Identification and Estimation in Manipulable Assignment Mechanisms

Nikhil Agarwal ¹  Paulo Somaini ²

¹MIT and NBER

²MIT, Cowles Visitor, and NBER
Estimates of agent preferences are often essential for economic analysis

- Primitives determining incentives and actions.
- Allocative efficiency and distributional consequences of counterfactual policies.
Introduction

- Estimates of agent preferences are often essential for economic analysis
  - Primitives determining incentives and actions.
  - Allocative efficiency and distributional consequences of counterfactual policies.

- Can sometimes directly interpret behavior in terms of preferences
  - A chosen product is revealed to provide the highest indirect utility
  - Agents (should) reveal their type to dominant strategy mechanisms
Estimates of agent preferences are often essential for economic analysis
- Primitives determining incentives and actions.
- Allocative efficiency and distributional consequences of counterfactual policies.

Can sometimes directly interpret behavior in terms of preferences
- A chosen product is revealed to provide the highest indirect utility
- Agents (should) reveal their type to dominant strategy mechanisms

Subtle link between actions and preferences if incentives are complex
- Bid shading in first price-auctions
- School choice: in the Boston Mechanism parents may avoid applying to a school that is too competitive even if they prefer it over all other options.
Estimates of agent preferences are often essential for economic analysis
▶ Primitives determining incentives and actions.
▶ Allocative efficiency and distributional consequences of counterfactual policies.

Can sometimes directly interpret behavior in terms of preferences
▶ A chosen product is revealed to provide the highest indirect utility
▶ Agents (should) reveal their type to dominant strategy mechanisms

Subtle link between actions and preferences if incentives are complex
▶ Bid shading in first price-auctions
▶ School choice: in the Boston Mechanism parents may avoid applying to a school that is too competitive even if they prefer it over all other options.

**Research Question**
✓ Can we learn agent preferences from data on ordinal reports submitted to manipulable assignment mechanisms?
<table>
<thead>
<tr>
<th>Manipulable Mechanism</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boston Mechanism</strong></td>
<td>Barcelona, Beijing, Boston (pre 2005), Charlotte-Macklenberg, Chicago (pre 2009), Denver, Miami-Dade, Minneapolis, Seattle (pre 1999 and post 2009), Tampa-St. Petersburg.</td>
</tr>
<tr>
<td>Deferred Acceptance</td>
<td>New York City, Ghanian Schools, various districts in England (since mid '00s)</td>
</tr>
<tr>
<td>w/ Truncated Lists</td>
<td></td>
</tr>
<tr>
<td><strong>Serial Dictatorships</strong></td>
<td>Chicago (2009 onwards)</td>
</tr>
<tr>
<td>w/ Truncated Lists</td>
<td></td>
</tr>
<tr>
<td><strong>First Priority First</strong></td>
<td>various districts in England (before mid '00s)</td>
</tr>
<tr>
<td><strong>Chinese Parallel</strong></td>
<td>Shanghai and several other Chinese provinces</td>
</tr>
<tr>
<td><strong>Cambridge Pan London Admissions</strong></td>
<td>Cambridge, London</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-Manipulable Mechanism</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Top Trading Cycles</strong></td>
<td>New Orleans</td>
</tr>
<tr>
<td><strong>Deferred Acceptance</strong></td>
<td>Boston (post 2005), Seattle (1999-2008)</td>
</tr>
<tr>
<td>w/ Unrestricted Lists</td>
<td></td>
</tr>
</tbody>
</table>
Overview: Model

- A single-unit assignment mechanism without monetary transfers
  - The mechanism uses reported ordinal preferences to allocate the objects
  - Examples include Serial Dictatorships, Deferred Acceptance, Boston Mechanism
Overview: Model

- A single-unit assignment mechanism without monetary transfers
  - The mechanism uses reported ordinal preferences to allocate the objects
  - Examples include Serial Dictatorships, Deferred Acceptance, Boston Mechanism

- Researcher has data from a market with many participants and few choices
  - Submitted reports (rank-order lists)
  - Knows the assignment mechanism and capacities
Overview: Model

- A single-unit assignment mechanism without monetary transfers
  - The mechanism uses reported ordinal preferences to allocate the objects
  - Examples include Serial Dictatorships, Deferred Acceptance, Boston Mechanism

- Researcher has data from a market with many participants and few choices
  - Submitted reports (rank-order lists)
  - Knows the assignment mechanism and capacities
  - ... and maybe, characteristics of schools and students
A single-unit assignment mechanism without monetary transfers
- The mechanism uses reported ordinal preferences to allocate the objects
- Examples include Serial Dictatorships, Deferred Acceptance, Boston Mechanism

Researcher has data from a market with many participants and few choices
- Submitted reports (rank-order lists)
- Knows the assignment mechanism and capacities
- ... and maybe, characteristics of schools and students
Overview: Analysis

Descriptive Evidence from Cambridge:
- Significant incentive to pick the top rank carefully
  - 82% of the students are assigned to their stated top choice
- Students respond to strategic incentives
  - Ranking behavior is discontinuous at the geographical boundary where proximity priority changes

Estimation Strategy:
- First step: estimate probability of assignment as a function of reports
  - Quantifies competition for school seats and risk faced by students
  - Derive conditions for consistency in large markets.
- Second step: estimate preference parameters, assuming reports are optimal
  - Key insight: From a student's perspective, each report yields a probability distribution over her assignment
  - Point identification under "special regressor".
  - We propose and implement a tractable estimation method based on MCMC.

Application (Cambridge):
- Compare preference estimates from alternative assumptions on behavior
- What fraction of students are assigned to their true top choice?
- Counterfactual: what would be the effect of switching to a strategy proof mechanism such as Deferred Acceptance.
Description of Evidence from Cambridge:
- Significant incentive to pick the top rank carefully
  - 82% of the students are assigned to their stated top choice
- Students respond to strategic incentives
  - Ranking behavior is discontinuous at the geographical boundary where proximity priority changes

Estimation Strategy:
- First step: estimate probability of assignment as a function of reports
  - Quantifies competition for school seats and risk faced by students
  - Derive conditions for consistency in large markets.
- Second-step: estimate preference parameters, assuming reports are optimal
Overview: Analysis

- **Descriptive Evidence from Cambridge:**
  - Significant incentive to pick the top rank carefully
    - 82% of the students are assigned to their stated top choice
  - Students respond to strategic incentives
    - Ranking behavior is discontinuous at the geographical boundary where proximity priority changes

- **Estimation Strategy:**
  - First step: estimate probability of assignment as a function of reports
    - Quantifies competition for school seats and risk faced by students
    - Derive conditions for consistency in large markets.
  - Second-step: estimate preference parameters, assuming reports are optimal
    - Key insight: From an student’s perspective, each report yields a probability distribution over her assignment.
    - Point identification under “special regressor”.
    - We propose and implement a tractable estimation method based on MCMC.

- **Application (Cambridge):**
  - Compare preference estimates from alternative assumptions on behavior
  - What fraction of students are assigned to their true top choice?
  - Counterfactual: what would be the effect of switching to a strategy proof mechanism such as Deferred Acceptance.
Related Literature

- **Estimation of Preferences in School Choice**
  - Abdulkadiroglu, Agarwal, Pathak (2014); He (2012); Hastings, Kane and Staiger (2009)

- **Large Market Matching Mechanisms and Core**
  - Azevedo and Leshno (2013); Azevedo and Budish (2012); Kojima and Pathak (2009)

- **Estimation and Identification of Valuations in Auctions**
  - Athey and Haile (2005); Guerre, Perrigne and Vuong (2000); Somaini (2014); amongst others

- **Identification of Multinomial Choice Models**
  - Berry and Haile (2010); Lewbel (2000); amongst others

- **Identification and Estimation of Discrete Games**
  - Ciliberto and Tamer (2009); Beresteanu and Molinari (2008)
Plan

1. Descriptive: Cambridge Public Schools
2. Estimating Preferences: Overview
3. First Step: Assignment Probabilities
4. Second Step: Preferences
5. Estimation and Estimates from Cambridge
Plan

1. Descriptive: Cambridge Public Schools
   - District Characteristics and the Controlled Choice Plan
   - Behavior and Strategic Incentives
   - Evidence on Strategic Behavior

2. Estimating Preferences: Overview

3. First Step: Assignment Probabilities

4. Second Step: Preferences

5. Estimation and Estimates from Cambridge
## Elementary Schools and Students

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: District Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Programs</td>
<td>24</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>25.6</td>
</tr>
<tr>
<td>Seats</td>
<td>473</td>
<td>456</td>
<td>476</td>
<td>508</td>
<td>438</td>
<td>470</td>
</tr>
<tr>
<td>Students</td>
<td>412</td>
<td>432</td>
<td>397</td>
<td>457</td>
<td>431</td>
<td>426</td>
</tr>
<tr>
<td>Free/Reduced Lunch</td>
<td>32%</td>
<td>38%</td>
<td>37%</td>
<td>29%</td>
<td>32%</td>
<td>34%</td>
</tr>
<tr>
<td>Paid Lunch</td>
<td>68%</td>
<td>62%</td>
<td>63%</td>
<td>71%</td>
<td>68%</td>
<td>66%</td>
</tr>
<tr>
<td><strong>Panel B: Student's Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>47%</td>
<td>47%</td>
<td>45%</td>
<td>49%</td>
<td>49%</td>
<td>47%</td>
</tr>
<tr>
<td>Black</td>
<td>27%</td>
<td>22%</td>
<td>24%</td>
<td>22%</td>
<td>23%</td>
<td>24%</td>
</tr>
<tr>
<td>Asian</td>
<td>17%</td>
<td>18%</td>
<td>15%</td>
<td>13%</td>
<td>18%</td>
<td>16%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9%</td>
<td>11%</td>
<td>10%</td>
<td>9%</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Panel C: Language spoken at home</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>72%</td>
<td>73%</td>
<td>73%</td>
<td>78%</td>
<td>81%</td>
<td>76%</td>
</tr>
<tr>
<td>Spanish</td>
<td>3%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Portuguese</td>
<td>0%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td><strong>Panel D: Distances (miles)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closest School</td>
<td>0.43</td>
<td>0.67</td>
<td>0.43</td>
<td>0.47</td>
<td>0.45</td>
<td>0.49</td>
</tr>
<tr>
<td>Average School</td>
<td>1.91</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.89</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Notes: Students participating in the January Kindergarten Lottery. Free/Reduced lunch based on student's application for Federal lunch subsidy.
The Cambridge Mechanism

- Two types of student priorities:
  1. Students with siblings that are attending that school get the highest priority
  2. Students receive priority at the two schools closest to their residence
The Cambridge Mechanism

Two types of student priorities:

1. Students with siblings that are attending that school get the highest priority
2. Students receive priority at the two schools closest to their residence

Students rank up to three school programs

Step 0: Generate a single lottery for each student
Step 1: Arrange first choice applicants by priority/lottery
   (a) Top student is assigned to the paid or free/reduced lunch program based on eligibility
   (b) Iterate until all students are considered
Step k: Students unassigned to previous $k - 1$ choices are considered
The Cambridge Mechanism

- Two types of student priorities:
  1. Students with siblings that are attending that school get the highest priority
  2. Students receive priority at the two schools closest to their residence

- Students rank up to three school programs

  Step 0: Generate a single lottery for each student
  Step 1: Arrange first choice applicants by priority/lottery
    (a) Top student is assigned to the paid or free/reduced lunch program based on eligibility
    (b) Iterate until all students are considered
  Step k: Students unassigned to previous $k - 1$ choices are considered

- Capacity specific to paid and free/reduced lunch programs
### Ranking Behavior

#### Panel A: Round of Assignment

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>81%</td>
<td>84%</td>
<td>85%</td>
<td>83%</td>
<td>75%</td>
<td>82%</td>
</tr>
<tr>
<td>Second</td>
<td>8%</td>
<td>3%</td>
<td>4%</td>
<td>7%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Third</td>
<td>5%</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Unassigned</td>
<td>6%</td>
<td>11%</td>
<td>9%</td>
<td>8%</td>
<td>16%</td>
<td>10%</td>
</tr>
</tbody>
</table>

#### Panel B: Round of Assignment: Paid Lunch Students

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>80%</td>
<td>77%</td>
<td>78%</td>
<td>79%</td>
<td>68%</td>
<td>76%</td>
</tr>
<tr>
<td>Second</td>
<td>5%</td>
<td>4%</td>
<td>5%</td>
<td>8%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Third</td>
<td>6%</td>
<td>3%</td>
<td>4%</td>
<td>2%</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td>Unassigned</td>
<td>9%</td>
<td>16%</td>
<td>14%</td>
<td>11%</td>
<td>24%</td>
<td>15%</td>
</tr>
</tbody>
</table>

#### Panel C: Round of Assignment: Free Lunch Students

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>85%</td>
<td>95%</td>
<td>98%</td>
<td>94%</td>
<td>89%</td>
<td>92%</td>
</tr>
<tr>
<td>Second</td>
<td>14%</td>
<td>1%</td>
<td>2%</td>
<td>4%</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>Third</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td>1%</td>
<td>4%</td>
<td>1%</td>
</tr>
<tr>
<td>Unassigned</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>

#### Panel D: Number of Programs Ranked

<table>
<thead>
<tr>
<th>Year</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>2%</td>
<td>5%</td>
<td>93%</td>
</tr>
<tr>
<td>2005</td>
<td>6%</td>
<td>6%</td>
<td>89%</td>
</tr>
<tr>
<td>2006</td>
<td>9%</td>
<td>9%</td>
<td>82%</td>
</tr>
<tr>
<td>2007</td>
<td>5%</td>
<td>7%</td>
<td>88%</td>
</tr>
<tr>
<td>2008</td>
<td>12%</td>
<td>7%</td>
<td>81%</td>
</tr>
<tr>
<td>Average</td>
<td>7%</td>
<td>7%</td>
<td>87%</td>
</tr>
</tbody>
</table>

Notes: Proximity priority as reported in the Cambridge Public School assignment files.
## Strategic Incentives

<table>
<thead>
<tr>
<th>School</th>
<th>Graham Parks</th>
<th>Haggerty</th>
<th>Baldwin</th>
<th>Morse</th>
<th>Amigos</th>
<th>Cambridgeport</th>
<th>King Open</th>
<th>Peabody</th>
<th>Tobin</th>
<th>Flet Mayn</th>
<th>Kenn Long</th>
<th>MLK</th>
<th>King Open Ola</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: All Students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ranked First</td>
<td>60</td>
<td>56</td>
<td>53</td>
<td>47</td>
<td>37</td>
<td>34</td>
<td>33</td>
<td>31</td>
<td>25</td>
<td>18</td>
<td>16</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Ranked Anywhere</td>
<td>192</td>
<td>120</td>
<td>166</td>
<td>102</td>
<td>75</td>
<td>113</td>
<td>114</td>
<td>105</td>
<td>64</td>
<td>48</td>
<td>54</td>
<td>51</td>
<td>6</td>
</tr>
<tr>
<td>Capacity</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>42</td>
<td>41</td>
<td>27</td>
<td>51</td>
<td>48</td>
<td>35</td>
<td>38</td>
<td>41</td>
<td>37</td>
<td>15</td>
</tr>
<tr>
<td>First Rejected</td>
<td>1-P</td>
<td>1-R</td>
<td>1-R</td>
<td>1-R</td>
<td>1-R</td>
<td>NR</td>
<td>NR</td>
<td>1-R</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Paid Lunch Students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ranked First</td>
<td>49</td>
<td>45</td>
<td>40</td>
<td>29</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>17</td>
<td>13</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Ranked Anywhere</td>
<td>152</td>
<td>95</td>
<td>128</td>
<td>60</td>
<td>51</td>
<td>87</td>
<td>70</td>
<td>65</td>
<td>33</td>
<td>9</td>
<td>21</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>Capacity</td>
<td>29</td>
<td>27</td>
<td>27</td>
<td>29</td>
<td>41</td>
<td>18</td>
<td>36</td>
<td>34</td>
<td>29</td>
<td>35</td>
<td>34</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>First Rejected</td>
<td>1-P</td>
<td>1-R</td>
<td>1-R</td>
<td>1-R</td>
<td>1-R</td>
<td>NR</td>
<td>NR</td>
<td>3-R</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Free Lunch Students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ranked First</td>
<td>9</td>
<td>12</td>
<td>12</td>
<td>17</td>
<td>12</td>
<td>11</td>
<td>13</td>
<td>10</td>
<td>12</td>
<td>16</td>
<td>10</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Ranked Anywhere</td>
<td>29</td>
<td>24</td>
<td>25</td>
<td>40</td>
<td>20</td>
<td>36</td>
<td>44</td>
<td>38</td>
<td>31</td>
<td>36</td>
<td>34</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Capacity</td>
<td>25</td>
<td>23</td>
<td>26</td>
<td>26</td>
<td>41</td>
<td>17</td>
<td>33</td>
<td>31</td>
<td>19</td>
<td>18</td>
<td>26</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td>First Rejected</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>1-R</td>
<td>1-R</td>
<td>2-P</td>
<td>NR</td>
<td>NR</td>
<td>1-R</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Strategic Behavior
Top Rank: Proximity Boundary
Strategic Behavior
Second Rank: Proximity Boundary
Strategic Behavior
Second and Third Closest Schools

![Graph showing the probability ranked 1st vs distance from the proximity boundary in miles. The x-axis represents the distance from the boundary in miles, ranging from -0.4 to 0.4. The y-axis represents the probability ranked 1st, ranging from 0 to 0.35. The graph includes error bars and a trend line.]
Strategic Behavior
Placebo with Two Closest Schools
Plan

1. Descriptive: Cambridge Public Schools
2. Estimating Preferences: Overview
3. First Step: Assignment Probabilities
4. Second Step: Preferences
5. Estimation and Estimates from Cambridge
Evidence questions the benchmark model of truthful behavior
Evidence questions the benchmark model of truthful behavior

Sound advice for ranking should be based on
1. Chances of getting into a school if ranked and where ranked
2. Considering if a long shot at a competitive school is preferable to an almost sure shot at a less competitive school
Evidence questions the benchmark model of truthful behavior

Sound advice for ranking should be based on
1. Chances of getting into a school if ranked and where ranked
2. Considering if a long shot at a competitive school is preferable to an almost sure shot at a less competitive school

This thought process suggests an estimation strategy
1. Estimate the probability of assignment as a function of reported ranking
2. Estimate preference parameters assuming optimal play
Estimating Preferences: Overview

- Evidence questions the benchmark model of truthful behavior

- Sound advice for ranking should be based on
  1. Chances of getting into a school if ranked and where ranked
  2. Considering if a long shot at a competitive school is preferable to an almost sure shot at a less competitive school

- This thought process suggests an estimation strategy
  1. Estimate the probability of assignment as a function of reported ranking
  2. Estimate preference parameters assuming optimal play
Estimating Preferences: Overview

- Evidence questions the benchmark model of truthful behavior
- Sound advice for ranking should be based on
  1. Chances of getting into a school if ranked and where ranked
  2. Considering if a long shot at a competitive school is preferable to an almost sure shot at a less competitive school
- This thought process suggests an estimation strategy
  1. Estimate the probability of assignment as a function of reported ranking
  2. Estimate preference parameters assuming optimal play
- Need to show that each step is valid given assumptions
Plan

1. Descriptive: Cambridge Public Schools

2. Estimating Preferences: Overview

3. First Step: Assignment Probabilities
   - Mechanisms
   - Large Market Approximations
   - Report-specific Priority + Cutoff Mechanisms

4. Second Step: Preferences

5. Estimation and Estimates from Cambridge
Mechanisms

- Reports, types and lotteries
  - A report $R_i \in \mathcal{R}_i$
  - Students have priority types $t_i \in T$
  - Allow for randomization through lotteries

Economy:
- Economy index by $n$ has students $i \in \{1, \ldots, n\}$
- School $j \in \{0, 1, \ldots, J\} = S$ has capacity $q_{nj}$, with school 0 denotes unmatched
- $m_n - 1$ is the measure of reports and priority types of other agents
- Probability of assignment is a vector-valued function $\phi_n((R_i, t_i), m_n - 1)$ that depends on $i$'s report, priority type and report + priority type of other students.
Mechanisms

- **Reports, types and lotteries**
  - A report \( R_i \in \mathcal{R}_i \)
  - Students have priority types \( t_i \in T \)
  - Allow for randomization through lotteries

- **Economy:**
  - Economy index by \( n \) has students \( i \in \{1, \ldots, n\} \)
  - School \( j \in \{0, 1, \ldots, J\} = S \) has capacity \( q_j^n \), with school 0 denotes unmatched
  - \( m^{n-1} \) is the measure of reports and priority types of other agents
Mechanisms

- Reports, types and lotteries
  - A report $R_i \in \mathcal{R}_i$
  - Students have priority types $t_i \in T$
  - Allow for randomization through lotteries

- Economy:
  - Economy index by $n$ has students $i \in \{1, \ldots, n\}$
  - School $j \in \{0, 1, \ldots, J\} = S$ has capacity $q_j^n$, with school 0 denotes unmatched
  - $m^{n-1}$ is the measure of reports and priority types of other agents

- Probability of assignment is a vector-valued function

$$\phi^n((R_i, t_i), m^{n-1})$$

that depends on $i$’s report, priority type and report + priority type of other students.
Large Market Asymptotics

- Typical datasets have many students but few schools and markets
  - Motivates a large market asymptotics with $n \to \infty$
    - Fixed number of schools, growing number of students and capacity
Typical datasets have many students but few schools and markets

- Motivates a large market asymptotics with $n \to \infty$
  - Fixed number of schools, growing number of students and capacity

Equilibrium depends on $n$, complicating analysis

- Assume that agents play equilibrium strategies of the limit game given by

$$\phi^\infty((R_i, t_i), m) = \lim_{n \to \infty} \phi^n((R_i, t_i), m)$$
Typical datasets have many students but few schools and markets

- Motivates a large market asymptotics with $n \to \infty$
  - Fixed number of schools, growing number of students and capacity

Equilibrium depends on $n$, complicating analysis

- Assume that agents play equilibrium strategies of the limit game given by

$$
\phi^\infty((R_i, t_i), m) = \lim_{n \to \infty} \phi^n((R_i, t_i), m)
$$

Justification based on a condition on the mechanism such that

- Equilibria of the limit game are $\epsilon$-equilibria of the finite game
- $\phi^\infty((R_i, t_i), m)$ are identified and can be consistently estimated
Typical datasets have many students but few schools and markets
  ✓ Motivates a large market asymptotics with $n \to \infty$
    ▶ Fixed number of schools, growing number of students and capacity

Equilibrium depends on $n$, complicating analysis
  ✓ Assume that agents play equilibrium strategies of the limit game given by

$$
\phi^\infty((R_i, t_i), m) = \lim_{n \to \infty} \phi^n((R_i, t_i), m)
$$

Justification based on a condition on the mechanism such that
  ✓ Equilibria of the limit game are $\epsilon$-equilibria of the finite game
  ✓ $\phi^\infty((R_i, t_i), m)$ are identified and can be consistently estimated
Large Market Approximation

Condition (Convergence of the mechanism)

Suppose the sequence of empirical measures $m^{n-1}$ on $\mathcal{R} \times T$ converges in probability to the population measure $m \in \mathcal{M}$, then for each $(R, t)$,
$$\phi^n((R, t), m^{n-1}) \xrightarrow{P} \phi^\infty((R, t), m).$$

Two takeaways:
1. Consistency: Justifies the use of data to estimate these probabilities
2. Justifies equilibrium assumption: limit equilibria are $\varepsilon$-equilibria of finite games

We will verify the Convergence condition for a large class of mechanisms that include Boston, Student Proposing DA, among others.
Large Market Approximation

Condition (Convergence of the mechanism)

Suppose the sequence of empirical measures $m^{n-1}$ on $\mathcal{R} \times T$ converges in probability to the population measure $m \in \mathcal{M}$, then for each $(R, t)$, $\phi^n((R, t), m^{n-1}) \xrightarrow{p} \phi^\infty((R, t), m)$.

Two takeaways:

1. Consistency: Justifies the use of data to estimate these probabilities
Large Market Approximation

Condition (Convergence of the mechanism)

Suppose the sequence of empirical measures $m^{n-1}$ on $\mathcal{R} \times T$ converges in probability to the population measure $m \in \mathcal{M}$, then for each $(R, t)$, $\phi^n((R, t), m^{n-1}) \xrightarrow{P} \phi^\infty((R, t), m)$.

Two takeaways:

1. Consistency: Justifies the use of data to estimate these probabilities
2. Justifies equilibrium assumption: limit equilibria are $\varepsilon$-equilibria of finite games
Large Market Approximation

Condition (Convergence of the mechanism)

Suppose the sequence of empirical measures $m^{n-1}$ on $R \times T$ converges in probability to the population measure $m \in \mathcal{M}$, then for each $(R, t)$, $\phi^n((R, t), m^{n-1}) \overset{P}{\to} \phi^\infty((R, t), m)$.

Two takeaways:

1. Consistency: Justifies the use of data to estimate these probabilities
2. Justifies equilibrium assumption: limit equilibria are $\varepsilon$-equilibria of finite games

We will verify the Convergence condition for a large class of mechanisms that include Boston, Student Proposing DA, among others.
Mechanisms are typically described in terms of algorithms. We introduce RSP+C mechanisms, which can be represented in terms of:

1. A function $f(R_i, t_i)$ uses a report to modify the priority type of the agent
2. A lottery $e_i$ for each student
3. Cutoffs $p_j$ for each school

Such that each student is assigned to her top ranked choice for which

$$f_j(R_i, t_i) + e_{ij} > p_j$$
Mechanisms are typically described in terms of algorithms.

We introduce RSP+C mechanisms, which can be represented in terms of:

1. A function $f(R_i, t_i)$ uses a report to modify the priority type of the agent.
2. A lottery $e_i$ for each student.
3. Cutoffs $p_j$ for each school.

Such that each student is assigned to her top ranked choice for which:

$$f_j(R_i, t_i) + e_{ij} > p_j$$

Excess demand (assignment) for school $j$:

$$D_j(p|m^{n-1}) - q_j$$
Report-specific Priority + Cutoff Mechanisms

- Mechanisms are typically described in terms of algorithms
- We introduce RSP+C mechanisms, which can be represented in terms of:
  1. A function $f(R_i, t_i)$ uses a report to modify the priority type of the agent
  2. A lottery $e_i$ for each student
  3. Cutoffs $p_j$ for each school

Such that each student is assigned to her top ranked choice for which

$$f_j(R_i, t_i) + e_{ij} > p_j$$

- Excess demand (assignment) for school $j$:

$$D_j(p|m^{n-1}) - q_j$$

- Market-clearing cutoff: $p^n$
Report-specific Priority + Cutoff Mechanisms

- Mechanisms are typically described in terms of algorithms.
- We introduce RSP+C mechanisms, which can be represented in terms of:
  1. A function $f(R_i, t_i)$ uses a report to modify the priority type of the agent.
  2. A lottery $e_i$ for each student.
  3. Cutoffs $p_j$ for each school.

Such that each student is assigned to her top ranked choice for which:

$$f_j(R_i, t_i) + e_{ij} > p_j$$

- Excess demand (assignment) for school $j$:

$$D_j(p|m^{n-1}) - q_j$$

- Market-clearing cutoff: $p^n$

**Theorem**

*Under smoothness conditions, if $\phi^n$ is a report-specific priority + cutoff mechanism, then $\phi^n$ satisfies Condition 1.*
Plan

1. Descriptive: Cambridge Public Schools

2. Estimating Preferences: Overview

3. First Step: Assignment Probabilities

4. Second Step: Preferences
   - Preference Model
   - Identification under Varying Choice Environments
   - Identification with a Special Regressor

5. Estimation and Estimates from Cambridge
Preference Model

- **Alphabet soup - similar to Berry and Haile, 2010**
  - Student $i$’s observables in market $b$, $z_{ib} = (z_{i1b}, \ldots, z_{ijb})$
  - School unobservables is a vector $\xi_b$
Preference Model

- Alphabet soup - similar to Berry and Haile, 2010
  - Student $i$'s observables in market $b$, $z_{ib} = (z_{i1b}, \ldots, z_{iJb})$
  - School unobservables is a vector $\xi_b$

- RUM: Want to identify the distribution $F_V(v_{i1}, \ldots, v_{iJ_b}|\xi_b, z_{ib})$ of
  \[
  v = (v_{i1}, \ldots, v_{iJ_b}), \text{ with } v_{i0} = 0
  \]

- ✓ Condition on market: Drop $\xi_b$.
- ✓ $F_V(v_{i1}, \ldots, v_{iJ_b}|z_{ib})$ admits a density $f_V(v_{i1}, \ldots, v_{iJ_b}|z_{ib})$
Choice Over Lotteries

We will talk about a choice over lotteries:

\[ \mathcal{L}_T = \{ L_R \in \Delta S : L_R = \phi^\infty((R, t), m_{\sigma^*}) \} \]

✓ In a limit BNE, agents know the probabilities associated with each \( R \)
Agent with utility $v_i$ choose lottery $L_R$ only if $v_i \cdot (L_R - L) \geq 0$ for all $L \in \mathcal{L}_r$.

✓ Report $R$ if utility is in the normal cone:

$$N_{\mathcal{L}_r}(L_R) = \{v \in \mathbb{R}^J : \forall L \in \mathcal{L}_r, v \cdot (L_R - L) \geq 0\}$$
Utility can only be known up to affine transformations

✓ Without loss of generality, normalize $\|v_i\| = 1$
Utility can only be known up to affine transformations
✓ Without loss of generality, normalize $\|v_i\| = 1$

Can identify mass accumulated in arcs, one for each rank-order list!
✓ $P(L \in \mathcal{L}_\Gamma | z) = \int 1\{ v \in N_{\mathcal{L}_\Gamma (L)} \} \text{d}F_{V|z}$
Identification with a Special Regressor

Another approach is based on a special regressor à la Lewbel (2000)

\[ \nu(z_{ijb}, \epsilon_i) = \nu(z_{ijb}^2, \epsilon_i) - z_{ijb}^1 \]

where \( \epsilon_i \perp z_{ijb}^1 \)

✓ Embeds scale normalization → drop \( \|v_i\| = 1 \)

✓ Distance-metric utility (Abdulkadiroglu, Agarwal and Pathak, 2014)
Identification with a Special Regressor

- Another approach is based on a special regressor á la Lewbel (2000)

\[ \nu(z_{ijb}, \epsilon_i) = \nu(z_{ijb}^2, \epsilon_i) - z_{ijb}^1 \]

where \( \epsilon_i \perp z_{ijb}^1 \)

- Embeds scale normalization \( \rightarrow \) drop \( \|\nu_i\| = 1 \)
- Distance-metric utility (Abdulkadiroglu, Agarwal and Pathak, 2014)

Consider markets conditioning on \( z^2 \) and drop it from notation

For a cone \( C \), the probability that \( \nu \in C \) given \( z^1 \) is

\[ h_C(z^1) = \int 1\{\nu \in C\} dF_V(\nu|z^1) \]

\[ = \int 1\{\nu - z^1 \in C\} g(\nu) d\nu \]
Identification: Two Dimensional Case

✓ Local variation is sufficient for local identification
Identification: Two Dimensional Case

✓ Local variation is sufficient for local identification
Plan

1. Descriptive: Cambridge Public Schools
2. Estimating Preferences: Overview
3. First Step: Assignment Probabilities
4. Second Step: Preferences
5. Estimation and Estimates from Cambridge
   - First Step
   - Second Step
Two-Step Estimation

- First step estimates $\phi^\infty((R, t), m)$: Simulate $\hat{\phi}(R, t) = \phi^n((R, t), \hat{m}^{n-1})$
  where $\hat{m}^{n-1}$ is either observed data or resampled from the data

- Cambridge mechanism satisfies Condition 1
92% of students have reports consistent with optimal play

Significant heterogeneity in prob of assignment, across schools and students
Second step can be implemented using Gibbs’ Sampling for a model

\[ v_{ij} = x_{ij} \beta_j - d_{ij} + \varepsilon_{ij} \]

where \( \varepsilon_{ij} \sim N(0, \Sigma) \), \( \beta \sim N(\bar{\beta}, \Sigma_{\beta}) \) and \( \Sigma \sim IW(\nu_0, S) \)

✓ \( x_{ij} \) are interactions between school dummies and paid lunch, ethnicity, home language dummies
Gibbs’ Sampler

- Second step can be implemented using Gibbs’ Sampling for a model

\[ v_{ij} = x_{ij} \beta_j - d_{ij} + \varepsilon_{ij} \]

where \( \varepsilon_{ij} \sim N(0, \Sigma) \), \( \beta \sim N(\bar{\beta}, \Sigma_\beta) \) and \( \Sigma \sim IW(\nu_0, S) \)

✓ \( x_{ij} \) are interactions between school dummies and paid lunch, ethnicity, home language dummies

- Agent \( i \) picks lottery \( L_i \); find associated cone: \( C_i = N_{L}(L_i) \)

- Data augmentation: start with a guess \( v_i \in C_i \)
Second step can be implemented using Gibbs’ Sampling for a model

\[ v_{ij} = x_{ij} \beta_j - d_{ij} + \varepsilon_{ij} \]

where \( \varepsilon_{ij} \sim N(0, \Sigma) \), \( \beta \sim N(\bar{\beta}, \Sigma_\beta) \) and \( \Sigma \sim IW(\nu_0, S) \)

✓ \( x_{ij} \) are interactions between school dummies and paid lunch, ethnicity, home language dummies

Agent \( i \) picks lottery \( L_i \); find associated cone: \( C_i = N_L(L_i) \)

Data augmentation: start with a guess \( v_i \in C_i \)

Gibbs Sampler:
1. Draw \( \beta^{k+1} \) and \( \Sigma^{k+1} \) given \( v \).

\[ v_{ij}^{k+1} | v_{i-j}^k, \beta_j^k, \Sigma^k, C_i \]
## Ranking and Assignment to True First Choice

<table>
<thead>
<tr>
<th></th>
<th>Panel A: All Students</th>
<th>Panel B: Free Lunch Students</th>
<th>Panel C: Paid Lunch Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Graham Parks</td>
<td>Haggerty</td>
<td>Baldwin</td>
</tr>
<tr>
<td>Preferred School</td>
<td>20.6</td>
<td>11.2</td>
<td>9.2</td>
</tr>
<tr>
<td>Ranked #1 (simul)</td>
<td>15.4</td>
<td>12.1</td>
<td>10.7</td>
</tr>
<tr>
<td>Ranked #1 (data)</td>
<td>14.3</td>
<td>12.6</td>
<td>11.9</td>
</tr>
<tr>
<td>Preferred and Ranked #1</td>
<td>13.9</td>
<td>10.1</td>
<td>8.3</td>
</tr>
<tr>
<td>Preferred and Assigned</td>
<td>9.4</td>
<td>8.3</td>
<td>6.6</td>
</tr>
<tr>
<td>Ranked #1 and Assigned</td>
<td>10.0</td>
<td>9.7</td>
<td>8.3</td>
</tr>
</tbody>
</table>

- **Stated ranks may overestimate assignment to true first choice**
  - ✓ Quantitative difference depends on student group
- **Stated ranks likely misestimate school desirability**
  - ✓ Bias isn’t obvious, depends on local competitiveness
### Ranking and Assignment to True First Choice

<table>
<thead>
<tr>
<th></th>
<th>Truthful All Students</th>
<th>Paid Lunch</th>
<th>Free Lunch</th>
<th>Sophisticated All Students</th>
<th>Paid Lunch</th>
<th>Free Lunch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assigned to First Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>66.74</td>
<td>56.95</td>
<td>86.12</td>
<td>70.00</td>
<td>60.53</td>
<td>88.74</td>
</tr>
<tr>
<td><strong>Assigned to Second Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.93</td>
<td>15.20</td>
<td>8.44</td>
<td>14.82</td>
<td>17.78</td>
<td>8.96</td>
</tr>
<tr>
<td><strong>Assigned to Third Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.95</td>
<td>8.52</td>
<td>0.85</td>
<td>4.36</td>
<td>6.15</td>
<td>0.81</td>
</tr>
<tr>
<td><strong>Assigned to Fourth Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.58</td>
<td>5.33</td>
<td>0.11</td>
<td>1.07</td>
<td>1.57</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Assigned to Fifth Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.15</td>
<td>3.23</td>
<td>0.02</td>
<td>0.16</td>
<td>0.24</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Panel A: Deferred Acceptance**

|                      |                       |            |            |                             |            |            |
| **Assigned to First Choice** |                       |            |            |                             |            |            |
|                      | 78.09                 | 73.65      | 86.88      | 74.38                       | 67.44      | 88.13      |
| **Assigned to Second Choice** |                       |            |            |                             |            |            |
|                      | 6.97                  | 7.17       | 6.58       | 13.26                       | 15.77      | 8.27       |
| **Assigned to Third Choice** |                       |            |            |                             |            |            |
|                      | 3.21                  | 4.09       | 1.48       | 3.24                        | 4.15       | 1.46       |
| **Assigned to Fourth Choice** |                       |            |            |                             |            |            |
|                      | 0                     | 0          | 0          | 0.84                        | 1.1        | 0.32       |
| **Assigned to Fifth Choice** |                       |            |            |                             |            |            |
|                      | 0                     | 0          | 0          | 0.13                        | 0.18       | 0.05       |

**Panel B: Cambridge Mechanism**

|                      |                       |            |            |                             |            |            |
| **Assigned to First Choice** |                       |            |            |                             |            |            |
|                      | Mean Utility DA - Cambridge | -0.02 | -0.04 | 0.01 | -0.10 | -0.15 | 0.00 |
| Std. Utility DA - Cambridge |                       |            |            |                             |            |            |
| Percent DA > Cambridge |                       |            |            |                             |            |            |
|                      | 51.49                 | 51.38      | 51.71      | 50.5                        | 48.92      | 53.63      |

**Panel C: Deferred Acceptance vs Cambridge**

Notes: Panels A and B present percentages of students assigned to true k-th choice. Panel C compares the expected utility difference between Deferred Acceptance and Cambridge Mechanism. Simulations of the Deferred Acceptance mechanism draw other student reports using the estimated utility distribution. All simulations based on the posterior means of the parameters and 1,000 draws.
Conclusion

- One can think of reporting a rank-order list as a choice over lotteries
  - ✔ Potentially generalizable to other settings
Conclusion

- One can think of reporting a rank-order list as a choice over lotteries
  ✓ Potentially generalizable to other settings

- We can identify and estimate preferences in a class of manipulable mechanism
  ✓ Gibbs’ sampler is computationally tractable

- Application to Cambridge public elementary schools
  ✓ Direct evidence of strategic incentives and strategic behavior

- Relative to the algorithm in place, Deferred Acceptance would reduce an aggregate measure of welfare for paid-lunch students (0.15 miles per student)
Conclusion

- One can think of reporting a rank-order list as a choice over lotteries
  - Potentially generalizable to other settings

- We can identify and estimate preferences in a class of manipulable mechanism
  - Gibbs’ sampler is computationally tractable

- Main restriction is on agent sophistication
  - Alternative to the current approach of treating lists as truthful
Conclusion

- One can think of reporting a rank-order list as a choice over lotteries
  - Potentially generalizable to other settings

- We can identify and estimate preferences in a class of manipulable mechanism
  - Gibbs’ sampler is computationally tractable

- Main restriction is on agent sophistication
  - Alternative to the current approach of treating lists as truthful

- Application to Cambridge public elementary schools
  - Direct evidence of strategic incentives and strategic behavior
  - Treating reports as truthful may overestimate fraction assigned to top choice
  - Paid-lunch face more competition and a more complex strategic environment
  - Relative to the algorithm in place, Deferred Acceptance would reduce an aggregate measure of welfare for paid-lunch students (0.15 miles per student)
Thank You!
Definition (Mechanism)

Fix a set of schools $S$ and a sequence of capacities $q_j^n$. A (feasible) **mechanism** $\{\Phi^n\}$ is a sequence of functions

$$\Phi^n : (\mathcal{R} \times T)^n \to (\Delta S)^n$$

such that for all $(R \times T) \in \mathcal{R}^n$,

$$\sum_{i=1}^{n} \Phi^n_{ij}(R, t) \leq q_j^n.$$
Semi-Anonymous Mechanisms

Definition

A mechanism is **semi-anonymous** if

1. For all \((R_i, t_i), R_{\neg i}, t_{\neg i}\) and \(i, i'\), we have
   \[ \Phi^n_i((R_i, t_i), R_{\neg i}, t_{\neg i}) = \Phi^n_{i'}((R_i, t_i), R_{\neg i}, t_{\neg i}). \]

2. For all \(R_i, t_i, R_{\neg i}, t_{\neg i}\), and all permutations \(\pi\) of \(-i = \{1, \ldots, i-1, i+1, \ldots, n\}\), we have that
   \[ \Phi^n_i((R_i, t_i), R_{\neg i}, t_{\neg i}) = \Phi^n_i((R_i, t_i), R_{\pi(-i)}, t_{\pi(-i)}). \]

- Relevant for considering priorities such as walk-zone, sibling etc
  - ✓ Restriction to finite \(T\) rules out fine metrics such as test scores
Semi-Anonymous Mechanisms

Definition

A mechanism is semi-anonymous if

1. For all \((R_i, t_i), R_{-i}, t_{-i}\) and \(i, i'\), we have
\[
\Phi_i^n((R_i, t_i), R_{-i}, t_{-i}) = \Phi_{i'}^n((R_i, t_i), R_{-i}, t_{-i}).
\]

2. For all \(R_i, t_i, R_{-i}, t_{-i}\), and all permutations \(\pi\) of \(-i = \{1, \ldots, i-1, i+1, \ldots, n\}\), we have that
\[
\Phi_i^n((R_i, t_i), R_{-i}, t_{-i}) = \Phi_i^n((R_i, t_i), R_{\pi(-i)}, t_{\pi(-i)}).
\]

- Relevant for considering priorities such as walk-zone, sibling etc
  
  ✓ Restriction to finite \(T\) rules out fine metrics such as test scores

- Enough to keep track of

\[
m(R_{-i}, t_{-i}) = \frac{1}{n-1} \sum_{i' \in -i} \delta_{R_{i'}, t_{i'}}
\]

and rewrite

\[
\phi^n((R_i, t_i), m(R_{-i}, t_{-i})) = \Phi_i^n((R_i, t_i), R_{-i}, t_{-i}).
\]
Limit Mechanism

- We will use the large market limit to discuss mechanisms á la Azevedo and Budish (2012)

**Definition**

The function $\phi^\infty : \mathcal{R} \times T \times \Delta(\mathcal{R} \times T) \rightarrow \Delta S$ is a **limit mechanism** of the sequence of anonymous mechanisms $\{\phi^n\}$ if for all $(R_i, t_i)$ and $m \in \Delta(\mathcal{R} \times T)$,

$$\phi^\infty(R_i, t_i, m) = \lim_{n \rightarrow \infty} \phi^n(R_i, m).$$

- Implicitly, reports of any single agent do not affect outcomes for other agents
- ✓ Mechanisms need not have a well-defined limit

- Azevedo and Budish (2012) show that some mechanisms are not strategy-proof in the large market e.g. Boston Mechanism
Report-specific Priority + Cutoff Mechanisms

- Student priority scores and individual demand
  - Let $e_i \in [0, 1]^J$ denotes the vector to student priority score including the lottery and the priority type $t_i$
  - At cutoff vector $p_j \in [0, 1]^J$, report $R_i$ expresses individual demand for $j$ if
    
    $$D_j^{(R_i, e_i)} = 1\{e_{ij} \geq p_j, jR_i\} \prod_{j' \neq j} 1\{jR_i j' \text{ or } e_{j'} < p_{j'}\}.$$
Report-specific Priority + Cutoff Mechanisms

- Student priority scores and individual demand
  - Let $e_i \in [0, 1]^J$ denotes the vector to student priority score including the lottery and the priority type $t_i$
  - At cutoff vector $p_j \in [0, 1]^J$, report $R_i$ expresses individual demand for $j$ if

$$D_j^{(R_i,e_i)} = 1\{e_{ij} \geq p_j, jR_i\} \prod_{j' \neq j} 1\{jR_{ij'} \text{ or } e_{ij'} < p_{j'}\}.$$  

- Given measure $\eta$ on $R \times [0, 1]^J$, aggregate demand is

**Definition (Aggregate Demand)**

For a cutoff vector $p \in [0, 1]^J$, the measure of students demanding $j \in S$ is given by

$$D_j(p) = \eta \left( \{(e_i, R_i) : e_{ij} \geq p_j, jR_i\} \bigcap \{(e_i, R_i) : jR_{ij'}\} \cup \{(e_i, R_i) : e_{ij'} < p_{j'}\} \right).$$
Report Specific Cutoff + Lottery Mechanism

Definition (Market Clearing Cutoff)

The vector of cutoffs $p$ is a market clearing cutoff for economy $(\eta, q)$ if for all $j \in S$, $D_j(p|\eta) \leq q_j$, with equality if $p_j > 0$. 

The Boston Mechanism, Deferred Acceptance, Pan London, Chinese Parallel and First Priority First are all Report-Specific Cutoff + Lottery Mechanisms.

$\frac{37}{31}$
Report Specific Cutoff+Lottery Mechanism

Definition (Market Clearing Cutoff)

The vector of cutoffs $p$ is a market clearing cutoff for economy $(\eta, q)$ if for all $j \in S$, $D_j(p|\eta) \leq q_j$, with equality if $p_j > 0$.

Definition (Report Specific Cutoff+Lottery Mechanism)

A mechanism $\phi^n$ is a report-specific priorities + cutoff mechanism if there exists a function $f : \mathcal{R} \times [0, 1]^J \rightarrow [0, 1]^J$ such that

(i) $f$ strictly increasing in the last $J$ arguments

(ii) $\phi^n_j((R_i, t_i), m(R_{-i}, t_{-i})) = \int \ldots \int D^{(R_i, f(R_i, e_i))}(p^n) d\eta_{R, e|t}$

(iii) $p^n$ are market clearing cutoffs for each profile of reports and lotteries
Report Specific Cutoff+Lottery Mechanism

Definition (Market Clearing Cutoff)

The vector of cutoffs $p$ is a market clearing cutoff for economy $(\eta, q)$ if for all $j \in S$, $D_j(p|\eta) \leq q_j$, with equality if $p_j > 0$.

Definition (Report Specific Cutoff+Lottery Mechanism)

A mechanism $\phi^n$ is a report-specific priorities + cutoff mechanism a there exists a function $f : \mathcal{R} \times [0, 1]^J \rightarrow [0, 1]^J$ such that

(i) $f$ strictly increasing in the last $J$ arguments

(ii) $\phi^n_j((R_i, t_i), m(R_{-i}, t_{-i})) = \int \ldots \int D^{(R_i, f(R_i, e_i))}(p^n)d\eta_{R,e|t}$

(iii) $p^n$ are market clearing cutoffs for each profile of reports and lotteries

✓ The Boston Mechanism, Deferred Acceptance, Pan London, Chinese Parallel and First Priority First are all Report-Specific Cutoff+Lottery Mechanisms
In the continuum economy, we make the following assumptions on $\eta$:

**Assumption (Non-degenerate lotteries)**

For each $p, p' \in [0, 1]^J$, $R$ and $j$,

\[
\eta(\{(e, R) : p_j \wedge p'_j \leq e_j \leq p_j \vee p'_j\}) \leq \kappa |p_j - p'_j|.
\]

**Assumption (Unique Cutoff)**

$(\eta, q)$ admits a unique market clearing cutoff $p^*$. 

Holds generically in $q$; Paper provides testable sufficient conditions on $\eta$.

Closely related to Azevedo (2011), and Azevedo and Leshno (2013).
In the continuum economy, we make the following assumptions on $\eta$:

**Assumption (Non-degenerate lotteries)**

For each $p, p' \in [0, 1]^J$, $R$ and $j$,

$$\eta(\{ (e, R) : p_j \land p'_j \leq e_j \leq p_j \lor p'_j \}) \leq \kappa |p_j - p'_j|.$$

**Assumption (Unique Cutoff)**

$(\eta, q)$ admits a unique market clearing cutoff $p^*$. 

- Non-degenerate lotteries can be verified with knowledge of tie-breaking rules
- Uniqueness of cutoffs rules out knife-edge cases
  - Holds generically in $q$
  - Paper provides testable sufficient conditions on $\eta$

✓ Closely related to Azevedo (2011), and Azevedo and Leshno (2013)
A (mixed) **strategy** is a function $\sigma_i : (\mathbb{R}^J \times T) \to \Delta \mathcal{R}_i$.

- Let $\sigma_i(u_i, t_i; R)$ be the weight on the profile $R$
- ✓ We will only consider symmetric equilibria

The ex-ante payoff from report $R_i$ in game $\phi^n$ is given by

$$V^n_i((R_i, t_i), m_\sigma) = \mathbb{E} \left[ \sum_j \phi^n_{ij}((R_i, t_i), m^{n-1}_\sigma)u_{ij} \right]. \quad (1)$$

where $m^{n-1}_\sigma$ is an empirical measure of $n - 1$ iid draws from $m_\sigma$, with $m_\sigma(R, t) = f_T(t) \times \int \sigma(u, t; R)dF_{U|T}$
Limit Equilibrium

Definition
The strategy \( \sigma^* \) is a limit equilibrium if for all \( i \), \( \sigma^*(u_i, t_i; R_i) > 0 \) implies that

\[
R_i \in \arg \max_{R'_i \in \mathcal{R}_i} \sum_j \phi_{ij}^\infty ((R_i, t_i), m_{\sigma}) u_{ij}.
\]

✓ We will focus on symmetric equilibria

Definition
The strategy \( \sigma^* \) is an \( \varepsilon \)-equilibrium of the \( n \)-th game if \( \sigma^*(u_i, t_i; R_i) > 0 \) implies

\[
V_i^n(R'_i, m_{\sigma^*}) - V_i^n(R_i, m_{\sigma^*}) \leq \varepsilon \|u_i\|_\infty.
\]
Proposition

Let $\sigma^*$ be a symmetric equilibrium of $\phi^\infty$. If Condition 1 holds, then there exists an $n_0$ such that for all $n > n_0$, $\sigma^*$ is an $\varepsilon$-equilibrium of $\phi^n$. 

Proof Sketch.

1. IID sampling from $m = \int \sigma^* dF_U$, $T$ implies that $m_n - 1_p \to m$.

2. Condition 1 implies that $\phi_n((R_i, t_i), m_n - 1_p) \to \phi_n((R_i, t_i), m)$.

3. By the continuous mapping theorem, $\|u_i\|_\infty |V_n_i(R_i', \sigma^*) - V_\infty_i(R_i', \sigma^*)| \to 0$.

4. For $\varepsilon > 0$, find $n_0$ large enough so that the gain from deviating from $\sigma^*$ is small with large enough probability.
Proposition

Let $\sigma^*$ be a symmetric equilibrium of $\phi^\infty$. If Condition 1 holds, then there exists an $n_0$ such that for all $n > n_0$, $\sigma^*$ is an $\varepsilon$-equilibrium of $\phi^n$.

Proof Sketch.
1. IID sampling from $m = \int \sigma^* \text{d}F_{U,T}$ implies that $m^{n-1} \xrightarrow{p} m$
2. Condition 1 implies that $\phi^n((R_i, t_i), m^{n-1}) \xrightarrow{p} \phi^n((R_i, t_i), m)$
3. By the continuous mapping theorem, $\frac{1}{\|u_i\|_\infty} |V_i^n(R'_i, \sigma^*) - V_i^\infty(R'_i, \sigma^*)| \xrightarrow{p} 0$
4. For $\varepsilon > 0$, find $n_0$ large enough so that the gain from deviating from $\sigma^*$ is small with large enough probability
Identification with Lottery Variation

- May observe variation in choice probabilities excludable from preferences
  - Variation in priorities $t$, mechanisms $\phi$ or play $m$
- Let $\mathcal{T}$ is a collection of markets with fixed $z, \xi$ and define
  \[
  \mathcal{N} = \{\text{int}(N_{\mathcal{L}_\Gamma}(L))\}_{\Gamma \in \mathcal{T}, L \in \mathcal{L}_\Gamma}
  \]
Identification with Lottery Variation

- May observe variation in choice probabilities excludable from preferences
  - Variation in priorities \(t\), mechanisms \(\phi\) or play \(m\)

- Let \(\mathcal{T}\) is a collection of markets with fixed \(z, \xi\) and define
  \[
  \mathcal{N} = \{\text{int}(N_{L\Gamma}(L))\}_{\Gamma \in \mathcal{T}, L \in \mathcal{L}_\Gamma}
  \]

- Let \(\mathcal{D}(\mathcal{N})\) be the smallest collection of subsets of \(\mathbb{R}^J\) such that
  1. \(\mathbb{R}^J, \mathcal{N} \in \mathcal{D}(\mathcal{N})\)
  2. For all \(N \in \mathcal{D}(\mathcal{N}), N^c \in \mathcal{D}(\mathcal{N})\)
  3. For all countable sequences of sets \(N_k \in \mathcal{D}(\mathcal{N})\) such that \(N_{k_1} \cap N_{k_2} = \emptyset\),
     \(\bigcup_k N_k \in \mathcal{D}(\mathcal{N})\)

**Lemma**

Given \(P(L \in \mathcal{L}_\Gamma | \Gamma)\) for each \(\Gamma \in \mathcal{T}\) and \(L \in \mathcal{L}_\Gamma\), the quantity

\[
h_D = \int 1\{v \in D\}dF_V(v)
\]

is identified for each \(D \in \mathcal{D}(\mathcal{N})\).
A convex cone in $C \subseteq \mathbb{R}^J$ is Simplicial if it is spanned by $J$ linearly independent vectors $w_1, \ldots, w_J$. 

**Theorem**

Let $C$ be a convex simplicial cone. If $h_C(z_1)$ is known on an open set containing $z_1$, then $g(z_1)$ is identified.

$\mathbf{\check{C}} = N_L \Gamma(L)$ for some $\Gamma$, $L$ or $C \in D(N_L)$ are particular cases

**Proof Sketch.**

$\nabla$ Let $W = (w'_1, \ldots, w'_J)$ and wlog $|W| = 1$. Evaluating $h_C$ at $Wx$

$$h_C(Wx) = \int_{\{\epsilon - Wx \in C\}} g(\epsilon) d\epsilon = \int_{x_1 - \infty}^{\infty} \ldots \int_{x_J - \infty}^{\infty} g(Wa) da$$

where we changed variables via $\epsilon = Wa$

$\nabla$ Since $W$ is full-rank, the partial derivatives of $h_L(Wx)$ wrt $x_1, \ldots, x_J$ identify $g$ at $z_1 = Wx$
Identification: Simplicial Cones

- A convex cone in $C \subseteq \mathbb{R}^J$ is **Simplicial** if it is spanned by $J$ linearly independent vectors $w_1, \ldots, w_J$

**Theorem**

*Let $C$ be a convex simplicial cone. If $h_C(z^1)$ is known on an open set containing $z^1$, then $g(z^1)$ is identified.*

✓ $C = N_{L\Gamma}(L)$ for some $\Gamma, L$ or $C \in \mathcal{D}(\mathcal{N})$ are particular cases
Identification: Simplicial Cones

A convex cone in $C \subseteq \mathbb{R}^J$ is **Simplicial** if it is spanned by $J$ linearly independent vectors $w_1, \ldots, w_J$.

**Theorem**

Let $C$ be a convex simplicial cone. If $h_C(z^1)$ is known on an open set containing $z^1$, then $g(z^1)$ is identified.

✓ $C = N_{\mathcal{L}_\Gamma}(L)$ for some $\Gamma, L$ or $C \in \mathcal{D}(\mathcal{N})$ are particular cases.

**Proof Sketch.**

- Let $W = (w'_1, \ldots, w'_J)'$ and wlog $|W| = 1$. Evaluating $h_C$ at $Wx$

  $$h_C(Wx) = \int 1\{\varepsilon - Wx \in C\} g(\varepsilon) d\varepsilon$$

  $$= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_J} g(Wa) da$$

  where we changed variables via $\varepsilon = Wa$

- Since $W$ is full-rank, the partial derivatives of $h_L(Wx)$ wrt $x_1, \ldots, x_J$ identify $g$ at $z^1 = Wx$.
Identification with Non-Simplicial Cones

Proof Sketch.

1. Define the linear operator $A$:

$$A \circ g(z) = \int_{N_L(L)} g(v + z) dv.$$ 

and suppose $g', g'' \in G$ such that $Ag' = Ag''$ but $g = g' - g'' \neq 0$.

2. Since $N_L(L)$ is salient, its dual $T_L(L)$ has a nonempty interior. Let $\varepsilon \in \text{int}(T_L(L))$, with $|\varepsilon|$ sufficiently small so that $g_{\varepsilon}(z) = g(z)e^{2\pi \langle \varepsilon, z \rangle} \in L^1$.

3. Note that the fourier transform $\hat{f}_{\varepsilon, N_L(L)}(\xi)$ of $1(v \in N_L(L))e^{-2\pi \langle \varepsilon, v \rangle}$ is an entire function.

4. Since $e^{-2\pi \langle \varepsilon, z \rangle} > 0$, we have that $\hat{f}_{\varepsilon, N_L(L)}(\xi) \cdot \overline{\hat{g}_{\varepsilon}(\xi)} = 0 \iff \forall z$

$$A \circ g(z) = e^{-2\pi \langle \varepsilon, z \rangle} \int 1(v \in N_L(L))e^{-2\pi \langle \varepsilon, v \rangle}e^{2\pi \langle \varepsilon, v + z \rangle} g(v + z) dv = 0.$$ 

5. Since $\hat{g}_{\varepsilon}$ is continuous, $(\hat{g}_{\varepsilon}^{-1}(0))^c$ is open. Since $g \neq 0$, the support of $\hat{g}_{\varepsilon}$ is non-empty. Hence, there is an open set $Z_{\varepsilon}$ where $\hat{g}_{\varepsilon}$ is different from zero, a contradiction.