An impact evaluation of class attendance policy on students' learning outcomes

Evidence from IIIT-Delhi

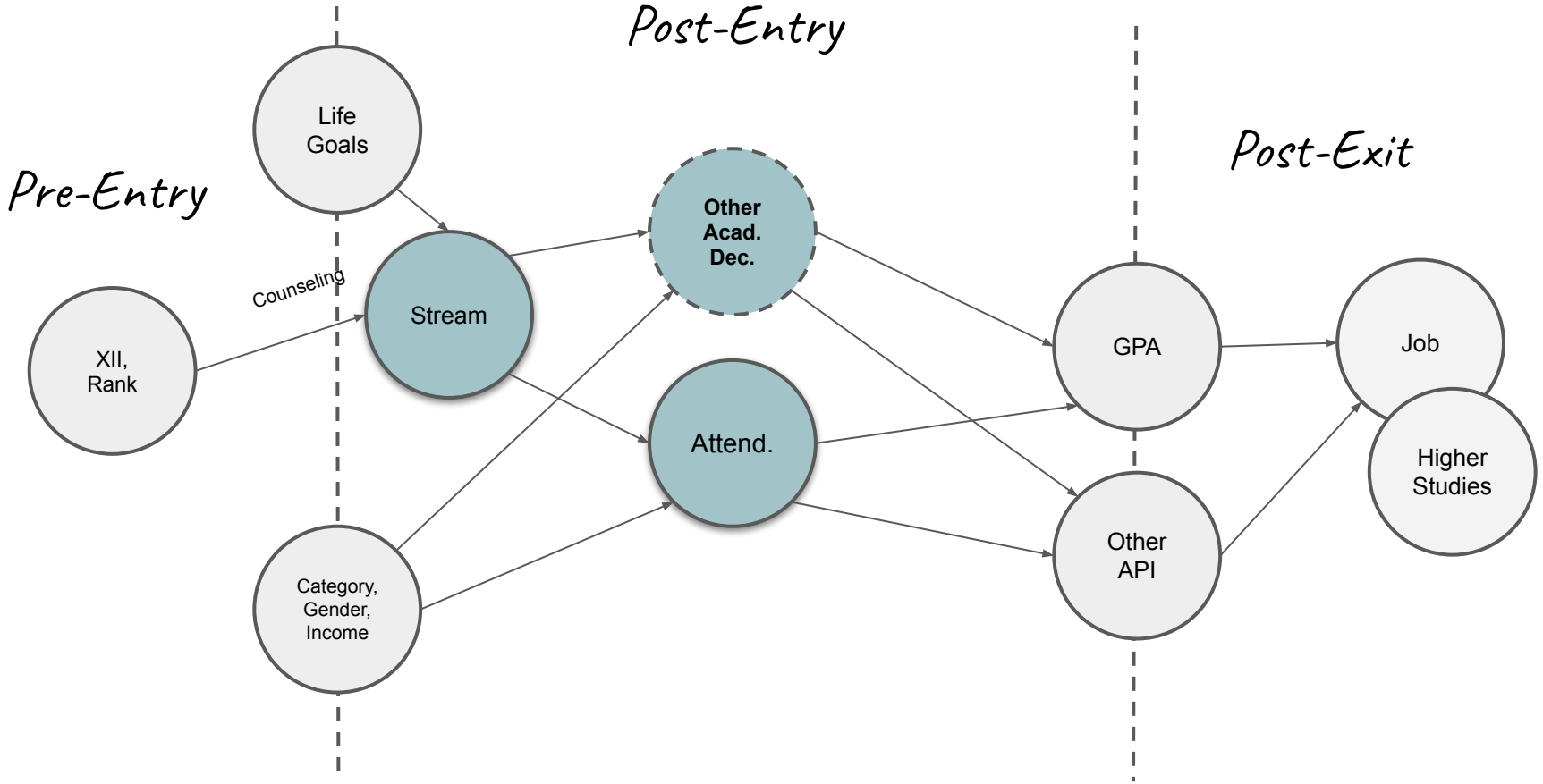
Anshika, Anvit, Mukul, Nischal, Pranay, Saif, Sumedha
Supervisor: Dr. Gaurav Arora
Advisers: Dr. Sanjit Kaul, Ms. Sheetu Ahuja, Ms. Anshu Dureja

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Background

- **Undergraduate student attendance far below perfect (U.K. Aliev, 2017)**
- **Learning happens in classrooms**
- **Attendance raises academic performance**
- **In class interaction improves students learning**
- **Why do then students miss classes?**

Journey Map of a IIITD entrant



Evidence

1. Anonymized data for 6 different batches of freshman year students
2. 6 core (required) courses in the 1st year. Everyone did all courses
    a. Courses grouped into 3:
        i. CSE (IP, DC, DSA)
        ii. Math (M1, PS)
        iii. COM
3. Two time periods 2014 - 2016 and 2017 - 2019 (n = 1228)

Source: Admin section, IIITD[1]

1 Thanks particularly to Ms. Anshu Dureja for spending time on the data.

Findings in a Nutshell

1. **GPA $\xleftarrow{+}$ (Attendance)**
   (impact is evident across courses, ~ 10% more[1] can earn an additional grade point)
2. **GPA $\xleftarrow{+}$ (Class XII, Course)**
   (stronger in CSE, Math courses)
3. **GPA $\leftarrow$ (Caste Category, Course)**
   (SC/ST < General. Gap large in Math despite similar attendance. Attendance is the first step.)
4. **GPA $\xleftarrow{+}$ (Hosteler), Attendance $\xleftarrow{-}$(Hosteler)**
   (Peer-group effect on students' learning outcome? Proximity to campus resources?)
5. **Attendance $\leftarrow$ (Winter/Monsoon)**
   (9am summer vs 9am winter)

$\xleftarrow{+}$ Positive Impact

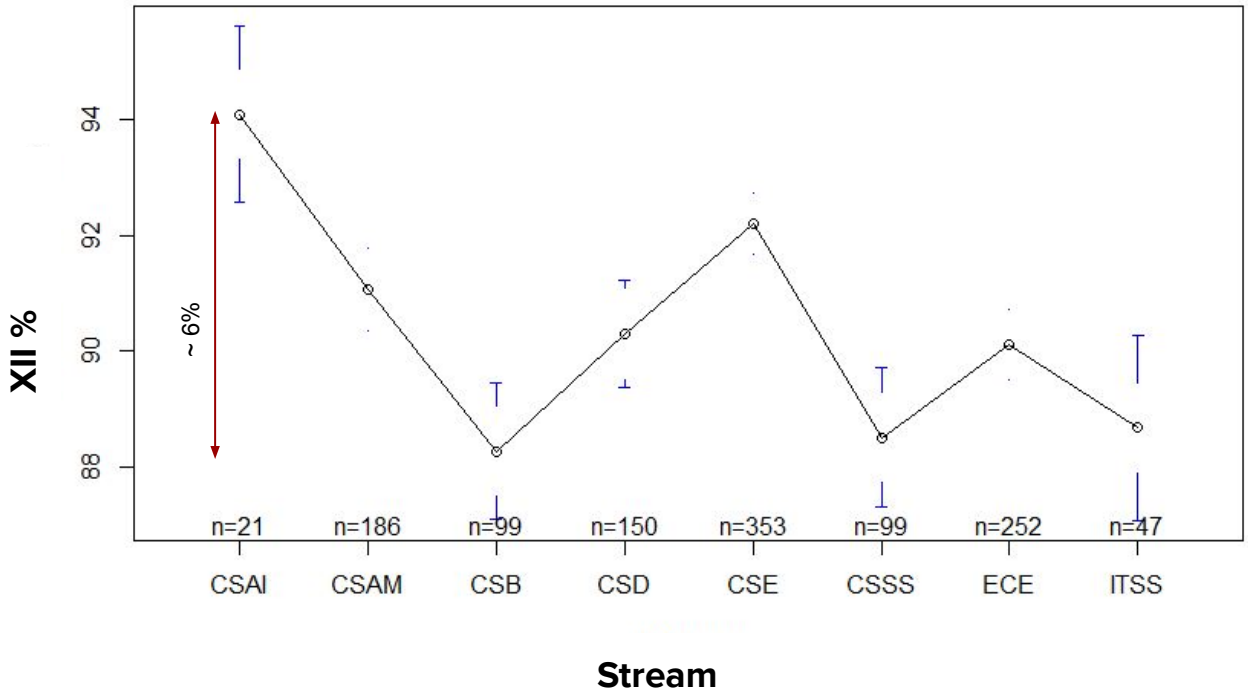
$\xleftarrow{-}$ Inverse/Negative Impact

$\leftarrow$ Causal Link

1. Approx. 5 lectures in a 16 week semester.

Relative Standing: Pre-Entry

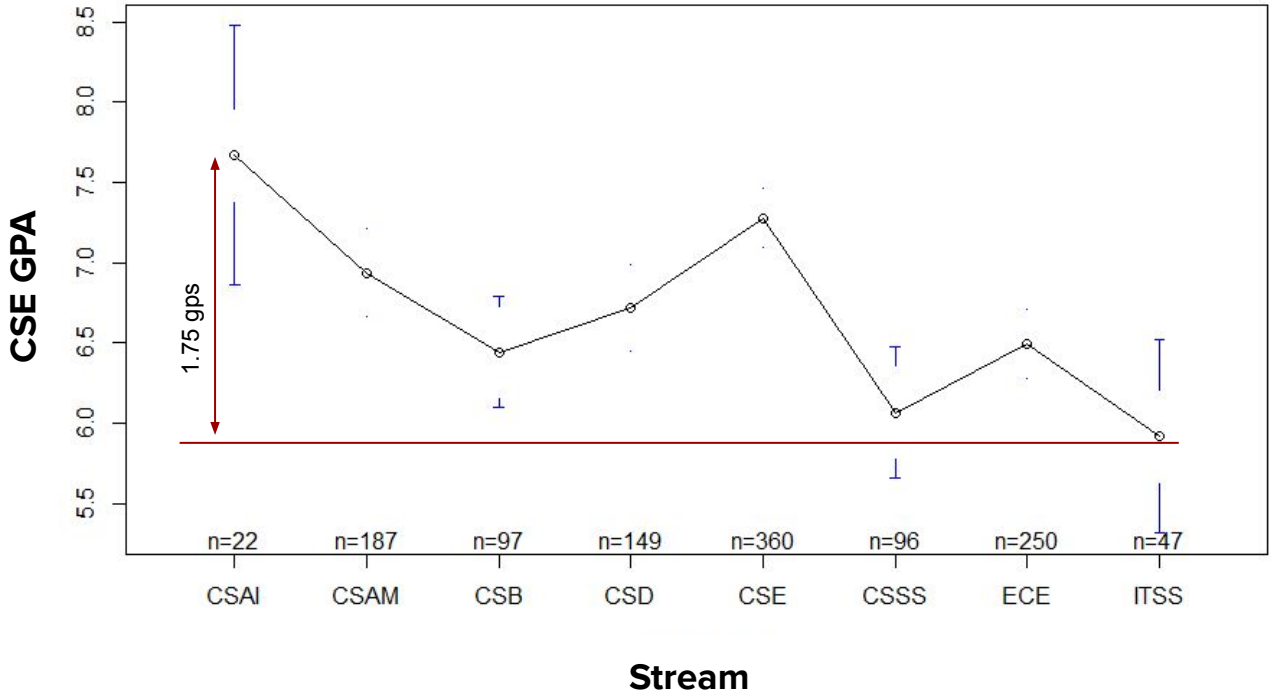
XIIth Percentage (by Stream)

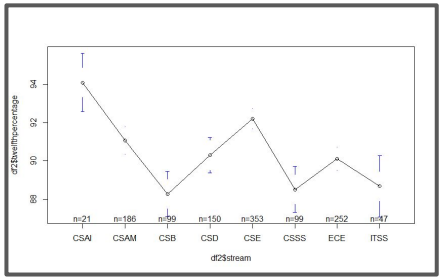


Same thing with rank.

Relative Standing Post-Entry

CSE GPA (by Stream)



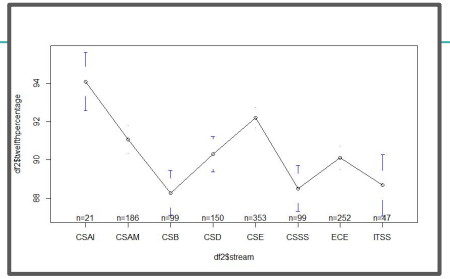


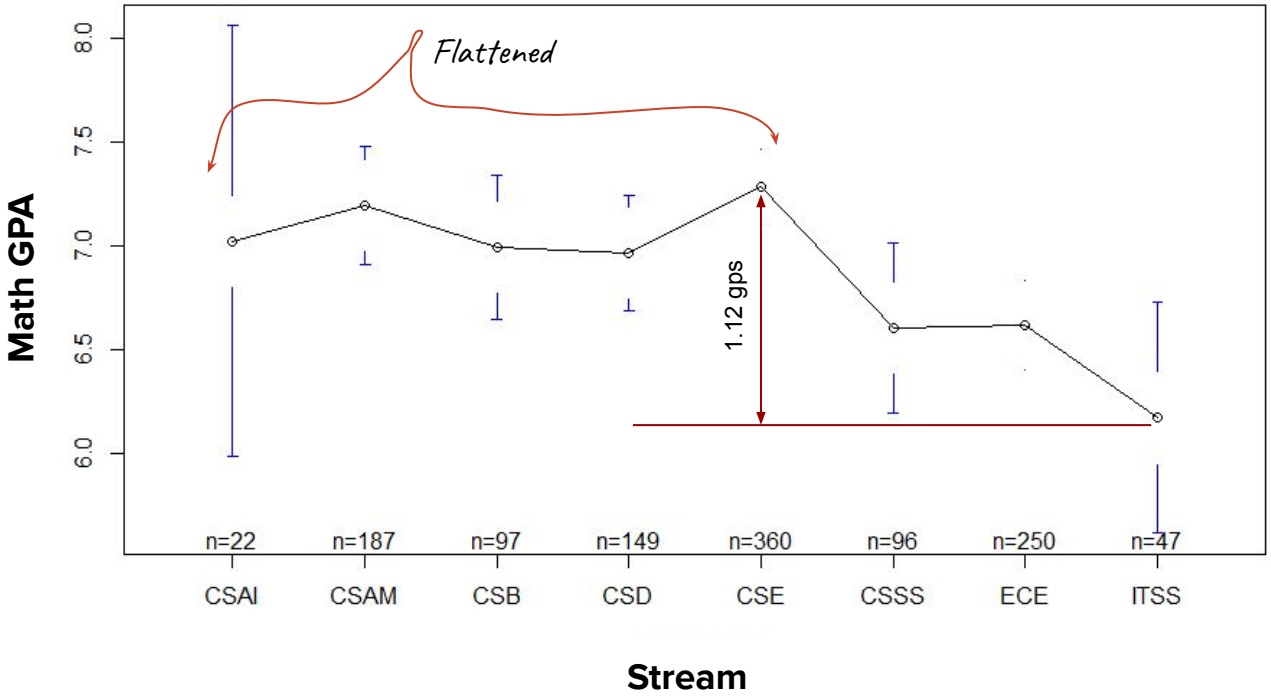
Same as pre-entry

Is Math different?

Relative Standing: Post-Entry

Math GPA (by Stream)



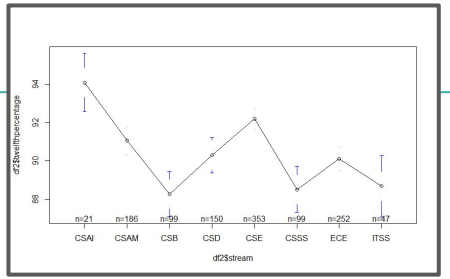


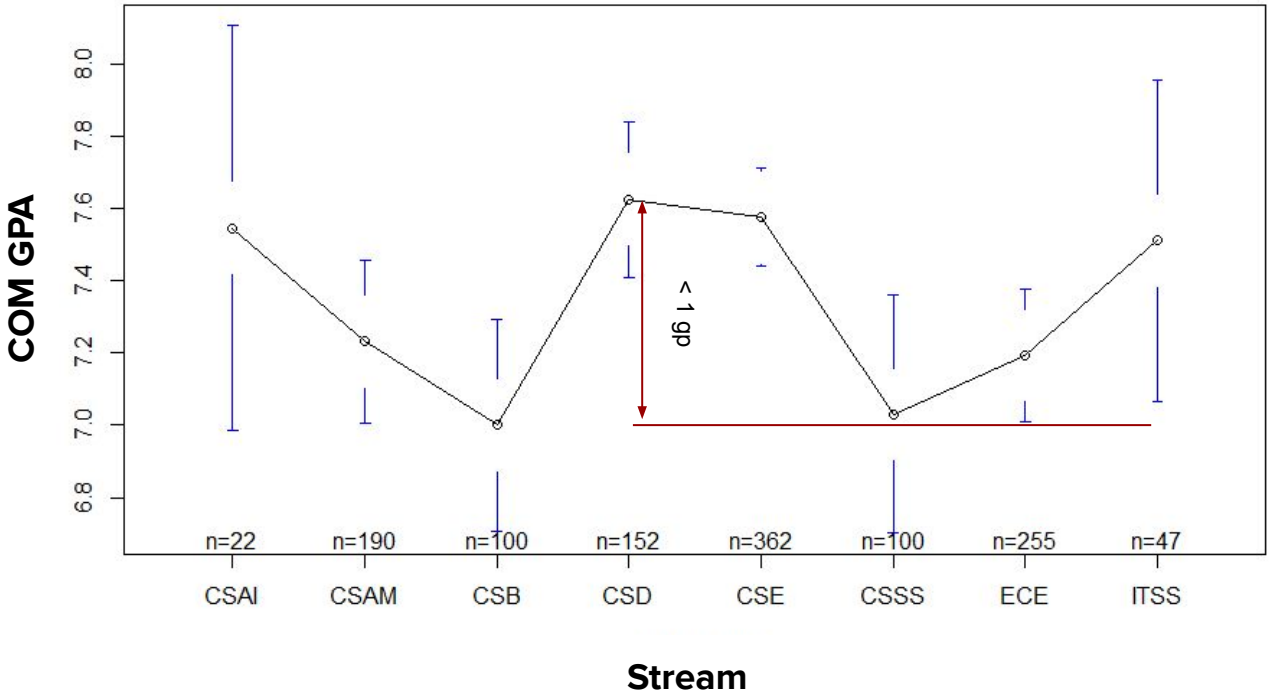
Not really the same as pre-entry.

The CSAI folks can't replicate pre-entry stardom in Mathematics.

Relative Standing: Post-Entry

COM GPA (by Stream)





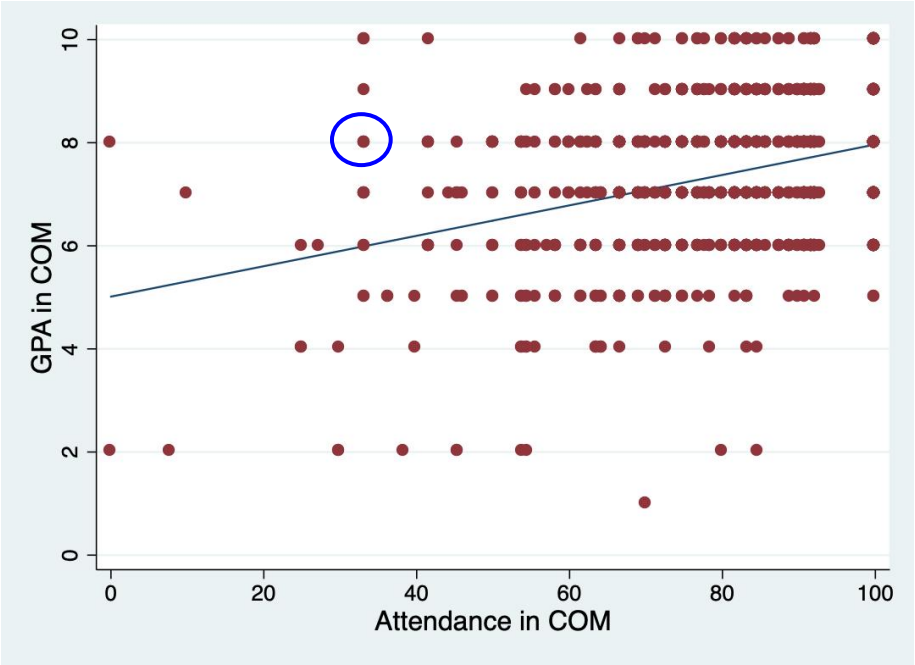
Clearly not the same as pre-entry.

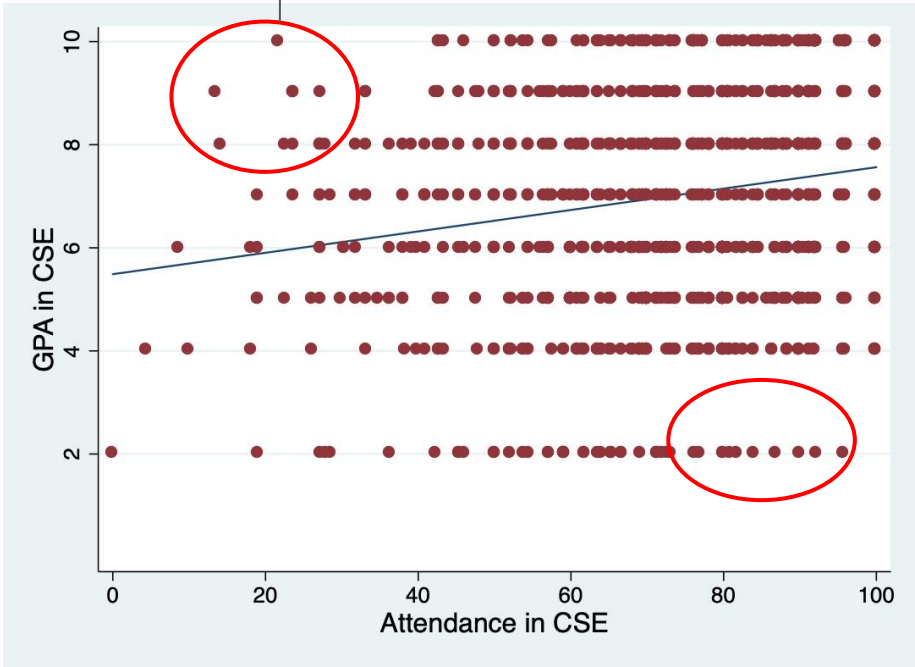
Design, ITSS shine. CSE and CSAI also hold on.

GPA vs Attendance

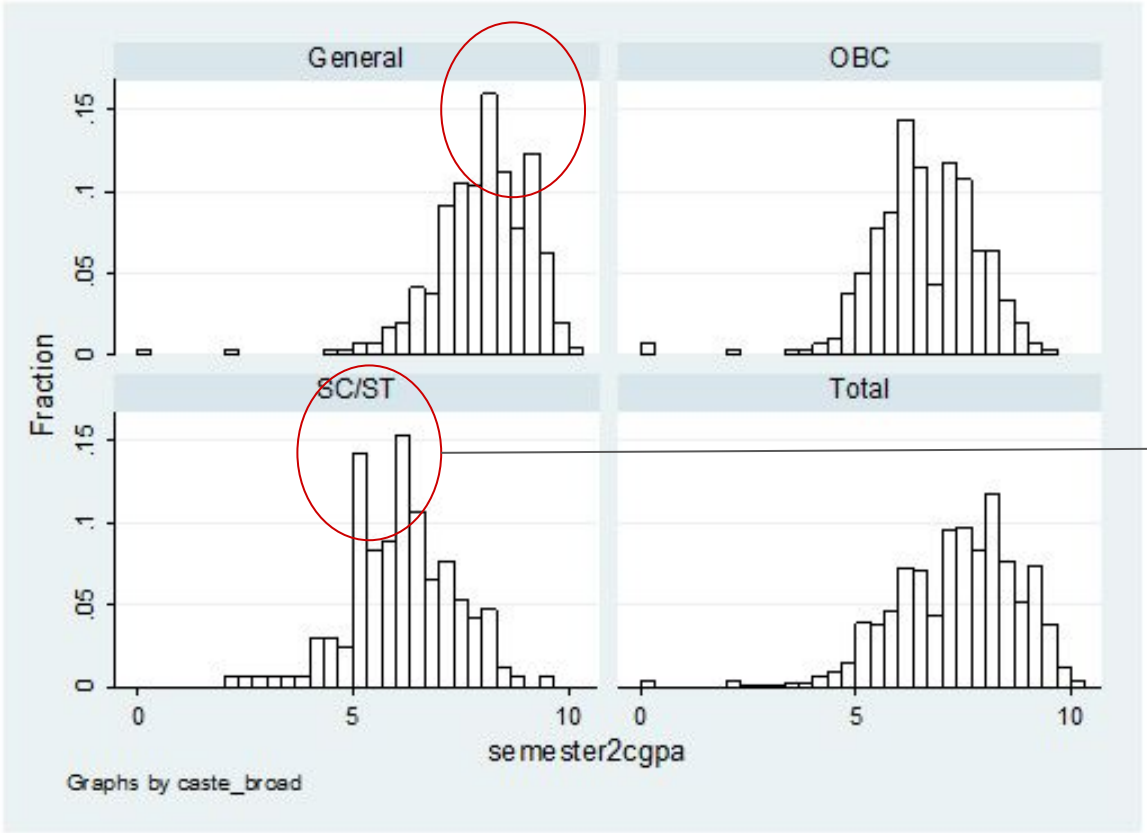
- GPA positively correlated with attendance
- Significant outliers

Students having very low attendance but securing high GPAs





GPA trends by social groups



General category individuals tend to have higher GPAs as compared to SC/STs.

With OBCs, this difference is still present but not as stark.

The probability of lower grades is higher.
We see a similar pattern in GPAs for all courses, especially Mathematics.

Social Groups Findings - Responses

Category	n	%
SC/ST	173	14.1
OBC	303	24.7
General	661	53.8
Missing	91	7.4
Total	1228	100

Social Groupings: Population

The CSE, Math and COM GPA fail the Shapiro Wilke's test of normality. Thus we have to correct for this.

CSE GPA by Social Groups

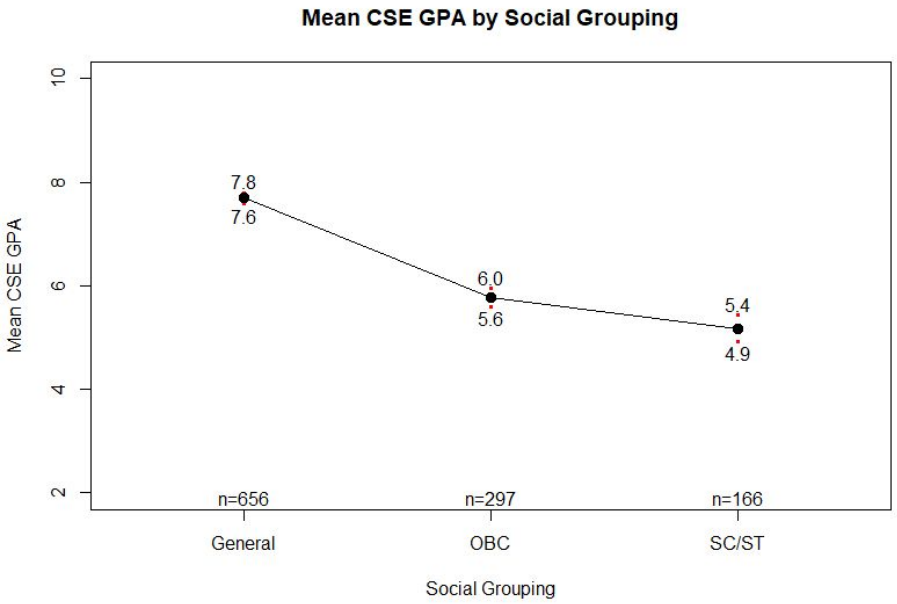


Table 1: Mean CSE GPA by Social Groupings. General is base category.

	Dependent variable:
	cse_gp
caste_broadOBC	−1.913***
	(0.105)
caste_broadSC/ST	−2.514***
	(0.131)
Constant	7.691***
	(0.059)
Observations	1,119
R^2	0.333
Adjusted R^2	0.332
Residual Std. Error	1.508 (df = 1116)
F Statistic	278.942*** (df = 2; 1116)
Note:	*p<0.1; **p<0.05; ***p<0.01

COM GPA by Social Groups

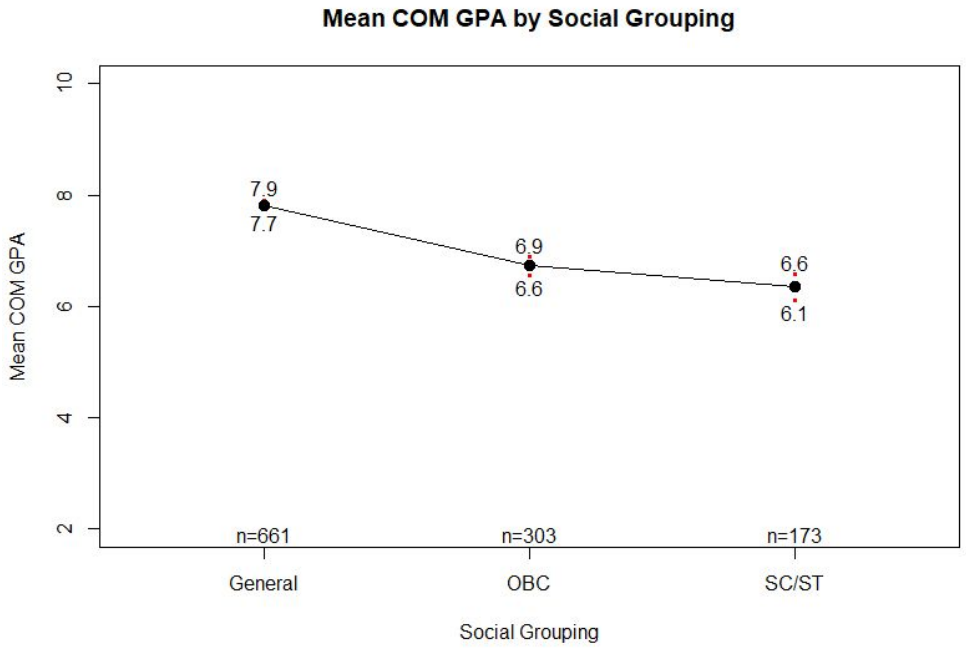


Table 3: Mean COM GPA by Social Groupings. General is base category.

	Dependent variable:
	com_gp
caste_broadOBC	−1.095***
	(0.090)
caste_broadSC/ST	−1.472***
	(0.111)
Constant	7.825***
	(0.051)
Observations	1,137
R^2	0.186
Adjusted R^2	0.184
Residual Std. Error	1.299 (df = 1134)
F Statistic	129.143*** (df = 2; 1134)
Note:	*p<0.1; **p<0.05; ***p<0.01

Math GPA by Social Grouping

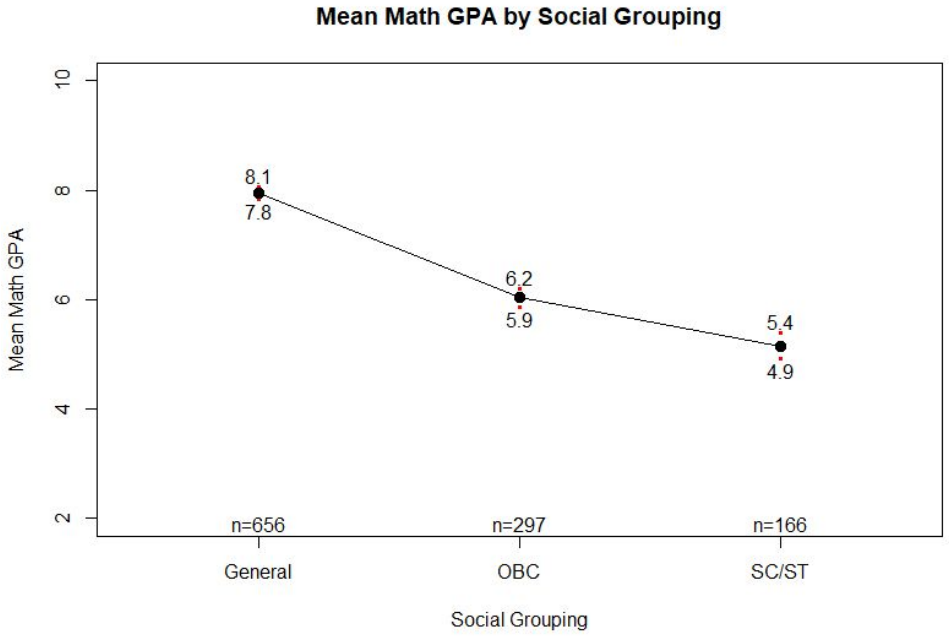
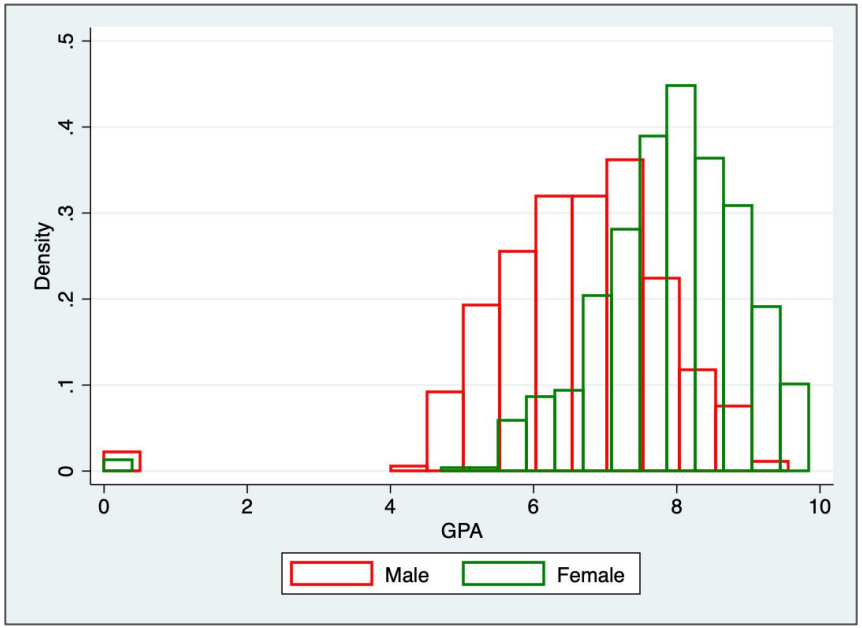


Table 2: Mean Math GPA by Social Groupings. General is base category.

	Dependent variable:
	mth_gp
caste_broadOBC	−1.905***
	(0.102)
caste_broadSC/ST	−2.787***
	(0.126)
Constant	7.944***
	(0.057)
Observations	1,119
R^2	0.377
Adjusted R^2	0.375
Residual Std. Error	1.453 (df = 1116)
F Statistic	336.966*** (df = 2; 1116)
Note:	*p<0.1; **p<0.05; ***p<0.01

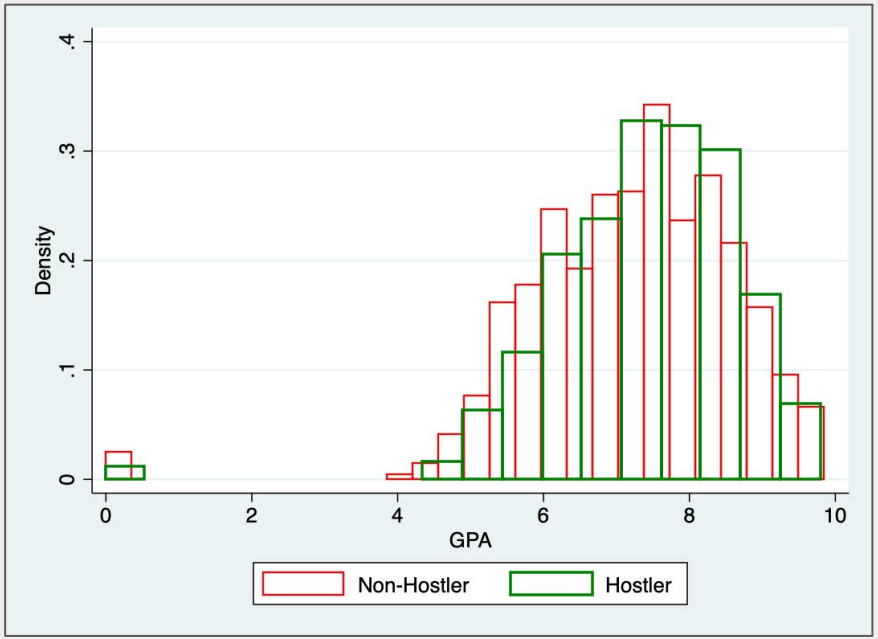
GPA trends by Gender

- Females have higher GPAs than males.



GPA trends by Hostel

- Hostlers have higher GPAs than non-hostlers.



Defining the Problem

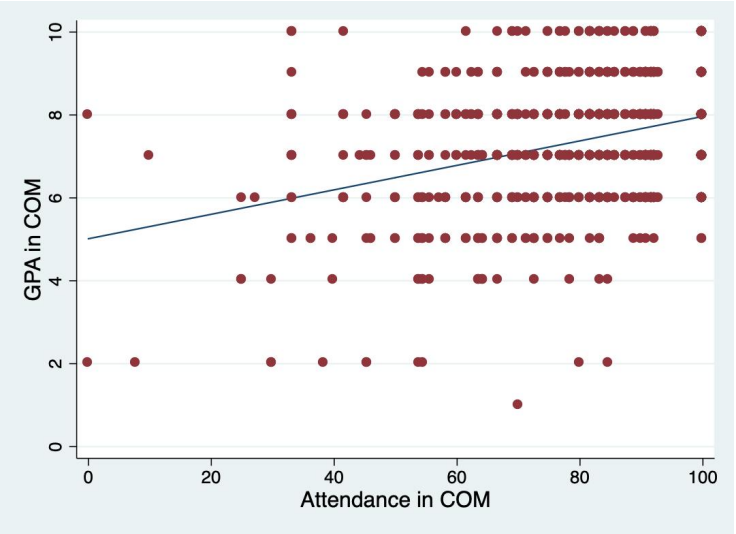
What is the impact of in-class attendance on course-specific GPAs at IIIT-Delhi?

- Do people attending more classes score higher?
  - Is attendance equally important for different courses?
- What are the other drivers of GPA?
  - Demographics
    - Caste Category
    - Gender
  - Hostellers vs. Day-scholars (peer-group effects)
  - Institutional Factors
    - Bonus
    - Fee-Waiver (indicative of income status)
    - Course offered in Winter/Monsoon

Correlation to Causality

Observation: Positive Correlation between

GPA and Attendance



But does higher attendance CAUSE higher GPA? → **Econometrics**

Exogeneity Assumption

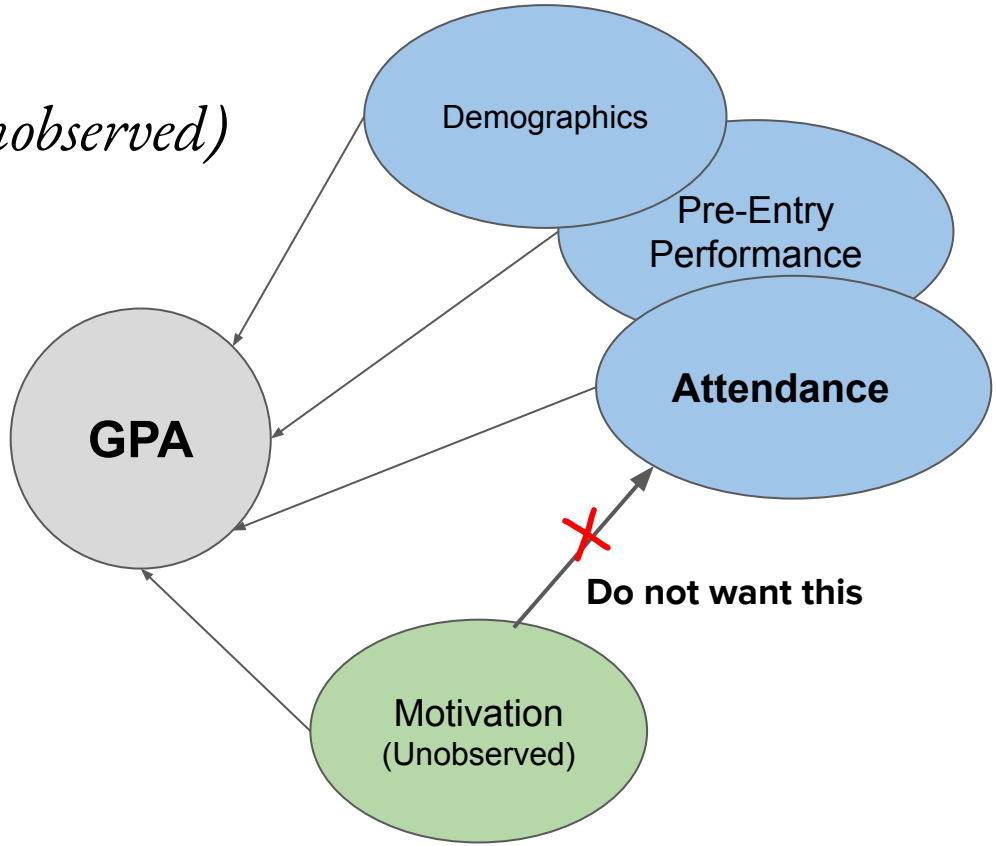
$$GPA = f(\textbf{\textit{Attendance}}, JEE, \ldots, Unobserved)$$

For Causality:

$$E[Unobserved \mid Attendance] = 0$$

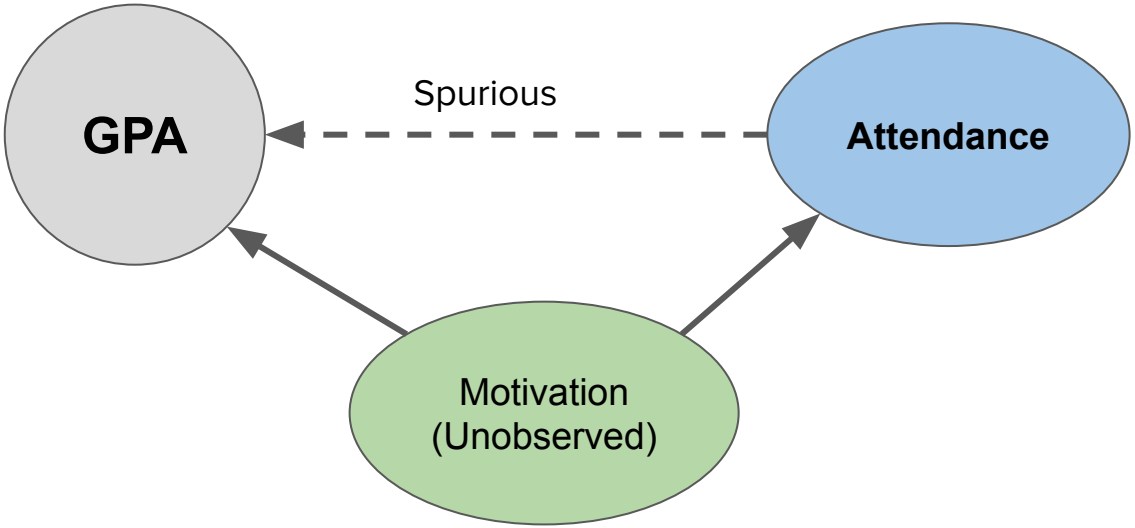
Blue and green are Uncorrelated.

Results are unbiased



Endogeneity

An **unobserved (hidden) factor** that simultaneously impacts both GPA and attendance **making it appear** that attendance causes higher GPA.



Attendance as a “Decision”

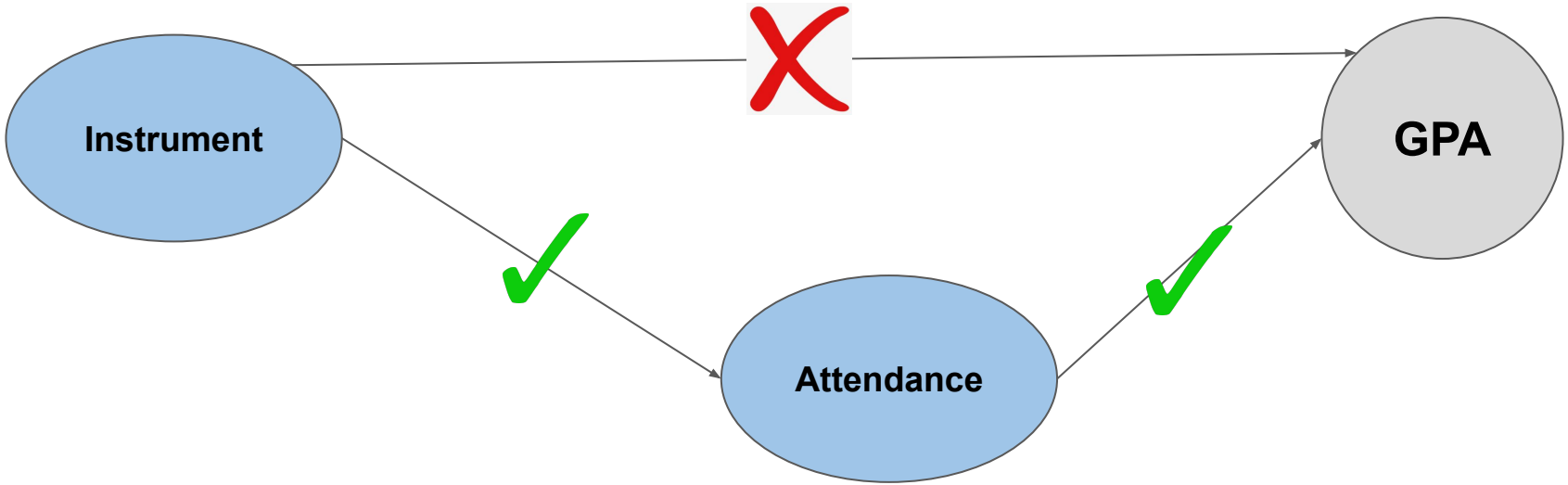
- Attendance is understood as a **decision.**
- Attendance percentage is an **accumulation of daily attendance decisions.**

What drives attendance?

- Willingness
- Attendance-related grading component
- Course instructor
- Ability
- Time of offering (Morning/Evening, Summer/Winter)
- Hosteller or Day Scholar
- Participation in extracurriculars
- AND MORE

Instrument Variable Approach

- To counter the endogeneity problem, we wish to find a variable (z) such that
  - z is CORRELATED with Attendance.
  - z does NOT affect grade point DIRECTLY, but THROUGH attendance.



- Allows to generate **unbiased** estimates of attendance.

Selecting the Instrument

- Hostel
  - Affects grade point directly through peer-learning
- Surprise Quiz
  - Is a component of grade point
- JEE Rank
  - Can be proxy of ability as it is the entry criteria
- Twelfth Percentage (XII %)
  - Pre-entry performance and not criteria
  - Strong correlation with attendance (Students scoring additional percent in class 12th attend more classes (~1%***) for all course groups.)
- Fee Waiver
  - Affects grade point only through attendance.

2 Stage Least Squares (2SLS) Model

First Stage ➔ Modelling Attendance

$$Att_{i,c(t)} = \beta_0 + \beta_1 Fee\_Waiver_{i,c(t)} + \beta_2 XII\%_i + \beta_3 Cat\_SC.ST_i + \beta_4 Cat\_OBC_i + \beta_5 Hostel_{i,c(t)} + \beta_6 Male_i + \beta_7 JEERank_i + \beta_8 Bonus_i + \beta_9 Trend_{c(t)} + u_{i,c(t)}$$

β_k 's are the parameters to be estimated. Important to have β_2 as significant.

$Att_{i,c(t)}$ represents the attendance percentage of a given **individual** in a certain **course** offered in a certain **year.**

c = {IP, DC, DSA, M1, PS, COM}

t = {2017, 2018, 2019}

2 Stage Least Squares (2SLS) Model

Second Stage ➔ Modelling GPA

$$GPA_{i,c(t)} = \lambda_0 + \lambda_1 \hat{Att}_{i,c(t)} + \lambda_2 Cat\_SC.ST_i + \lambda_3 Cat\_OBC_i + \lambda_4 Hostel_{i,c(t)} + \lambda_5 Male_i + \lambda_6 JEERank_i + \lambda_7 Bonus_i + \lambda_8 Trend_{c(t)} + \nu_{i,c(t)}$$

We will be using **unbiased estimates** of attendance as generated from the first stage

Aim is to generate unbiased estimates of λ_1

A one-percent change in attendance causes λ_1 change in GPA

Results

Modelling Attendance in CS Courses:

Explanatory Variables	Coefficient (SE)
Social Group (SC/ST = 1)	-5.0 * (2.6)
Social Group (OBC = 1)	-3.8 * (1.6)
Hostel Accommodation (Yes = 1)	-7.7*** (1.6)
Gender (Male = 1)	-1.9 (1.3)
JEE Rank / 1000	0.03** (0.01)
Bonus Marks	1.0 (1.0)
Trend	1.7 (1.9)
Twelfth Percentage (Instrument)	0.9*** (0.2)
Fee Waiver (Instrument)	1.5 (2.1)
Constant	-7.5 (15.0)

N = 1022 | *: p < 0.10 , **: p < 0.05, ***: p < 0.01

STAGE 1

Modelling GPA in CS (With Attendance Instrumented):

Independent Variables	Coefficient (SE)
Attendance in CS	0.1*** (0.0)
Caste = SC/ST	-0.2 (0.2)
Caste = OBC	-0.6*** (0.1)
Hostel	0.9*** (0.1)
Bonus Marks Awarded	0.5*** (0.0)
Male	0.1 (0.2)
JEE Rank / 1000	-0.01*** (0.00)
Trend	0.0 (0.0)
Constant	-1.0 (0.9)

N = 1022 | *: p < 0.10 , **: p < 0.05, ***: p < 0.01

STAGE 2

Hostellers

- Hostellers' **attendance is lower** by ~8% as compared to their day-scholar counter parts.
- But, their **GPA is higher** by 0.8 points.
- Strong peer group effect (?)
- True for Maths and CS courses.

Social Groups (SC/ST)

1. **Computer Science:**

Lower GPA outcome : **Caused by lower attendance (by 1-10%*)** + **Other factors**

2. **Maths:**

(UNRELATED to this social grouping)

Lower GPA outcome: **Not caused by attendance.**

3. **Communication Skills:**

Lower GPA outcome: **Caused by lower attendance (by 5-15%*)** + **Other factors**

Social Groups (OBC)

Across all course groups:

Lower GPA outcome: **Caused by lower attendance (by 0.5-7%*)** + **Other factors**

(RELATED to this social grouping)

Trend over the years:

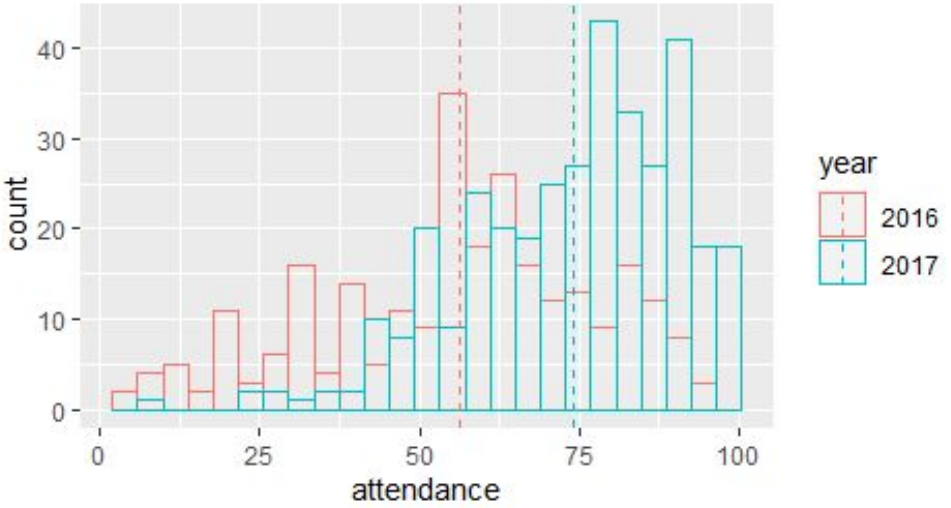
1. On average, GPA for Maths courses increased by 0.22*** grade points every year.
2. For Communication Skills, the GPA has fallen by 0.25* grade points every year.

There are no significant trends in attendance.

Winter v/s Monsoon -- DC

- DC offered in Winter for 2016 batch
- But in Monsoon for 2017 batch

YEAR	MEAN
2016	56.15
2017	74.07



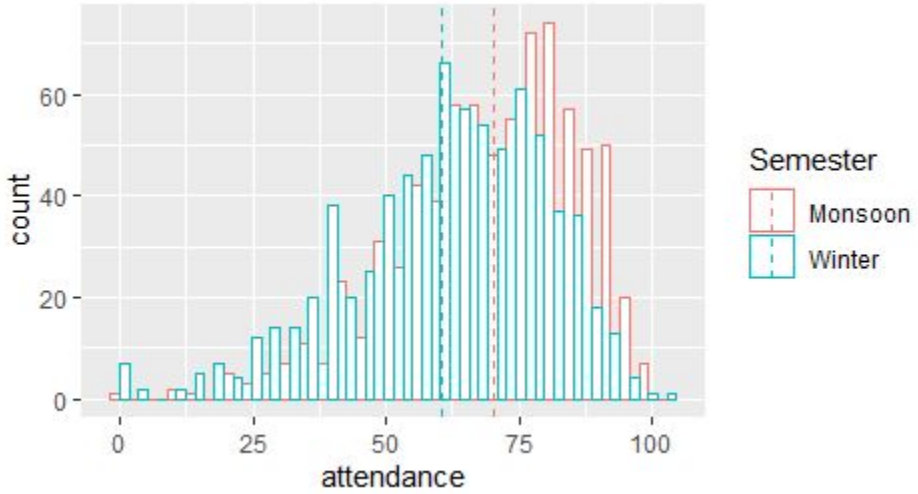
*Mean values are statistically different

p-value < 2.2e-16

Winter vs Monsoon -- 2017 & 2018

- Monsoon Courses considered
  - IP, DC & M-1
- Winter Courses considered
  - DSA & PS

SEMESTER	MEAN
Monsoon	70.25
Winter	60.53



*Mean values are statistically different

p-value < 2.2e-16

Conclusion

Attendance is important but it is only the first step to inclusive learning.

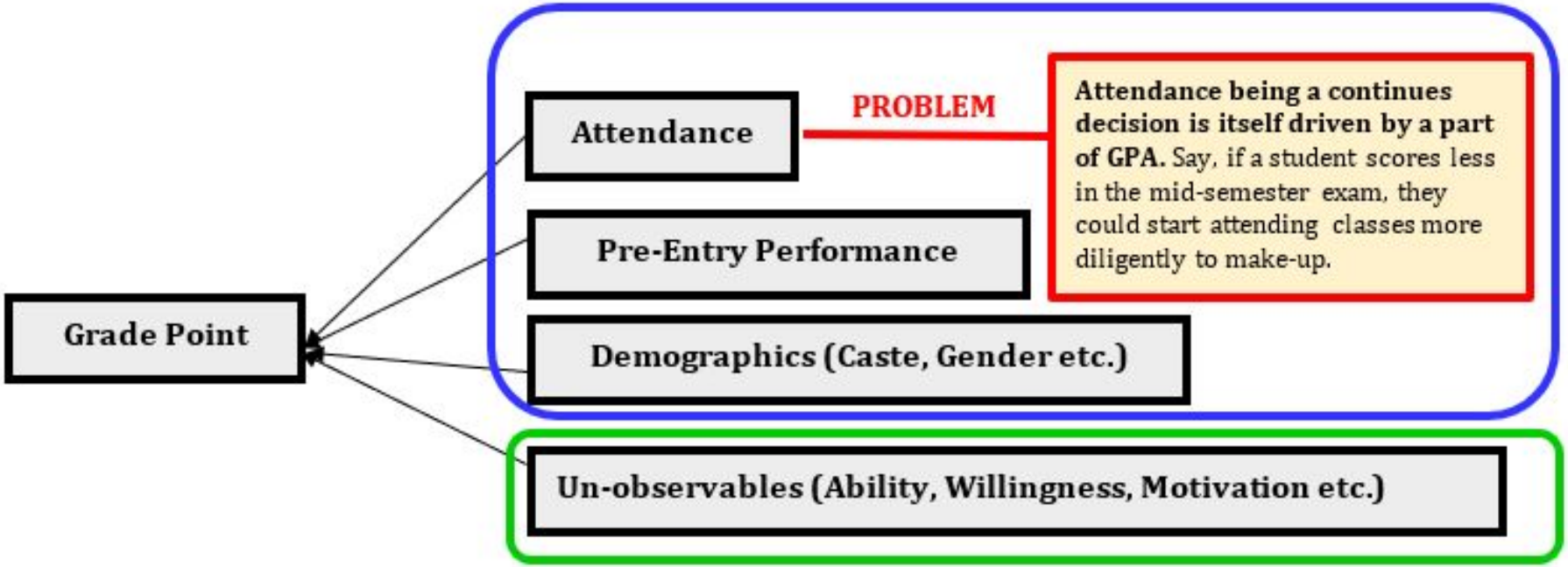
Evidence for positive peer learning effects.

Mild evidence for winter and summer timings affecting attendance.

Thank you!

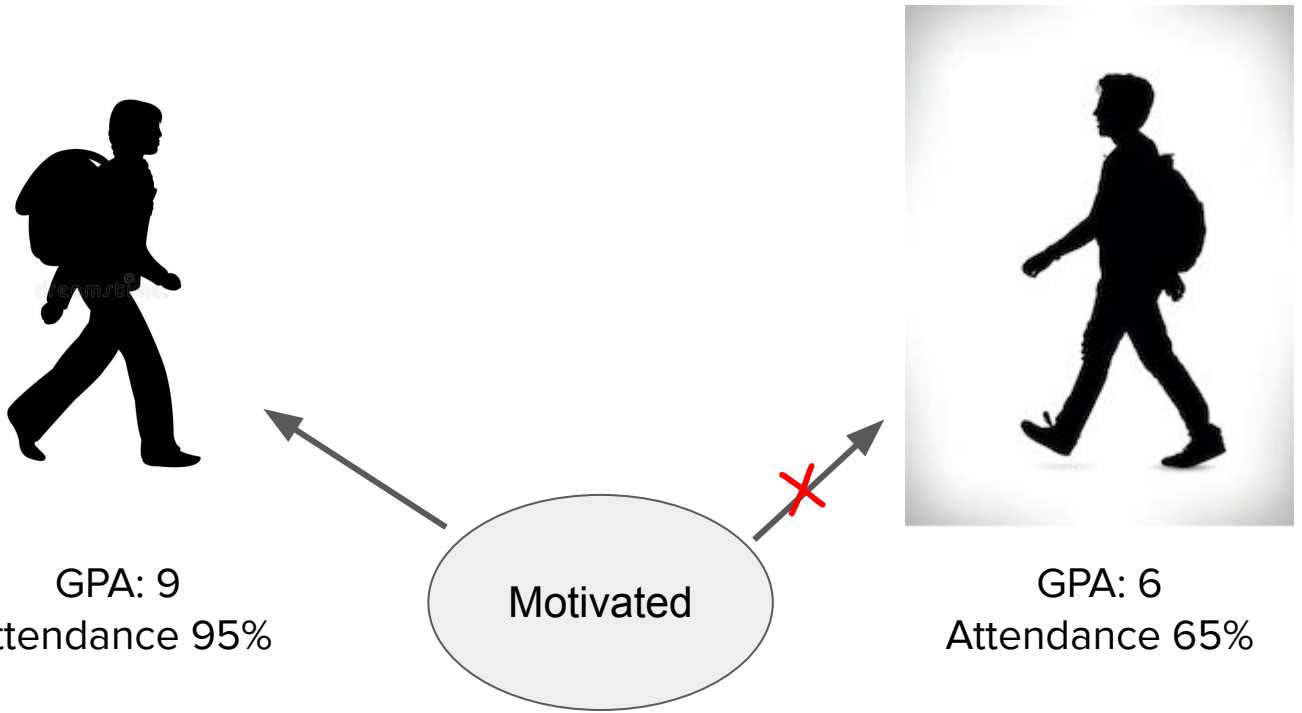
Extra Slides

Variable	Computer Science	Maths	Communication Skills
Hostellers	Lower attendance (**stage 1**); Higher GPA (**stage 2**)	Lower attendance ; Higher GPA	No impact on attendance or GPA.
Gender (Males = 1)	No impact on attendance or GPA.	No impact on attendance; Lower GPA.	No impact on attendance ; Lower GPA
Pre-entry Group (SC/ST)	Lower Attendance; No impact on GPA.	No impact on attendance; Lower GPA	Lower Attendance; No impact on GPA.
Pre-entry Group (OBC)	Lower Attendance; Lower GPA	Lower Attendance; Lower GPA	Lower Attendance; Lower GPA
Pre-Entry Criteria (Rank)	No impact on attendance ; Positive impact on GPA	No impact on attendance ; Positive impact on GPA	No impact on attendance ; Positive impact on GPA
Pre-Entry Criteria (Bonus)	No impact on attendance ; Positive impact on GPA	Positive impact on attendance and GPA.	No impact on attendance ; Positive impact on GPA
Pre-Entry Criteria (XIIth %)	Positive impact on attendance.	Positive impact on attendance.	Positive impact on attendance.
Pre-Entry Criteria (Waiver)	No impact on attendance	No impact on attendance	No impact on attendance



Generally, we assume that the blue and green portions are uncorrelated. This is crucial for the estimates to be unbiased, i.e. on – average be equal to the true parameters. But, given that attendance may be driven by GPA, it would also be related to factors like motivation. So, the correlation between the blue and green is no longer zero.

Endogeneity



Independent Variables	Computer Science GPA (Attendance Instrumented)
	N=1022
Attendance in CS	**0.11***** (0.01)
Caste = SC/ST	-0.20 (0.22)
Caste = OBC	**-0.56***** **(0.11)**
Hostel	**0.94***** (0.11)
Bonus Marks Awarded	**0.45***** (0.04)
Male	0.10 (0.16)
JEE Rank	**-1 x 10^(-5)***** **(2.3 x 10^(-6))**
Trend	-0.02 (0.02)
Constant	-1.02 (0.88)

Independent Variables	Mathematics GPA (Attendance Instrumented)
	N=1022
Attendance in MTH	**0.09***** (0.01)
Caste = SC/ST	**-0.89*** (0.37)
Caste = OBC	**-0.72***** (0.20)
Hostel	**0.79***** (0.06)
Bonus Marks Awarded	**0.06 .** (0.03)
Male	**-0.12 .** (0.07)
JEE Rank	**-1 x 10^(-5) **** (3 x 10^(-6))
Trend	**0.22***** (0.02)
Constant	0.35 (1.10)

Independent Variables	Communication GPA (Attendance Instrumented)
	N=1118
Attendance in COM	**0.07***** (0.01)
Caste = SC/ST	0.08 (0.30)
Caste = OBC	**-0.31*** (0.13)
Hostel	0.14 (0.25)
Bonus Marks Awarded	**0.29***** (0.05)
Male	**-0.58***** (0.01)
JEE Rank	**-3 x 10^(-6)*** **(1.5 x 10^(-6))**
Trend	**-0.25 .** (0.13)
Constant	**-2.27 .** **(1.35)**

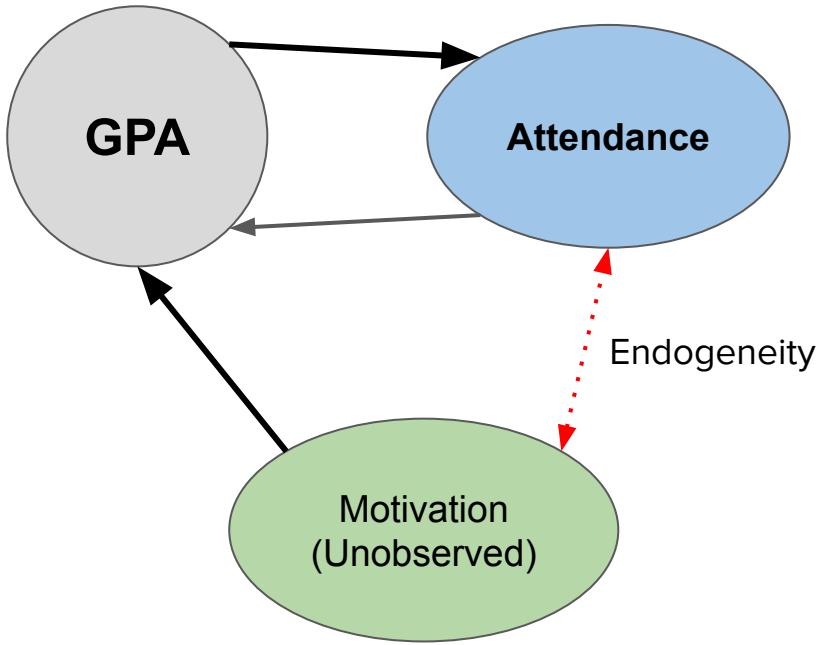
Endogeneity

Circular Causality $\Rightarrow$ Endogeneity

Temporal element to attendance.

I did well in the mid-sem so I attended fewer classes after that. GPA impacts Attendance

I attended fewer classes post mid-sem and did badly in the final. Attendance -> GPA



Pre-Entry Performance (Class XIIth) as an IV:

Students scoring an additional percent in class XIIth also attend more classes (by ~1%*) for all course groups.**

2 Stage Least Squares (2SLS) Model

- **Stage 1 → Modelling Attendance**

$$Att_{i,c(t)} = \beta_0 + \beta_1 Fee\_Waiver_{i,c(t)} + \beta_2 XII\%_i + \beta_3 Caste\_SC.ST_i + \beta_4 Caste\_OBC_i + \beta_5 Hostel_{i,c(t)} + \beta_6 Gender\_Male_i + \beta_7 JEERank_i + \beta_8 Bonus_i + \beta_9 Trend_{c(t)} + u_{i,c(t)}$$

- t = { 2017,2018,2019 }
- c = {IP, DC, DSA, M1, PS, COM}

- **Stage 2 → Modelling GPA**

$$GPA_{i,c(t)} = \beta_0 + \beta_1 Att_{i,c(t)} + \beta_2 Caste\_SC.ST_i + \beta_3 Caste\_OBC_i + \beta_4 Hostel_{i,c(t)} + \beta_5 Gender\_Male_i + \beta_6 JEERank_i + \beta_7 Bonus_i + \beta_8 Trend_{c(t)} + u_{i,c(t)}$$

- Attendance used in stage 2 is unbiased as estimated from stage 1

ATTENDANCE MODELLING			
Dependent Variable = Attendance	**Computer Science**	**Mathematics**	**Communication Skills**
	N=1022 || $R^2 = 0.15$	N=1022 || $R^2 = 0.12$	N=1022 || $R^2 = 0.07$
Explanatory Variables	**Coefficient (SE)**	**Coefficient (SE)**	**Coefficient (SE)**
Caste = SC/ST	**-5.0 *** **(2.6)**	-3.7 (2.7)	**-9.9*** ** **(2.6)**
Caste = OBC	**-3.8 *** **(1.6)**	**-3.7** ** **(1.7)**	**-3.4** ** **(1.6)**
Hostel Accommodation (Yes = 1)	**-7.7*** ** **(1.6)**	**-7.5*** ** **(1.4)**	1.0 (1.6)
Male = 1	-1.9 (1.3)	-1.2 (1.4)	-0.5 (1.3)
JEE Rank / 1000	**0.03** ** **(0.01)**	**0.03 *** **(0.01)**	**0.03 *** **(0.01)**
Bonus Marks	1.0 (1.0)	**2.6*** ** **(0.9)**	0.6 (1.4)
Trend	**1.7 *** **(0.9)**	0.1 (0.9)	0.0 (1.0)
Twelfth Percentage (Instrument)	**0.9*** ** **(0.2)**	**0.9*** ** **(0.2)**	**0.7*** ** **(0.2)**
Fee Waiver (Instrument)	1.5 (2.1)	0.2 (2.5)	-5.0 (3.6)
Constant	-7.5 (15.0)	-8.5 (15.0)	20.1 (16.2)

Significance codes: 0 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1

Independent Variables	Computer Science GPA (Attendance Instrumented)	Mathematics GPA (Attendance Instrumented)	Communication GPA (Attendance Instrumented)
	N=1022	N=1022	N=1118
Attendance in CS	**0.1*** (0.0)	**0.1*** (0.0)	**0.1*** (0.0)
Caste = SC/ST	-0.2 (0.2)	**-0.9**** (0.4)	0.1 (0.3)
Caste = OBC	**-0.6*** (0.1)**	**-0.7*** (0.2)**	**-0.3**** (0.1)
Hostel	**0.9*** (0.1)**	**0.8*** (0.1)**	0.1 (0.3)
Bonus Marks Awarded	**0.5*** (0.0)**	**0.1*** (0.0)**	**0.3*** (0.2)**
Male	0.1 (0.2)	**-0.1*** (0.1)**	**-0.6*** (0.0)**
JEE Rank / 1000	**-0.01*** (0.00)**	**-0.01*** (0.00)**	**0.00* (0.00)**
Trend	0.0 (0.0)	**0.20*** (0.0)**	**-0.3*** (0.1)**
Constant	-1.0 (0.9)	0.4 (1.1)	**-2.3* (1.4)**

Significance codes: 0 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY DELHI
(State University established by Govt. of NCT of Delhi)

IIITD/IRB/9/24/2020-4

24/September/2020
Dr. Gaurav Arora
Indraprastha Institute of Information Technology, Delhi

IIITD INSTITUTIONAL REVIEW BOARD APPROVAL

Project Title: **An impact evaluation of attendance and plagiarism policies on students' learning outcomes: Evidence from IIIT-Delhi**

We refer to your application for ethics approval with respect to the above project. The Board has deliberated on your request and reviewed your documents. Your study aims to quantify the causal impact of the attendance and plagiarism policies at IIIT-Delhi that are aimed at enhancing student's learning outcomes.

The documents reviewed are
a) Filled IRB form in the format needed.
b) Project description

The Board is satisfied with the bioethical consideration for the project and approves the ethics application under Expedited review. The approval period is from 24 September 2020 to 24 September 2021.

The following protocol and compliances are to be observed upon IIITD IRB approval.
1) The data provided by academic deprtment should be anonymised before it is provided to PI or project team.
2) The project should adher to the aims mentioned in the application and the data available from academic department should not be used to evaluate any student or faculty.

Dr Vibhor Kumar
Chair, Institutional Review Board
Indraprastha Institute of Information Technology, Delhi