Safety First: Perceived Risk of Street Harassment and Educational Choices of Women

Girija Borker

World Bank

24 July 2020
Educational Choices of Women

- Women in India choose to attend worse quality colleges than men
  - in absolute terms
Educational Choices of Women

- Women in India attend worse quality colleges than men
  - within their choice set
Impact of Educational Choices of Women

- Such sorting affects women’s:
  - academic training (Zhang 2005)
  - network of peers (Winston and Zimmerman 2004)
  - access to labor opportunities (Pascarella and Terenzini 2005)
  - lifetime earnings (Brewer, Eide, and Ehrenberg 1999)

- This (mis)allocation has consequences for long term economic growth (Hsieh et al. 2019)
Impact of Educational Choices of Women

- Such sorting affects women’s:
  - academic training (Zhang 2005)
  - network of peers (Winston and Zimmerman 2004)
  - access to labor opportunities (Pascarella and Terenzini 2005)
  - lifetime earnings (Brewer, Eide, and Ehrenberg 1999)

- This (mis)allocation has consequences for long term economic growth (Hsieh et al. 2019)
Possible Explanations

- Many possible reasons why we see this allocation:
  - value of academics
  - value of time
  - competitiveness (Niederle and Vesterlund 2007, Niederle and Vesterlund 2011, Buser et al. 2014)

- I study one possible explanation = Safety in public spaces
  - in a context where majority of students live at home and travel to college daily and
  - men don’t face sexual harassment in public spaces
Street Harassment

- 95% of females aged 16 to 49 years in Delhi feel unsafe in public spaces (ICRW 2013)

- 84% women under 40 in India report avoiding an area in their city because of harassment (Livingston 2015)
  - Delhi college students = 71%
Literature

- Fear of imagined dangers → individual behavior (Becker and Rubinstein 2011)

- Qualitative evidence in the psychology and criminology literature
  - women incur significant psychological costs from harassment (Langton and Truman 2014)
  - women change their behavior in response (Pain 1997):
    - mobility patterns (Hsu 2011, Keane 1998, Portel et al. 2011)
    - labor force participation at the extensive margin (Chakraborty et al. 2018, Siddique 2018) and intensive margin (Cook et al. 2020)
      ★ first study to measure the misallocation effects

- School attributes that matter from the school choice literature
  - proximity (Carneiro, Das, and Reis 2013)
  - academic attainment (Gallego and Hernando 2009)
  - student composition (Hastings, Kane, and Staiger 2009)
    ★ first study to evaluate travel safety

- Spatial frictions → gender disparities in human capital acquisition
  - choice of location (Mukherjee 2012, Burde and Linden 2013)
  - better roads or provision of transport (Muralidharan and Prakash 2017, Jacoby and Mansuri 2015, Cheema et al. 2020)
    ★ first study to measure the extent to which safety matters
Question

Does street harassment affect women’s college choice?
This Paper

- **Unique data**
  - students’ travel and college choices
  - route mapping using Google Maps
  - mobile app data on perceived safety

- **Descriptive evidence**
  - exploits admissions procedure to approximate changes in students’ choice set

- **Structural estimation**
  - mixed logit model
  - uses spatial variation in students’ origin, college locations and area safety
Background
Delhi University

- Delhi University (DU) is one of the best and largest universities in India:
  - 180,000 undergraduate students (2013-14) = 8% of students who passed high school exams in India

- DU is composed of 77 colleges:
  - each college has its own campus, staff, classes and placements
Colleges in DU

Size

- 700 - 1,500
- 1,501 - 2,200
- 2,201 - 3,000
- 3,001 - 3,700
- 3,701 - 5,000
Delhi University

- Majority of the students (72%) are from Delhi NCR
  - 99.1% live at home with their parents and travel to college every day

- Students travel mostly by public transport
  - 83% use public transport

- Admissions in DU
  - strictly based on students’ high school exam scores
  - cutoff scores for each college
    - determine each student’s complete choice set of colleges
Delhi University

- Majority of the students (72%) are from Delhi NCR
  - 99.1% live at home with their parents and travel to college every day

- Students travel mostly by public transport
  - 83% use public transport

- Admissions in DU
  - strictly based on students’ high school exam scores
  - cutoff scores for each college
    - determine each student’s complete choice set of colleges
Majority of the students (72%) are from Delhi NCR
  - 99.1% live at home with their parents and travel to college every day

Students travel mostly by public transport
  - 83% use public transport

Admissions in DU
  - strictly based on students’ high school exam scores
  - cutoff scores for each college
    * determine each student’s complete choice set of colleges
Data
Data Overview

1. Student Information
2. Travel Routes
3. Safety Data
Data Overview

1. Student Information
   - Choice set of colleges

2. Travel Routes
   - Potential routes to choice set colleges

3. Safety Data
Data Overview

1. Student Information
   - Choice set of colleges

2. Travel Routes
   - Potential routes to choice set colleges

3. Safety Data

- Actual travel route
- Google Maps
- Survey Data
Data Overview

1. **Student Information**
   - Choice set of colleges

2. **Travel Routes**
   - Potential routes to choice set colleges

3. **Safety Data**
   - Area safety and safety by travel mode

- **Survey Data**
- **Google Maps**
- **SafetiPin mobile app**
- **SafeCity analytical data**

- **Actual travel route**
Data Overview

1. **Student Information**
   - Choice set of colleges
   - Actual travel route

2. **Travel Routes**
   - Potential routes to choice set colleges
   - Safety of each travel route

3. **Safety Data**
   - Area safety and safety by travel mode

**Survey Data**

**Google Maps**

**SafetiPin mobile app**

**Safecity analytical data**
Detailed survey data of 3,800 male and female students across 8 colleges

- subject wise high school exam scores
- current and parental residential location
- exact travel route and modes
- household characteristics
- exposure to harassment

Data collected during class at a time scheduled with the professor.
Sample for Analysis

Delhi students who live with their parents and travel to college everyday

![Sample college map]

- Sample college
- Males
- Females

Kilometers

0 2 4 8 12 16
Route Mapping using Google Maps

- Map potential routes to chosen college and each college in choice set
  - up to four routes per Google based travel option
    - public transit
    - driving only
    - walking only
    - broken up into “travel legs” based on travel mode
- Map actual travel routes
  - landmarks
  - travel modes
  - departure time
- Overlap of reported routes with Google Map routes: 90%
Data

1. Student Information
2. Travel Routes
3. Safety Data
Safety Data

**SafetiPin mobile app data on perceived safety** (November 2013 - January 2016)

- characterizes the safety of location based on 9 parameters
- partially crowdsourced and in part collected by trained auditors
  - 98% contributors are 39 years or younger
  - 70% contributors are female
- over 26,000 audits across Delhi National Capital Region
- used as the base level of safety of an area
Safety Data

SafetiPin mobile app data on perceived safety (November 2013 - January 2016)

- characterizes the safety of location based on 9 parameters
- partially crowdsourced and in part collected by trained auditors
  - 98% contributors are 39 years or younger
  - 70% contributors are female
- over 26,000 audits across Delhi National Capital Region
- used as the base level of safety of an area
Safety Data by Travel Mode

Analytical data based on Safecity mobile app
- records personal stories of harassment in public spaces
  - data based on over 5,500 crowd-sourced reports

<table>
<thead>
<tr>
<th>Travel Mode</th>
<th>% Harassment Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus</td>
<td>0.40</td>
</tr>
<tr>
<td>Metro</td>
<td>0.16</td>
</tr>
<tr>
<td>Taxi</td>
<td>0.14</td>
</tr>
<tr>
<td>Auto Rickshaw</td>
<td>0.07</td>
</tr>
<tr>
<td>Train</td>
<td>0.05</td>
</tr>
</tbody>
</table>

- used to weight area safety by travel mode
How it all comes together
<table>
<thead>
<tr>
<th>Travel mode</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rickshaw</td>
<td>Green</td>
</tr>
<tr>
<td>Metro</td>
<td>Cyan</td>
</tr>
<tr>
<td>Walk</td>
<td>Orange</td>
</tr>
<tr>
<td>Bus</td>
<td>Pink</td>
</tr>
</tbody>
</table>

### Reported Travel Route

- **Student**: The student's route is marked with a blue triangle.
- **Chosen college**: The chosen college is marked with a star.

**Travel modes**:
- **Rickshaw**: The student uses a Rickshaw for part of the journey.
- **Metro**: The student takes the Metro for part of the journey.
- **Walk**: The student walks part of the journey.
- **Bus**: The student takes the bus for part of the journey.
Reported Travel Route

Travel mode:
- Student
- Chosen college
- Rickshaw
- Metro
- Walk
- Bus

Distance: ±0.5 Kilometers
Potential Routes to Choice Set Colleges

Travel mode
- Bus
- Driving
- Metro
- Walking

- Student
- Chosen college
- Choice set colleges
Travel across Safety Zones

Travel mode
- Rickshaw
- Metro
- Walk
- Bus
Safety Score of Travel Route

\[
\text{Safety of travel route} = \sum_m \sum_p \left[ \frac{\text{Area safety}_p \times \text{Route length}_{mp}}{\text{Total route length}} \times (1 - \text{Harassment}_{mp}) \right]
\]
Descriptive Evidence
Changes in Students’ Choice Set

- The ideal experiment = random allocation of college choice sets
  - student responses to variation in college and route attributes → underlying trade-offs

- Exploit DU’s admissions procedure to approximate the ideal experiment
  - strictly based on high school exam scores
Changes in Students’ Choice Set

- The ideal experiment = random allocation of college choice sets
  - student responses to variation in college and route attributes → underlying trade-offs

- Exploit DU’s admissions procedure to approximate the ideal experiment
  - strictly based on high school exam scores
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

1,228 unique pairs
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

1,228 unique pairs
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

1,228 unique pairs
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

1,228 unique pairs
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

1,228 unique pairs
Choice Relative to Neighbors

- Case-control match - compare choice of index student with neighbor:
  - lives in a 1.5km radius
  - same gender
  - same major
  - same year of admission

\[ \{1,228 \text{ unique pairs}\} \]
Choice Relative to Neighbors

△ index student’s exam score relative to neighbor → (changes students’ choice set)

△ index student’s choice response:
  - safety
  - quality
  - time
  - cost
(underlying preferences)
Choice Relative to Neighbors

△ index student’s exam score relative to neighbor →
(changes students’ choice set)

△ index student’s choice response:
- safety
- quality
- time
- cost

(underlying preferences)
women choose relatively safer routes as their choice set expands
men don’t respond in terms of safety

positive quality gradient for both women and men
significantly greater for men
Structural Model
Student Choice Model

Student $i$’s utility is given by:

$$U_{ir}^c = \beta_{iq} Q_{ir}^c + \beta_{is} S_{ir}^c + \beta_{it} T_{ir}^c + \beta_{ip} P_{ir}^c + \epsilon_{ir}^c$$

where

- $Q_{ir}^c$ is college quality
- $S_{ir}^c$ is safety of the travel route to college
- $T_{ir}^c$ is the daily travel time to college
- $P_{ir}^c$ is the monthly travel cost to college
Student Choice Model

Student $i$'s utility is given by:

$$U_{ir}^c = \beta_{iq} Q_{i}^c + \beta_{is} S_{ir}^c + \beta_{it} T_{ir}^c + \beta_{ip} P_{ir}^c + \varepsilon_{ir}^c$$

where

- $Q_{i}^c$ is college quality = cutoff score
- $S_{ir}^c$ is safety of the travel route to college
- $T_{ir}^c$ is the daily travel time to college
- $P_{ir}^c$ is the monthly travel cost to college
Student Choice Model

Student $i$’s utility is given by:

$$U_{ir}^c = \beta_{iq} Q_i^c + \beta_{ir} S_{ir}^c + \beta_{it} T_{ir}^c + \beta_{ip} P_{ir}^c + \epsilon_{ir}^c$$

where

- $Q_i^c$ is college quality
- $S_{ir}^c$ is safety of the travel route to college
- $T_{ir}^c$ is the daily travel time to college
- $P_{ir}^c$ is the monthly travel cost to college
Student Choice Model

Student \( i \)'s utility is given by:

\[
U_{ir}^c = \beta_{iq} Q_i^c + \beta_{is} S_{ir}^c + \beta_{it} T_{ir}^c + \beta_{ip} P_{ir}^c + \epsilon_{ir}^c
\]

where

- \( Q_i^c \) is college quality
- \( S_{ir}^c \) is safety of the travel route to college
- \( T_{ir}^c \) is the daily travel time to college, in minutes
- \( P_{ir}^c \) is the monthly travel cost to college
Student Choice Model

Student $i$'s utility is given by:

$$U_{ir}^c = \beta_{iq} Q_{i}^c + \beta_{is} S_{ir}^c + \beta_{it} T_{ir}^c + \beta_{ip} P_{ir}^c + \epsilon_{ir}^c$$

where

- $Q_{i}^c$ is college quality
- $S_{ir}^c$ is safety of the travel route to college
- $T_{ir}^c$ is the daily travel time to college
- $P_{ir}^c$ is the monthly travel cost to college, in thousand Rs.
Student Choice Model

- Student $i$'s chooses college $c$ and route $r$ ($d_{ir}^c = 1$) such that the choice maximizes his or her utility over all possible colleges and routes in their choice set:

$$d_{ir}^c = 1 \text{ if and only if } U_{ic} > U_{is} \quad \forall b \neq c \quad \forall r \neq s$$

$$d_{ir}^c = 0 \quad \text{otherwise}$$

- Main variable of interest:

$$MRS_{i_{QS}} \equiv \frac{\Delta Q_{ic}}{\Delta S_{ir}^c} = \frac{\beta_{is}}{\beta_{iq}}$$

(1)

Identification assumption:

- the location and attributes of the students, colleges, and possible routes are exogenous to the process of college and route choice
Mixed Logit Model

- Flexible substitution patterns albeit by imposing structure on the distribution of preferences

- Heterogeneity of preferences
  - eg. dislike harassment vs. + differential exposure
  - eg. value quality vs. + high decision making costs

- Estimation
  - random coefficients on safety, quality, time and cost
  - separately estimated for men and women
  - augment with fixed coefficients on additional college and route attributes
Distributional assumptions

- Random coefficients:
  - route safety, $\beta_s$ and college quality, $\beta_q \sim$ triangular
  - travel time, $\beta_t$ and travel costs, $\beta_c \sim$ restricted triangular, assumed to be non-positive
Identification

Several aspects of the context and data help in identification of parameters

1. Students live at home with their parents
   - parents unlikely to base their residential choices on location of their children’s future preferred colleges
   - high rates of home ownership: 82% of Delhi residents live in owned houses
   - identify parameters separately from residential sorting

2. No sorting of colleges by quality and neighborhood or student characteristics
   - student characteristics: Exam scores, Gender, SES
   - area characteristics: Area safety
   - route characteristics: Route safety

3. Admission cutoffs do not seem to take into account safety concerns
   - unable to predict the advantage given to women in cutoffs
   - Predict advantage
Identification

Several aspects of the context and data help in identification of parameters

1. Students live at home with their parents
   - parents unlikely to base their residential choices on location of their children’s future preferred colleges
   - high rates of home ownership: 82% of Delhi residents live in owned houses
   - identify parameters separately from residential sorting

2. No sorting of colleges by quality and neighborhood or student characteristics
   - student characteristics
   - area characteristics
   - route characteristics

3. Admission cutoffs do not seem to take into account safety concerns
   - unable to predict the advantage given to women in cutoffs
Identification

Several aspects of the context and data help in identification of parameters

1. Students live at home with their parents
   - parents unlikely to base their residential choices on location of their children’s future preferred colleges
   - high rates of home ownership: 82% of Delhi residents live in owned houses
   - identify parameters separately from residential sorting

2. No sorting of colleges by quality and neighborhood or student characteristics
   - student characteristics Exam scores Gender SES
   - area characteristics Area safety
   - route characteristics Route safety

3. Admission cutoffs do not seem to take into account safety concerns
   - unable to predict the advantage given to women in cutoffs
Results
## Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>Female (1)</th>
<th>Male (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS (Safety, Score)</td>
<td>-8.80</td>
<td>-2.11</td>
</tr>
<tr>
<td>pp per SD of safety</td>
<td>[-39.05, -3.15]</td>
<td>[-3.30, -1.56]</td>
</tr>
<tr>
<td>MRS (Safety, Cost)</td>
<td>1.45</td>
<td>0.83</td>
</tr>
<tr>
<td>’000 Rs. per SD of safety</td>
<td>[1.33, 1.75]</td>
<td>[0.76, 1.05]</td>
</tr>
<tr>
<td>MRS (Safety, Time)</td>
<td>26.77</td>
<td>20.80</td>
</tr>
<tr>
<td>minutes per SD of safety</td>
<td>[24.63, 30.47]</td>
<td>[19.02, 24.41]</td>
</tr>
</tbody>
</table>

**Notes:** MRS are measured in terms of the SD of route safety within a students’ choice set. Confidence interval in square bracket, computed using delta method.
## Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>Female (1)</th>
<th>Male (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS (Safety, Score)</td>
<td>-8.80</td>
<td>-2.11</td>
</tr>
<tr>
<td>pp per SD of safety</td>
<td>[-39.05, -3.15]</td>
<td>[-3.30, -1.56]</td>
</tr>
<tr>
<td>MRS (Safety, Cost)</td>
<td>1.45</td>
<td>0.83</td>
</tr>
<tr>
<td>'000 Rs. per SD of safety</td>
<td>[1.33, 1.75]</td>
<td>[0.76, 1.05]</td>
</tr>
<tr>
<td>MRS (Safety, Time)</td>
<td>26.77</td>
<td>20.80</td>
</tr>
<tr>
<td>minutes per SD of safety</td>
<td>[24.63, 30.47]</td>
<td>[19.02, 24.41]</td>
</tr>
</tbody>
</table>

**Notes:** MRS are measured in terms of the SD of route safety *within* a students’ choice set. Confidence interval in square bracket, computed using delta method.
## Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>Female (1)</th>
<th>Male (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MRS (Safety, Score)</strong></td>
<td>-8.80</td>
<td>-2.11</td>
</tr>
<tr>
<td>pp per SD of safety</td>
<td>[-39.05, -3.15]</td>
<td>[-3.30, -1.56]</td>
</tr>
<tr>
<td><strong>MRS (Safety, Cost)</strong></td>
<td>1.45</td>
<td>0.83</td>
</tr>
<tr>
<td>'000 Rs. per SD of safety</td>
<td>[1.33, 1.75]</td>
<td>[0.76, 1.05]</td>
</tr>
<tr>
<td><strong>MRS (Safety, Time)</strong></td>
<td>26.77</td>
<td>20.80</td>
</tr>
<tr>
<td>minutes per SD of safety</td>
<td>[24.63, 30.47]</td>
<td>[19.02, 24.41]</td>
</tr>
</tbody>
</table>

**Notes:** MRS are measured in terms of the SD of route safety *within* a students' choice set. Confidence interval in square bracket, computed using delta method.

---

**Rupees per SD of safety**

- Male: 9,950
- Female: 17,500
Robustness

- Alternative majors
- Alternative safety measures
- Alternative samples
Conclusion

- Highlights the long term consequences of everyday harassment
  - 17% decline in PDV of post-college salaries using estimates from Sekhri (2019)

- Gender gaps in parental investments → wider gender gaps in lifetime earnings
  - this paper identifies an additional mechanism – the lack of safety

- Implications for other economic decisions
  - e.g. can help explain the puzzle of low female labor force participation in India
Thank you

gborker@worldbank.org
Delhi University

- Of the 58 general education undergraduate colleges:
  - 22 colleges are women only
  - 8 of the colleges are evening colleges

- 14 colleges have on-campus accommodation
  - for 5% of the students in the University

- 21 colleges have lower cutoffs for women (2014)
  - advantage varies from 1pp to 5pp

- DU fees is on average $\frac{1}{18}th$ to $\frac{1}{9}th$ of the fees in a private university in Delhi
<table>
<thead>
<tr>
<th></th>
<th>Light (Night)</th>
<th>Openness</th>
<th>Visibility</th>
<th>People</th>
<th>Security</th>
<th>Walk Path</th>
<th>Public Transport</th>
<th>Gender Usage</th>
<th>Feeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None. No street or other lights.</td>
<td>Not Open. Many blind corners and no clear sightline.</td>
<td>No eyes. No windows or entrances (to shops or residences) or street vendors.</td>
<td>Deserted. No one in sight.</td>
<td>None. No guards or police visible in surrounding area.</td>
<td>None. No walking path available.</td>
<td>Unavailable. No metro or bus stop, auto/ rickshaw within 10 minutes walk.</td>
<td>Not diverse. No one in sight, or only men.</td>
<td>Frightening. Will never venture here without sufficient escort.</td>
</tr>
<tr>
<td></td>
<td>Little. Can see lights, but barely reaches this spot.</td>
<td>Partly Open. Able to see a little ahead and around.</td>
<td>Few eyes. Less than 5 windows or entrances or street vendors.</td>
<td>Few people. Less than 10 people in sight.</td>
<td>Minimal. Some private security visible in surrounding area but not nearby.</td>
<td>Poor. Path exists but in very bad condition.</td>
<td>Distant. Metro or bus stop, auto/ rickshaw within 10 minutes walk.</td>
<td>Somewhat diverse. Mostly men, very few women or children.</td>
<td>Uncomfortable. Will avoid whenever possible.</td>
</tr>
</tbody>
</table>
Score distribution

- High school exam score: Female vs. Male

- Probability distribution for high school exam scores:
  - Probability: 0.01, 0.02, 0.03, 0.04
  - Score range: 50 to 100

Graph showing distribution of high school exam scores for females (dashed line) and males (solid line).
Colleges’ Quality Distribution

The sample colleges are fairly evenly spread across the quality distribution.

Each colored bar represents a college in the Full Survey Sample.

Quality = first cutoff score for general category male students.
Confidential administrative data for all students in the 8 Full Survey Sample colleges
- current and parental residential location
- social category

Short survey data for 800 male and female students across 32 other colleges in DU
- combination of intercept and online survey
- current and parental residential location
- high school exam scores
## Comparing Samples

<table>
<thead>
<tr>
<th>Data</th>
<th>Female Full Survey</th>
<th>Female Short Survey</th>
<th>Female Admin. Data</th>
<th>Male Full Survey</th>
<th>Male Short Survey</th>
<th>Male Admin. Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi residents</td>
<td>1,767</td>
<td>459</td>
<td>11,450</td>
<td>946</td>
<td>171</td>
<td>8,288</td>
</tr>
<tr>
<td>Proportion of surveyed</td>
<td>0.74</td>
<td>0.82</td>
<td>0.80</td>
<td>0.67</td>
<td>0.62</td>
<td>0.68</td>
</tr>
</tbody>
</table>

### Social Category

<table>
<thead>
<tr>
<th>Social Category</th>
<th>Female</th>
<th>Male</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>0.75</td>
<td>0.56</td>
<td>[2.19]</td>
<td>[0.81]</td>
<td>[-3.05]</td>
</tr>
<tr>
<td>SC</td>
<td>0.12</td>
<td>0.20</td>
<td>[0.14]</td>
<td>[0.14]</td>
<td>0.19</td>
</tr>
<tr>
<td>ST</td>
<td>0.01</td>
<td>0.02</td>
<td>[-0.02]</td>
<td>[0.02]</td>
<td>0.03</td>
</tr>
<tr>
<td>OBC</td>
<td>0.11</td>
<td>0.22</td>
<td>[3.69]</td>
<td>[-0.94]</td>
<td>0.27</td>
</tr>
</tbody>
</table>

### High school exam score (%)

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>84.88</td>
<td>83.90</td>
<td>-</td>
<td>82.70</td>
<td>82.03</td>
</tr>
</tbody>
</table>

### Distance to college (kms.)

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13.06</td>
<td>13.05</td>
<td>15.63</td>
<td>14.30</td>
<td></td>
</tr>
</tbody>
</table>

### Distance to center (kms.)

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.42</td>
<td>15.99</td>
<td>15.93</td>
<td>16.49</td>
<td>15.99</td>
</tr>
</tbody>
</table>
Short Survey Sample

- Males
- Females

Kilometers ±
Colleges in Sample

- Full Survey Sample
- Short Survey Sample
- Other colleges

Distance Scale:
0 1.5 3 6 9 12 Kilometers
Proportion of accepted students who enrolled

![Graph showing the proportion of students enrolled based on cutoff scores.](image)

- Proportion of accepted students who enrolled
- Cutoff score range: 70 to 100
- Proportion of students: 0.3 to 0.6
Exam Scores of Students

Each bar represents average exam scores for students in the neighborhood of a college. The colleges are sorted in ascending order of quality.
Proportion of Female Students

Each bar represents proportion of female students in the neighborhood of a college. The colleges are sorted in ascending order of quality.
Each bar represents average SES index for students in the neighborhood of a college. The colleges are sorted in ascending order of quality.
Each bar represents neighborhood safety for a college. The colleges are sorted in ascending order of quality.
Each bar represents neighborhood safety for a college. The colleges are sorted in ascending order of quality.
# Proportion of Female Students

## Advantage to Women in 2014

<table>
<thead>
<tr>
<th></th>
<th>Advantage to women in 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion female in 2013</td>
<td>-0.088</td>
</tr>
<tr>
<td></td>
<td>(2.049)</td>
</tr>
<tr>
<td>College neighborhood safety</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(0.244)</td>
</tr>
<tr>
<td>Boarding</td>
<td>-1.967</td>
</tr>
<tr>
<td></td>
<td>(0.626)</td>
</tr>
<tr>
<td>Number of majors</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
</tr>
<tr>
<td>Size of college</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Annual tuition</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.597</td>
</tr>
<tr>
<td></td>
<td>(0.994)</td>
</tr>
<tr>
<td>Mean of Y</td>
<td>1.415</td>
</tr>
<tr>
<td>Observations</td>
<td>41</td>
</tr>
</tbody>
</table>
Estimating the MRS

1. Predict optimal route for each student x college
2. Assume predicted route is chosen
3. Predict college based on predicted route

- Chosen route is predicted route in 23% of cases
- SD in safety within choice set is 26% lower than overall SD
- Predicted college has same rank as predicted in 39% of cases
Mixed Logit Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th></th>
<th>Male</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Random coefficients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Route safety</td>
<td>0.705</td>
<td>1.010</td>
<td>0.468</td>
<td>0.570</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.033)</td>
<td>(0.034)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Cutoff score</td>
<td>0.065</td>
<td>0.045</td>
<td>0.143</td>
<td>0.162</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Daily travel time</td>
<td>-0.014</td>
<td>-0.023</td>
<td>-0.014</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Monthly travel cost</td>
<td>-0.256</td>
<td>-0.115</td>
<td>-0.352</td>
<td>-0.303</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.026)</td>
<td>(0.029)</td>
</tr>
<tr>
<td><strong>Fixed coefficients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College neighborhood safety</td>
<td>0.159</td>
<td></td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td></td>
<td>(0.042)</td>
<td></td>
</tr>
<tr>
<td>Size of college</td>
<td>0.001</td>
<td></td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Women’s only college</td>
<td>0.591</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public transport mode</td>
<td>1.581</td>
<td></td>
<td>0.601</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td></td>
<td>(0.109)</td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>1,767</td>
<td>1,767</td>
<td>946</td>
<td>946</td>
</tr>
<tr>
<td>Observations</td>
<td>289,121</td>
<td>298,121</td>
<td>112,958</td>
<td>112,958</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-7985.66</td>
<td>-6825.76</td>
<td>-3901.95</td>
<td>-3807.22</td>
</tr>
</tbody>
</table>
## Augmented Specification

<table>
<thead>
<tr>
<th></th>
<th>Female (1)</th>
<th>Male (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MRS (Safety, Score)</strong></td>
<td>-13.52</td>
<td>-2.93</td>
</tr>
<tr>
<td>pp per SD of safety</td>
<td>[-28.79, -7.69]</td>
<td>[-12.20, -1.19]</td>
</tr>
<tr>
<td><strong>MRS (Safety, Cost)</strong></td>
<td>4.47</td>
<td>1.13</td>
</tr>
<tr>
<td>'000 Rs. per SD of safety</td>
<td>[4.28, 4.86]</td>
<td>[1.04, 1.39]</td>
</tr>
<tr>
<td><strong>MRS (Safety, Time)</strong></td>
<td>22.70</td>
<td>18.13</td>
</tr>
<tr>
<td>minutes per SD of safety</td>
<td>[19.70, 27.74]</td>
<td>[16.11, 22.62]</td>
</tr>
</tbody>
</table>

Notes: MRS are measured in terms of the SD of route safety within a students' choice set. Confidence interval in square bracket, computed using delta method. This specification includes controls for college size, college neighborhood safety, an indicator for women's only colleges and an indicator for whether the dominant mode of the travel route is public transport characterized by group travel.
Major Overlap

- All results conditional on major choice
- Students apply to several majors → could be a margin of choice

<table>
<thead>
<tr>
<th>Number of Majors</th>
<th>% students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.41</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
</tr>
<tr>
<td>5</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*Notes: This table shows the number of majors students applied for at the time of admission and the proportion of students.

- Related majors have significant overlap

<table>
<thead>
<tr>
<th>Related Majors</th>
<th>Overlap in choice set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts General, Commerce General</td>
<td>0.96</td>
</tr>
<tr>
<td>Commerce, Commerce General</td>
<td>0.78</td>
</tr>
<tr>
<td>Commerce, Economics</td>
<td>0.84</td>
</tr>
<tr>
<td>History, English</td>
<td>0.77</td>
</tr>
<tr>
<td>Political Science, Hindi</td>
<td>0.93</td>
</tr>
<tr>
<td>Political Science, English</td>
<td>0.82</td>
</tr>
<tr>
<td>Political Science, History</td>
<td>0.76</td>
</tr>
</tbody>
</table>

*Notes: This table shows the percentage overlap in college choice sets for related majors.
## Alternative Safety Measures

<table>
<thead>
<tr>
<th></th>
<th>Female (1)</th>
<th>Male (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Excluding &quot;Crowd&quot;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Route safety (SD)</td>
<td>0.504</td>
<td>0.394</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Cutoff score</td>
<td>0.062</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Route time (mins.)</td>
<td>-0.015</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Monthly travel cost (thousand Rs.)</td>
<td>-0.173</td>
<td>-0.278</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.024)</td>
</tr>
<tr>
<td><strong>Panel B: Excluding &quot;Light&quot;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Route safety (SD)</td>
<td>0.496</td>
<td>0.325</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Cutoff score</td>
<td>0.060</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Daily travel time (mins.)</td>
<td>-0.015</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Monthly travel cost (thousand Rs.)</td>
<td>-0.175</td>
<td>-0.279</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Number of students</td>
<td>1,767</td>
<td>946</td>
</tr>
<tr>
<td>Observations</td>
<td>289,121</td>
<td>112,958</td>
</tr>
</tbody>
</table>
Reported Crime and Perceived Safety

Figure: Reported Crime and Perceived Safety

Notes: The figure shows the coefficient from a district level regression of log of rapes in 2013 on average area safety and log of the number of the 15 to 34 year old females. Data on crimes is from the National Crime Records Bureau. The four types of crime are all the crimes against women that could potentially take place in public spaces.