\text{Nadal wins} \mid \theta) = \Phi(\beta_0 + \beta^S_N + \beta^C_N \cdot P_iM(\theta) - \beta^R_F - \beta^C_F \cdot P_iM(\theta)) \]

- PiM is defined as a positive variable
- Having a positive critical ability \((\beta^C_N)\) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
- We have to demean PiM at the match level
  \[ P_{iM\text{Demeaned}}(\theta) = P_iM(\theta) - \text{mean}(P_iM(\theta)) \]
Demeaning PiM

\[
P(Nadal \text{ wins } | \theta) = \Phi \left( \beta_0 + \beta_N^S + \beta_N^C \cdot P_iM(\theta) - \beta_F^R - \beta_F^C \cdot P_iM(\theta) \right)
\]

- PiM is defined as a positive variable
- Having a positive critical ability (\( \beta_N^C \)) implies winning more points in general
- We want critical ability to imply winning more important points and less unimportant points
- We have to demean PiM at the match level
  \[
P_{Demeaned}(\theta) = P_iM(\theta) - \text{mean}(P_iM(\theta))
\]

After the demeaning, the critical ability does not affect the average probability of winning a point
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot PiM) \]

PiM represents demeaned PiM
94*3=282 variables. We do not run tests at the individual level
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n}(\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot PiM) \]

PiM represents demeaned PiM
94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th></th>
<th>p-value</th>
<th>p-value</th>
<th>p-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Serving</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Returning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Critical (PiM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Performing Best when it Matters Most
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P_{iM}) \]

PiM represents demeaned PiM
94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th>Serving</th>
<th>Returning</th>
<th>Critical (PiM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
</tr>
<tr>
<td>0***</td>
<td></td>
<td>0.06*</td>
</tr>
<tr>
<td>0***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Performing Best when it Matters Most  
J. González-Díaz, B. Rogers, and O. Gossner
Motivation

Tennis and point importance

Data & Methodology

Results

Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} \left( \beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot \text{PiM} \right) \]

PiM represents demeaned PiM

94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th>Serving</th>
<th>Returning</th>
<th>Critical (PiM)</th>
<th>Critical (BP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p )-value</td>
<td>( p )-value</td>
<td>( p )-value</td>
<td>( p )-value</td>
</tr>
</tbody>
</table>
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P_i M) \]

PiM represents demeaned PiM
94*3 = 282 variables. We do not run tests at the individual level

**Joint significance tests**

<table>
<thead>
<tr>
<th></th>
<th>Serving</th>
<th>Returning</th>
<th>Critical (PiM)</th>
<th>Critical (BP)</th>
<th>Critical (PiM-0.63)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>p-value</strong></td>
<td>0***</td>
<td>0***</td>
<td>0.06*</td>
<td>0.79</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0***</td>
<td>0***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0***</td>
<td>0***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot PIM) \]

PiM represents demeaned PiM
94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th>Serving</th>
<th>p-value</th>
<th>Returning</th>
<th>p-value</th>
<th>Critical (PiM)</th>
<th>p-value</th>
<th>Critical (BP)</th>
<th>p-value</th>
<th>Critical (PiM-0.63)</th>
<th>p-value</th>
<th>Point id (demeaned)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0***</td>
<td></td>
<td>0***</td>
<td></td>
<td>0***</td>
<td></td>
<td>0***</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta^S_i \delta^S_i + \beta^R_i \delta^R_i + \beta^C_i \delta^C_i \cdot \text{PiM}) \]

PiM represents demeaned PiM

94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th></th>
<th>Serving</th>
<th>Returning</th>
<th>Critical (PiM)</th>
<th>Critical (BP)</th>
<th>Critical (PiM-0.63)</th>
<th>Point id (demeaned)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0.06*</td>
<td>0.0002***</td>
</tr>
<tr>
<td></td>
<td>0***</td>
<td>0***</td>
<td>0.79</td>
<td>0.38</td>
<td>0.01***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P_iM) \]

PiM represents demeaned PiM

94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th></th>
<th>(p)-value</th>
<th>(p)-value</th>
<th>(p)-value</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
</tr>
<tr>
<td>Returning</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
</tr>
<tr>
<td>Critical (PiM)</td>
<td>0.06*</td>
<td>0.79</td>
<td>0.38</td>
<td>0.0002***</td>
</tr>
<tr>
<td>Critical (BP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical (PiM-0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point id (demeaned)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- There is heterogeneity in serving and returning abilities
Motivation

Tennis and point importance

Data & Methodology

Results

Results of the first regression

\[ y = \beta_0 + \sum_{i=1}^{n} (\beta_i^S \delta_i^S + \beta_i^R \delta_i^R + \beta_i^C \delta_i^C \cdot P_{iM}) \]

PiM represents demeaned PiM

94*3=282 variables. We do not run tests at the individual level

Joint significance tests

<table>
<thead>
<tr>
<th></th>
<th>Serving</th>
<th>Returning</th>
<th>Critical (PiM)</th>
<th>Critical (BP)</th>
<th>Critical (PiM-0.63)</th>
<th>Point id (demeaned)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
</tr>
<tr>
<td>Serving</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0.06*</td>
<td>0.0002***</td>
</tr>
<tr>
<td>Returning</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0***</td>
<td>0.79</td>
<td>0.01***</td>
</tr>
<tr>
<td>Critical (PiM)</td>
<td></td>
<td></td>
<td>0***</td>
<td>0***</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Critical (BP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical (PiM-0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point id (demeaned)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- There is heterogeneity in serving and returning abilities
- It seems there is also heterogeneity in critical abilities
### Results of the first regression

<table>
<thead>
<tr>
<th>Top 25</th>
<th>Serving</th>
<th>Point Estimate</th>
<th>Returning</th>
<th>Point Estimate</th>
<th>Critical</th>
<th>Point Estimate</th>
<th>ATP Rating</th>
<th>log(rating)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A.RODDICK</td>
<td>0.268</td>
<td>L.HEWITT</td>
<td>-0.457</td>
<td>T.ROBREDO</td>
<td>6.026</td>
<td>P.SAMPRA</td>
<td>8.109</td>
</tr>
<tr>
<td>2</td>
<td>P.SAMPRAS</td>
<td>0.203</td>
<td>R.FEDERER</td>
<td>-0.485</td>
<td>A.CORRETJA</td>
<td>4.491</td>
<td>R.FEDERER</td>
<td>8.077</td>
</tr>
<tr>
<td>3</td>
<td>R.KRAJICEK</td>
<td>0.145</td>
<td>K.KUCERA</td>
<td>-0.5</td>
<td>J.FERRERO</td>
<td>2.892</td>
<td>M.STICH</td>
<td>7.856</td>
</tr>
<tr>
<td>4</td>
<td>R.FEDERER</td>
<td>0.065</td>
<td>A.AGASSI</td>
<td>-0.519</td>
<td>A.COSTA</td>
<td>2.631</td>
<td>L.HEWITT</td>
<td>7.813</td>
</tr>
<tr>
<td>5</td>
<td>M.MIRNYI</td>
<td>0.036</td>
<td>J.BJORKMAN</td>
<td>-0.534</td>
<td>M.ROSSET</td>
<td>1.434</td>
<td>A.AGASSI</td>
<td>7.780</td>
</tr>
<tr>
<td>6</td>
<td>M.STICH</td>
<td>0.017</td>
<td>M.YOUSHNY</td>
<td>-0.549</td>
<td>M.ZABALET</td>
<td>1.43</td>
<td>A.RODDICK</td>
<td>7.779</td>
</tr>
<tr>
<td>7</td>
<td>A.AGASSI</td>
<td>0.000</td>
<td>N.ESCUDE</td>
<td>-0.558</td>
<td>G.POZZI</td>
<td>1.324</td>
<td>Y.KAFELNIKOV</td>
<td>7.766</td>
</tr>
<tr>
<td>8</td>
<td>P.RAFTER</td>
<td>-0.002</td>
<td>Y.KAFELNIKOV</td>
<td>-0.562</td>
<td>R.SCHUETTLER</td>
<td>1.257</td>
<td>G.KUERTEN</td>
<td>7.742</td>
</tr>
<tr>
<td>9</td>
<td>G.RUSEDSKI</td>
<td>-0.005</td>
<td>D.NALBANDIAN</td>
<td>-0.574</td>
<td>A.RODDICK</td>
<td>1.043</td>
<td>T.MUSTIER</td>
<td>7.702</td>
</tr>
<tr>
<td>10</td>
<td>N.ESCUDE</td>
<td>-0.024</td>
<td>G.CORIA</td>
<td>-0.577</td>
<td>G.IVANISEVIC</td>
<td>0.872</td>
<td>J.FERRERO</td>
<td>7.583</td>
</tr>
<tr>
<td>11</td>
<td>G.KUERTEN</td>
<td>-0.040</td>
<td>J.BLAKE</td>
<td>-0.582</td>
<td>B.BLACK</td>
<td>0.841</td>
<td>P.RAFTER</td>
<td>7.573</td>
</tr>
<tr>
<td>12</td>
<td>L.HEWITT</td>
<td>-0.047</td>
<td>P.KORDA</td>
<td>-0.587</td>
<td>M.WOODFORDE</td>
<td>0.799</td>
<td>R.NADAL</td>
<td>7.523</td>
</tr>
<tr>
<td>13</td>
<td>M.LARSSON</td>
<td>-0.047</td>
<td>A.RODDICK</td>
<td>-0.603</td>
<td>B.KARBACHER</td>
<td>0.762</td>
<td>P.KORDA</td>
<td>7.431</td>
</tr>
<tr>
<td>14</td>
<td>M.SAFIN</td>
<td>-0.081</td>
<td>G.CANAS</td>
<td>-0.617</td>
<td>P.SAMPRAS</td>
<td>0.655</td>
<td>T.HENMAN</td>
<td>7.407</td>
</tr>
<tr>
<td>15</td>
<td>B.BECKER</td>
<td>-0.087</td>
<td>D.HRBATY</td>
<td>-0.634</td>
<td>L.HEWITT</td>
<td>0.637</td>
<td>C.MOYA</td>
<td>7.377</td>
</tr>
<tr>
<td>16</td>
<td>T.MARTIN</td>
<td>-0.087</td>
<td>P.RAFTER</td>
<td>-0.638</td>
<td>N.ESCUDE</td>
<td>0.569</td>
<td>D.NALBANDIAN</td>
<td>7.37</td>
</tr>
<tr>
<td>17</td>
<td>J.BLAKE</td>
<td>-0.092</td>
<td>V.SPADEA</td>
<td>-0.651</td>
<td>A.MEDVEDEV</td>
<td>0.261</td>
<td>A.CORRETJA</td>
<td>7.293</td>
</tr>
<tr>
<td>18</td>
<td>X.MALISSE</td>
<td>-0.101</td>
<td>H.LEE</td>
<td>-0.655</td>
<td>P.RAFTER</td>
<td>0.151</td>
<td>M.SAFIN</td>
<td>7.281</td>
</tr>
<tr>
<td>19</td>
<td>W.ARTHURS</td>
<td>-0.104</td>
<td>S.SARGSIAN</td>
<td>-0.664</td>
<td>S.DOSEDEL</td>
<td>0.098</td>
<td>B.BECKER</td>
<td>7.264</td>
</tr>
<tr>
<td>20</td>
<td>M.ZABALET</td>
<td>-0.112</td>
<td>J.COURIER</td>
<td>-0.674</td>
<td>M.SAFIN</td>
<td>0.096</td>
<td>R.KRAJICEK</td>
<td>7.236</td>
</tr>
<tr>
<td>21</td>
<td>M.DAMM</td>
<td>-0.121</td>
<td>M.ZABALET</td>
<td>-0.678</td>
<td>W.ARTHURS</td>
<td>0.009</td>
<td>G.CORIA</td>
<td>7.22</td>
</tr>
<tr>
<td>22</td>
<td>J.COURIER</td>
<td>-0.134</td>
<td>T.ENQVIST</td>
<td>-0.68</td>
<td>A.AGASSI</td>
<td>0.0</td>
<td>J.COURIER</td>
<td>7.192</td>
</tr>
<tr>
<td>23</td>
<td>D.NALBANDIAN</td>
<td>-0.135</td>
<td>A.CLEMENT</td>
<td>-0.683</td>
<td>R.FEDERER</td>
<td>-0.196</td>
<td>C.PIOLINE</td>
<td>7.188</td>
</tr>
<tr>
<td>24</td>
<td>R.GINEPRI</td>
<td>-0.135</td>
<td>T.HAAS</td>
<td>-0.685</td>
<td>C.PIOLINE</td>
<td>-0.236</td>
<td>T.ROBREDO</td>
<td>7.168</td>
</tr>
<tr>
<td>25</td>
<td>J.FERRERO</td>
<td>-0.139</td>
<td>M.SAFIN</td>
<td>-0.689</td>
<td>T.MARTIN</td>
<td>-0.238</td>
<td>A.MEDVEDEV</td>
<td>7.16</td>
</tr>
</tbody>
</table>
Towards our second regression
Towards our second regression

- How much serving, returning, and critical abilities explain of a player’s success?
Towards our second regression

- How much serving, returning, and critical abilities explain of a player’s success?
- We regress on ATP ratings and rankings
Towards our second regression

- How much serving, returning, and critical abilities explain of a player's success?
- We regress on ATP ratings and rankings

\[ \text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical} \]
Towards our second regression

- How much serving, returning, and critical abilities explain of a player’s success?
- We regress on ATP ratings and rankings

\[
\text{ATP rating} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

\[
\text{ATP ranking} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]
Towards our second regression

- How much serving, returning, and critical abilities explain of a player’s success?
- We regress on ATP ratings and rankings

\[
\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

\[
\text{ATP}_{\text{ranking}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

Actually, log(ratings) and log(rankings)
Towards our second regression

- How much serving, returning, and critical abilities explain of a player’s success?
- We regress on ATP ratings and rankings

\[
\text{ATP}^{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

\[
\text{ATP}^{\text{ranking}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

Actually, log(ratings) and log(rankings)

<table>
<thead>
<tr>
<th>correlations</th>
<th>log(rating)</th>
<th>Serving</th>
<th>Returning</th>
<th>Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(rating)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Serving</td>
<td>0.57</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Returning</td>
<td>0.39</td>
<td>0.26</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Critical</td>
<td>0.38</td>
<td>0.34</td>
<td>0.20</td>
<td>1</td>
</tr>
</tbody>
</table>
Towards our second regression

\[ \text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical} \]
Towards our second regression

\[ \text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical} \]

Technical problems
Towards our second regression

\[
\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

Technical problems

- What regression to run? OLS?
Towards our second regression

$$\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}$$

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
Towards our second regression

$$\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}$$

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
- Their errors may be correlated
Towards our second regression

\[ \text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical} \]

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
- Their errors may be correlated
- Standard OLS may lead to wrong confidence intervals
Towards our second regression

\[
\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
- Their errors may be correlated
- Standard OLS may lead to wrong confidence intervals

What do we do?
Towards our second regression

\[
\text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical}
\]

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
- Their errors may be correlated
- Standard OLS may lead to wrong confidence intervals

What do we do?

- We run a standard OLS
Towards our second regression

\[ \text{ATP}_{\text{rating}} = \alpha_0 + \alpha_1 \cdot \text{Serving} + \alpha_2 \cdot \text{Returning} + \alpha_3 \cdot \text{Critical} \]

Technical problems

- What regression to run? OLS?
- Serving, Returning, and Critical are estimated variables
- Their errors may be correlated
- Standard OLS may lead to wrong confidence intervals

What do we do?

- We run a standard OLS
- We check robustness of results via GLS and bootstrap
## Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>No controls</th>
<th>Some controls</th>
<th>More controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>No controls</th>
<th>Some controls</th>
<th>More controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>7.86***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serving</td>
<td>1.48***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td>0.78***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td>0.036**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>No controls</th>
<th>Some controls</th>
<th>More controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>7.86***</td>
<td>7.32***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.15)</td>
<td></td>
</tr>
<tr>
<td>Serving</td>
<td>1.48***</td>
<td>1.22***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.29)</td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td>0.78***</td>
<td>0.77***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.29)</td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td>0.036**</td>
<td>0.031**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>Endurance (via pointid)</td>
<td></td>
<td>66.85</td>
<td>(52.63)</td>
</tr>
</tbody>
</table>
## Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>No controls</th>
<th>Some controls</th>
<th>More controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Intercept)</strong></td>
<td>7.86***</td>
<td>7.32***</td>
<td>6.14</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.15)</td>
<td>(20)</td>
</tr>
<tr>
<td><strong>Serving</strong></td>
<td>1.48***</td>
<td>1.22***</td>
<td>1.23***</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.33)</td>
</tr>
<tr>
<td><strong>Returning</strong></td>
<td>0.78***</td>
<td>0.77***</td>
<td>0.87***</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.29)</td>
<td>(0.33)</td>
</tr>
<tr>
<td><strong>Critical</strong></td>
<td>0.036**</td>
<td>0.031**</td>
<td>0.034**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.016)</td>
</tr>
<tr>
<td><strong>Endurance (via pointid)</strong></td>
<td>66.85</td>
<td>53.75</td>
<td>(56.45)</td>
</tr>
<tr>
<td></td>
<td>(52.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Birth year</strong></td>
<td></td>
<td></td>
<td>0.00056</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td><strong>Height</strong></td>
<td></td>
<td></td>
<td>0.00065</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0077)</td>
</tr>
<tr>
<td><strong>Lefty</strong></td>
<td></td>
<td>0.098</td>
<td>(0.12)</td>
</tr>
<tr>
<td><strong>GDP</strong></td>
<td>−0.0000003</td>
<td></td>
<td>(0.0000002)</td>
</tr>
<tr>
<td></td>
<td>(0.000002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bollettieri</strong></td>
<td></td>
<td>0.15</td>
<td>(0.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td><strong>USA</strong></td>
<td></td>
<td>−0.18</td>
<td>(0.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13)</td>
<td></td>
</tr>
</tbody>
</table>

Performing Best when it Matters Most, J. González-Díaz, B. Rogers, and O. Gossner
Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>Coefficient in R²</th>
<th>Standard deviation</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>Coefficient in R2</th>
<th>Standard deviation</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving</td>
<td>1.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td>0.036</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>Coefficient in R2</th>
<th>Standard deviation</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving</td>
<td>1.48</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Returning</td>
<td>0.78</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Critical</td>
<td>0.036</td>
<td>2.56</td>
<td></td>
</tr>
</tbody>
</table>
Results of the second regression

<table>
<thead>
<tr>
<th></th>
<th>Coefficient in R2</th>
<th>Standard deviation</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving</td>
<td>1.48</td>
<td>0.15</td>
<td>0.22</td>
</tr>
<tr>
<td>Returning</td>
<td>0.78</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>Critical</td>
<td>0.036</td>
<td>2.56</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Performing Best when it Matters Most

J. González-Díaz, B. Rogers, and O. Gossner
Conclusions

1. There is heterogeneity in agent’s reactions to changes in the importance of the situation.
2. This heterogeneity has a significant impact on an agent’s career.
Performing Best when it Matters Most: Evidence from professional tennis

Julio González-Díaz 1 Olivier Gossner 2 Brian Rogers 3

1Research Group in Economic Analysis
Universidad de Vigo

2London School of Economics
University of London

3Kellogg School of Management
Northwestern University

October 16th, 2009