14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Education Quality

Esther Duflo

14.771

School quality in Developing Countries

- There has been rapid improvement in school enrollment in developing countries over the last 10-15 years.
- However these improvements have not been matched by improvement in school quality:
  - Low learning performance (ASER study in India)
  - Massive Teacher absence (Chaudhury and other: 24% in India)
- Education quality has been an extremely active domain of research, and in particular there are a series of randomized evaluation paper on various issues:
  - "Production function" issues: class size, textbooks, flipcharts, etc.
  - Incentives for students, parents, and teachers
  - School systems:
    - Pedagogy (curriculum etc.)
    - Para-teachers vs regular teachers
    - Parent information/mobilization (report cards, school commitees etc.)

Duflo, Hanna, Ryan: Incentives for Para-teachers

- In India, regular teachers have essentially no incentives (tenure, no increase in salary)
- Para-teachers and incentives
  - It should be easier to provide them with good incentives
  - However, in India, they are no more likely to be present
  - Could be because they are actually not provided with incentives
- Motivating questions for this paper:
  - Can an incentive programs for para-teachers increase their presence?
  - Would increase presence lead to increase in learning or would it be undermined by:
    - Multitasking
    - Loss in intrinsic motivation
    - Incompetence

What the paper does

1. A randomized Experiment in teacher incentives
2. A regression discontinuity Design scheme to interpret the results: We estimate the change in teacher behavior just before and just after the end of a month, and this suggests that they respond to financial incentives
3. Use the treatment group to estimate a structural model; The non-linear nature of the attendance rules allows for estimation of a simple dynamic labor supply model, where teacher chooses every day between going to school or staying home and getting an outside option

The Context

- We worked with Seva Mandir, an NGO in rural Rajasthan
- They run 150 “non-formal education center” (NFE): single teacher school for students who do not attend regular school.
- Students are 7-14 year old, completely illiterate when they join.
- Schools teach basic hindi and math skills and prepare students to “graduate” to primary school.
- In 1997, 20 million children were served by such NFEs

The Intervention

- Teacher in Intervention school were provided with a camera with non-temperable time and date stamp

A picture

Photograph of children in school removed due to copyright restrictions.

The Intervention

- Teacher in Intervention school were provided with a camera with non-temperable time and date stamp
- Instructed to take a picture of themselves and the children every day (morning and afternoon). A valid pairs of picture has:
  - Two pictures taken the same day, separated by at least 5 hours each.
  - At least 8 children per picture
- Payment is calculated each month and is a non-linear function of attendance:
  - Up to 10 days: Rs 500.
  - Each day above 10 days: Rs 50.
- In non-intervention schools, teachers receive Rs 1000, and are reminded by attending at least 20 days is compulsory.

The Evaluation

- We originally picked 120 schools, out of which 7 closed immediately after they were picked to be in the study (unrelated to the study).
- 57 treatment schools, the rest control.
- Data collection:
  - Teacher and student attendance: Monthly random checks.
  - In treatment schools: Camera data
  - Students learning: tests in September 03-April 04-Oct 04
  - Long term impact: a new sets of random checks was done in 2006-2007, and a new set of test scores were done in 2007

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Power

- We know that $E[Y_i(0)|W_i = 1] = [Y_i(0)|W_i = 0]$
- But in a finite sample, it may or may not hold.
- Size (level) of a test (e.g. test H_0 ATE=0): Probability of a type I error: I reject H_0 when H_0 is true
- Generally we set the size at 5%.
- Power of a test: 1-probability of type II error.
- Type II error: for a given size, I do not reject 0, when I should have.
- Power depend on effect of program, and on precision of the estimate:
  - Sample size
  - Level of Randomization: If I randomize at the group level, I need to cluster at this group level: need to adjust power calculation for that (it will depend on size of the group, and expected correlation of outcomes within the group).

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Checking the Balance in the Camera Experiment

Table 1: Is School Quality Similar in Treatment and Control Groups Prior to Program?

	Treatment	Control	Difference
	(1)	(2)	(3)
A. Teacher Attendance			
School Open	0.66	0.64	0.02
			(0.11)
	41	39	80
B. Student Participation (Random Check)			
Number of Students Present	17.71	15.92	1.78
			(2.31)
	27	25	52
C. Teacher Qualifications			
Teacher Test Scores	34.99	33.62	1.37
			(2.01)
	53	56	109
Teacher Highest Grade Completed	10.21	9.80	0.41
			(0.46)
	57	54	111

School quality

Table 1: Is School Quality Similar in Treatment and Control Groups Prior to Program?

	Treatment (1)	Control (2)	Difference (3)
D. Teacher Performance Measures (Random Check)			
Percentage of Children Sitting Within Classroom	0.83	0.84	0.00
			(0.09)
	27	25	52
Percent of Teachers Interacting with Students	0.78	0.72	0.06
			(0.12)
	27	25	52
Blackboards Utilized	0.85	0.89	-0.04
			(0.11)
	20	19	39
E. School Infrastructure			
Infrastructure Index	3.39	3.20	0.19
			(0.30)
	57	55	112
Fstat(1,110)			1.21
p-value			(0.27)

Students

Table 2: Are Students Similar Prior To Program?

	Levels			Normalized by Control		
	Treatment	Control	Difference	Treatment	Control	Differenc
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Can the Child Write?					
Took Written Exam	0.17	0.19	-0.02			
			(0.04)			
	1136	1094	2230			
	B. Oral Exam					
Math Score on Oral Exam	7.82	8.12	-0.30	-0.10	0.00	-0.10
			(0.27)			(0.09)
	940	888	1828	940	888	1828
Language Score on Oral Exam	3.63	3.74	-0.10	-0.03	0.00	-0.03
			(0.30)			(0.08)
	940	888	1828	940	888	1828
Total Score on Oral Exam	11.44	11.95	-0.51	-0.08	0.00	-0.08
			(0.48)			(0.07)
	940	888	1828	940	888	1828
	C. Written Exam					
Math Score on Written Exam	8.62	7.98	0.64	0.23	0.00	0.23
			(0.51)			(0.18)
	196	206	402	196	206	402
Language Score on Written Exam	3.62	3.44	0.18	0.08	0.00	0.08
			(0.46)			(0.20)
	196	206	402	196	206	402
Total Score on Written Exam	12.17	11.41	0.76	0.16	0.00	0.16
			(0.90)			(0.19)
	196	206	402	196	206	402

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Attrition

- At the school level: some schools got lost, for reasons not related to the program
- At the individual level for the test: we have substantial attrition
  - Why is that a potential problem?
  - When will it be a problem?
  - What should we check?
    - percentage attrition is not differential by group
    - observable characteristics of attritors are no different in T and C group
  - If not what can we do?
    - Assume a selection process, and correct for it (we lose main advantage of a random sample)
    - Provide bounds

Attrition

Table 9: Descriptive Statistics for Mid Test and Post Test

	Mid Test			Post Test		
	Treatment	Control	Difference	Treatment	Control	Difference
	A. Attrition Process					
Percent Attrition	0.11	0.22	-0.10	0.24	0.21	0.03
			(0.05)			(0.04)
Difference in Percent Written of Pre-Test attriters-stayers	0.01	0.03	0.02	0.06	-0.03	0.10
			(0.06)			(0.06)
Difference in Verbal Test of Pre-Test attriters-stayers	0.05	0.08	-0.03	0.02	0.12	-0.10
			(0.14)			(0.14)
Difference in Written Test of Pre-Test attriters-stayers	-0.41	-0.23	-0.18	-0.19	-0.13	-0.06
			(0.34)			(0.29)
	B. Exam Score Means					
Took Written	0.36	0.33	0.03	0.61	0.57	0.04
			(0.04)			(0.05)
Math	0.14	0.00	0.14	-0.08	-0.24	0.16
			(0.10)			(0.15)
Language	0.14	0.00	0.14	1.71	1.60	0.11
			(0.10)			(0.11)
Total	0.14	0.00	0.14	0.35	0.24	0.12
			(0.10)			(0.11)

The Randomized evaluation Checklist

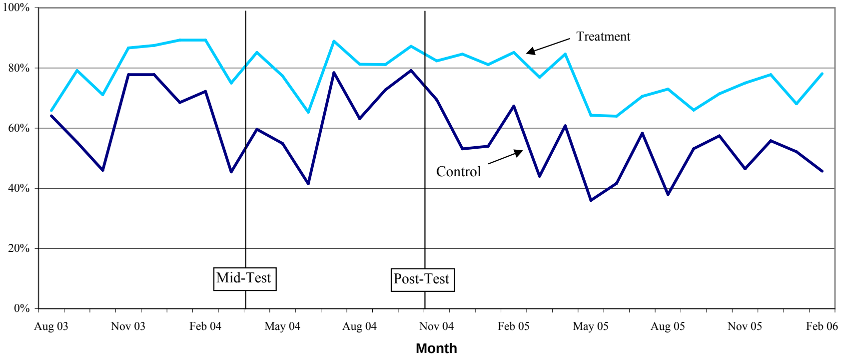
1. What was the power of the Experiment?
    - At what level was the experiment randomized?
    - We need to take into account clustering at that level in computing our standard error
    - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
    - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
    - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
    - If so, how did we deal with it?
4. Did we have non-compliance?
    - If so how did we deal with it? (next lecture)
5. Did we have contagion (externalities) between treatment and control group?

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Attendance: Graphical Evidence

Figure 2: Percentage of Schools Open during Random Checks



Attendance: tables

Table 3: Teacher Attendance

Sept 2003-Feb 2006			Difference Between Treatment and Control Schools		
Treatment	Control	Diff	Until Mid-Test	Mid to Post Test	After Post Test
(1)	(2)	(3)	(4)	(5)	(6)
A. All Teachers					
0.79	0.58	0.21	0.20	0.20	0.23
		(0.03)	(0.04)	(0.04)	(0.04)
1575	1496	3071	882	660	1529
B. Teachers with Above Median Test Scores					
0.78	0.63	0.15	0.15	0.15	0.14
		(0.04)	(0.05)	(0.05)	(0.06)
843	702	1545	423	327	795
C. Teachers with Below Median Test Scores					
0.78	0.53	0.24	0.21	0.14	0.32
		(0.04)	(0.05)	(0.06)	(0.06)
625	757	1382	412	300	670

Cheating?

Table 4: Comparing Random Checks to Photo Data for Treatment Schools

Scenario	Number	Percent of Total
A. Possible Scenarios		
School Open and Valid Photos	879	66%
School Open and Invalid Photos	179	13%
School Closed and Valid Photos	88	7%
School Closed and Invalid Photos	191	14%
B. Out of 179 where School is Open, the photos are invalid because….		
School not open for full 5 hours	43	24%
Only one photo	90	50%
Not enough Children	36	20%
Instructor not in Photo	9	5%
Don't Know	1	1%
C. Out of 88 where School is Closed and the photos are valid…..		
Random check completed after the school closed	13	15%
Camera broke/excused meeting	21	24%
Teacher left in the middle of the day	54	61%

No evidence of Multitasking

Table 7: Teacher Performance

	Sept 2003-Feb 2006			Difference Between Treatment and Control Schools		
	Treatment	Control	Diff	Until Mid-Test	Mid to Post Test	After Post Test
	(1)	(2)	(3)	(4)	(5)	(6)
Percent of Children Sitting Within	0.72	0.73	-0.01	0.01	0.04	-0.01
Classroom			(0.01)	(0.89)	(0.03)	(0.02)
	1239	867	2106	643	480	983
Percent of Teachers Interacting with	0.55	0.57	-0.02	-0.02	0.05	-0.04
Students			(0.02)	(0.04)	(0.05)	(0.03)
	1239	867	2106	643	480	983
Blackboards Utilized	0.92	0.93	-0.01	-256766.00	0.01	-0.01
			(0.01)	(0.02)	(0.02)	(0.02)
	990	708	1698	613	472	613

Notes: (1) Teacher Performance Measures from Random Checks only includes schools that were open during the random check. (2) Standard errors are clustered by school.

No increase on conditional attendance, more days worked

Table 8: Child Attendance

	Sept 03-Feb 06			Difference Between Treatment and Control Schools		
	Treatment	Control	Diff	Until Mid-Test	Mid to Post Test	After Post Test
	(1)	(2)	(3)	(4)	(5)	(6)
A. Attendance Conditional on School Open						
Attendance of Students Present at Pre-Test Exam	0.46	0.46	0.01	0.02	0.03	0.00
			(0.03)	(0.03)	(0.04)	(0.03)
	23495	16280	39775			
Attendance for Children who did not leave NFE	0.62	0.58	0.04	0.02	0.04	0.05
			(0.03)	(0.03)	(0.04)	(0.03)
	12956	10737	23693			
B. Total Instruction Time (Presence)						
Presence for Students Present at Pre-Test Exam	0.37	0.28	0.09	0.10	0.10	0.08
			(0.03)	(0.03)	(0.04)	(0.03)
	29489	26695	56184			
Presence for Student who did not leave NFE	0.50	0.36	0.13	0.10	0.13	0.15
			(0.03)	(0.04)	(0.05)	(0.04)
	16274	17247	33521			
C. Presence, by Student Learning Level at Program Start (for those who did not leave)						
Took Oral Pre-Test	0.50	0.36	0.14	0.11	0.14	0.15
			(0.03)	(0.03)	(0.05)	(0.04)
	14778	14335	29113			
Took Written Pre-Test	0.48	0.39	0.10	0.07	0.07	0.11
			(0.06)	(0.07)	(0.06)	(0.07)
	1496	2912	4408			

Notes: (1) Standard errors are clustered at the level of the school. (2) Child attendance data were collected during random checks. (3) The attendance at the pre-test exam determined the child enrollment at the start of the program.

Regression

$$\text{Score}_{ikj} = \beta_1 + \beta_2 \text{Treat}_j + \beta_3 \text{Pre\_Writ}_{ij} + \beta_4 \text{Pre\_oral}_{ij} + \beta_5 \text{Writ} + \epsilon_{ijk}$$

Test Score results

Table 10: Estimation of Treatment Effects for the Mid- and Post-Test

Mid-Test				Post-Test			
Took Written	Math	Lang	Total	Took Written	Math	Lang	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All Children							
0.04	0.15	0.16	0.17	0.06	0.21	0.16	0.17
(0.03)	(0.07)	(0.06)	(0.06)	(0.04)	(0.12)	(0.08)	(0.09)
1893	1893	1893	1893	1760	1760	1760	1760
B. With Controls							
0.02	0.13	0.13	0.14	0.05	0.17	0.13	0.15
(0.03)	(0.07)	(0.05)	(0.06)	(0.04)	(0.10)	(0.07)	(0.07)
1893	1893	1893	1893	1760	1760	1760	1760

Results by Pre-test score

Table 10: Estimation of Treatment Effects for the Mid- and Post-Test

Mid-Test			**Post-Test**			
Math	Lang	Total	Took Written	Math	Lang	Total
(2)	(3)	(4)	(5)	(6)	(7)	(8)
C. Took Pre-Test Oral						
0.14	0.13	0.15		0.2	0.13	0.16
(0.08)	(0.06)	(0.07)		(0.14)	(0.09)	(0.10)
1550	1550	1550		1454	1454	1454
D. Took Pre-Test Written						
0.19	0.28	0.25		0.28	0.28	0.25
(0.12)	(0.11)	(0.11)		(0.18)	(0.11)	(0.12)
343	343	343		306	306	306

Graduation to government school

Table 11: Dropouts and Movement into Government Schools

	Treatment (1)	Control (2)	Diff (3)
Child Left NFE	0.44	0.36	0.08
			(0.04)
Child Enrolled in Government School	0.26	0.16	0.10
			(0.03)
Child Dropped Out of School	0.18	0.20	-0.02
			(0.03)
N	1136	1061	2197

Estimating the impact of teacher absence

- Suppose we want to use this experiment to estimate the impact of teacher absence on test score?
- What would the strategy be?
    - Use Treatment dummy as instrument for teacher attendance
    - Wald estimate: divide effect of program on test score by effect of program on attendance
- What would the potential threat to validity of the strategy
- What do we think about the severity of this threat?

Estimating the impact of teacher absence

Table 12: Does the Random Check Predict Test Scores?

Method:	OLS	OLS	OLS	2SLS
Sample:	Control Schools	Treatment Schools	Treatment Schools	All Schools
Data:	Random Check	Random Check	Photographs	Random Check
	(1)	(2)	(3)	(4)
A. Mid-test (Sept 03-April 04)				
Took Written	0.02	0.28	0.36	0.26
	(0.10)	(0.08)	(0.11)	(0.19)
Total Score	0.20	0.39	0.87	1.07
	(0.19)	(0.21)	(0.22)	(0.43)
N	878	1015	1015	1893
B. Post-test (Sept 03 -Oct 04)				
Took Written	0.24	0.51	0.59	0.33
	(0.16)	(0.15)	(0.20)	(0.22)
Total Score	0.58	1.17	0.98	0.97
	(0.35)	(0.36)	(0.53)	(0.47)
N	883	877	877	1760

Monitoring or Incentives? Preliminary Evidence

- Are teachers sensitive to increased monitoring or to incentives?
- Preliminary evidence based on *Regression Discontinuity Design*
- Consider a case where treatment is assigned when the treatment is assigned based on a strict threshold:
  - Sharp RD: $W_i = 1[X_i > c]$
  - Fuzzy RD: $lim_{x \downarrow c} pr(W_i = 1|Xi = x) \neq lim_{x \uparrow c} pr(W_i = 1|Xi = x)$
- Identification assumption for RD: $lim_{x \downarrow c} E[Y_i(0)|Xi = x] = lim_{x \uparrow c} E[Y_i(0)|Xi = x]$
- Estimator: we try to approximate: $lim_{x \downarrow c} E[Y_i|Xi = x] - lim_{x \uparrow c} E[Y_i|Xi = x]$
  - In the sharp RD: this will be the treatment effect
  - In the fuzzy RD: we use the treshold as instrument: compute our friend the Wald estimate.

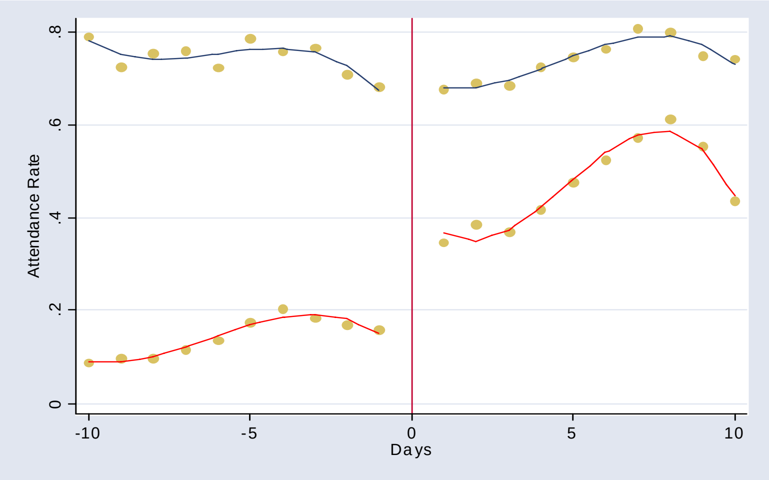
RD in the teacher case

- In practice: We try to estimate a smooth (non-parametric) function of the relationship between Y and X (here: day in the month and whether teacher works).
- We then use this to estimate the limits at the threshold, from the left and the right.
- When we switch from the last day of the month to the first day of the month:
  - A teacher who has attended 9 days or less in the rest of the month faces no incentive at the end of month t and faces incentives again at the end of month $t+1$.
  - A teacher who has attended more than 10 days in the rest of the month face a Rs 50 incentives at the end of month t and slightly smaller at the beginning of the next month
- Graphical Evidence
- Regression:

$$W_{itm} = \alpha + \beta 1_m(d > 10) + \gamma F + \lambda 1_m(d > 10) * F + \upsilon_i + \mu_m \epsilon_{is}, \tag{1}$$

Regression Discontinuity Design: Graphical Evidence

Figure 5: RDD Representation of Teacher Attendance at the Start and End of the Month



Regression Discontinuity Design: Regressions

Table 5 : Do Teachers Work More When They are "In the Money"?

	(1)	(2)	(3)	(4)
Beginning of Month	0.19	0.12	0.46	0.39
	(0.05)	(0.06)	(0.04)	(0.03)
In the Money	0.52	0.37	0.6	0.48
	(0.04)	(0.05)	(0.03)	(0.01)
Beginning of the Month * In the Money	-0.19	-0.12	-0.34	-0.3
	(0.06)	(0.06)	(0.04)	(0.02)
Observations	2813	2813	27501	27501
R-squared	0.06	0.22	0.08	0.16
Sample	1st and last day of month	1st and last day of month	1st 10 and last 10 days of month	1st 10 and last 10 days of month
Third Order Polynomial on Days on each side			X	X
Teacher Fixed Effects		X		X
Month Fixed Effects		X		X
Clustered Standard Errors	X		X	

The Model

- Each day, a teacher chooses whether or not to attend school, by comparing the value of attending school to that of staying home or doing something else.
- State space $s = (t, d)$, where t is the current time and d is the days worked previously in the current month.
- Payoffs:
  - If the teacher does not attend school: $\mu + \epsilon_t$
  - Payoff of attending school is calculated at the end of the month according to:

$$\pi(d) = 500 + \max\{0, d - 10\} \tag{2}$$

- T takes value between 1 and $T = 25$.
- Transitions: Each day, t increases by one, unless $t = T$, in which case it resets to $t = 1$. If a teacher has worked in that period d increases by one, otherwise it remains constant.

Value function

Given this payoff structure, for $t < T$, we can write the value function for each teacher as follows:

$$V(t,d) = \max\{\mu + \epsilon_t + EV(t+1,d), EV(t+1,d+1)\}. \quad (3)$$

At time T, we have:

$$V(T,d) = \max\{\mu + \epsilon_T + \beta\pi(d) + EV(1,0), \beta\pi(d+1) + EV(1,0)\}, \quad (4)$$

where β is marginal utility of income.
$EV(1,0)$ enters both side and can thus be ignored: we can solve each month independently, backwards from time T.

Identification

- Identification is constructive, and based on partitions of the state space.
- At time T, the agent faces a static decision; work if:

$$\mu + \epsilon_T + \beta\pi(d) > \beta\pi(d+1). \tag{5}$$

- The probability of this event is:

$$\begin{aligned} Pr(work|d,\theta) &= Pr(\epsilon_T > \beta(\pi(d+1) - \pi(d)) - \mu) & (6) \\ &= 1.0 - \Phi(\beta(\pi(d+1) - \pi(d)) - \mu), & (7) \end{aligned}$$

Identification with iid innovation in outside option

- When $d < 10$, the difference between $\pi(d+1)$ and $\pi(d)$ is zero, and β does not enter the equation.
- The resulting equation is:

$$Pr(work|d, \theta) = 1 - \Phi(\mu), \tag{8}$$

  which is a simple probit.
- If all teachers share same μ, μ is identified by teachers who are out of the money, and then β from teachers in the money.
- var(ϵ) normalized to be equal to 1.
- If teachers have different μ model still identified by comparing different teachers with themselves over time (teacher fixed effect).

Identification with AR(1) innovation in outside option

- If ϵ is serially correlated, identification is more complicated.
- Suppose that the shock follows an AR(1) process:

$$\epsilon_t = \rho \epsilon_{t-1} + \nu_t, \tag{9}$$

- ϵ_T will be correlated with d, as teachers with very high draws on ϵ_T are more likely to be in the region where $d < 10$ if ρ is positive (the converse will be true if ρ is negative).
- This will bias our estimates of μ and β.

iid model, with or without fixed effect

Simply write the empirical counterpart of the maximization problem.
The log likelihood is:

$$LLH(\theta) = \sum_{i=1}^{N} \sum_{m=1}^{M_i} \sum_{t=1}^{T_m} [1(\text{work}) Pr(\text{work}|t, d, \theta)$$

$$+1(\text{not work})(1 - Pr(\text{work}|t, d, \theta)],$$

where:

$$\begin{aligned} Pr(\text{work}|t, d, \theta) &= Pr(\mu + \epsilon_t + EV(t+1, d) < EV(t+1, d+1)) \\ &= Pr(\epsilon_t < EV(t+1, d+1) - EV(t+1, d) - \mu) \\ &= \Phi(EV(t+1, d+1) - EV(t+1, d) - \mu), \quad (10) \end{aligned}$$

Serial correlation

- Both estimation and identification are a little complicated...
- Use method of simulated moment: simulate work history for different parameters, and try to match a distribution of days worked at the beginning of the month.
- Can introduce heterogeneity by drawing p teacher from a distribution with high outside option, and $1 - p$ from distribution with low outside option.

Results from the structural Model

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
β	0.049	0.024	0.059	0.051	0.014	0.019
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
μ_1	1.55		2.315	2.063	-0.107	0.012
	(0.013)		(0.013)	(0.012)	(0.040)	(0.028)
ρ			0.682	0.547	0.461	
			(0.010)	(0.023)	(0.039)	
σ_1^2				0.001	0.153	0.135
				(0.011)	(0.053)	(0.027)
μ_2					3.616	1.165
					(0.194)	(0.101)
σ_2^2					0.26	0.311
					(0.045)	(0.051)
p					0.047	0.131
					(0.007)	(0.015)
Heterogeneity	None	FE	None	RC	RC	RC

Prediction on days worked (real=20.23 days)

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
Heterogeneity	None	FE	None	RC	RC	RC
$\in_{\text{Bonus}}$	3.52	1.687	6.225	10.08	0.306	0.370
	(1.550)	(0.098)	(0.634)	(1.249)	(0.038)	(0.029)
$\in_{\text{bonus\_cutoff}}$	-75.49	-16.04	-50.22	-63.11	-1.29	-1.78
	(6.506)	(1.264)	(2.612)	(3.395)	(0.479)	(0.449)
Predicted Days Worked	20.50	19.00	15.30	12.15	20.23	21.36
	(0.031)	(0.062)	(0.058)	(0.102)	(3.512)	(0.373)
Days Worked BONUS=0	1.60	6.02	1.29	1.318	13.55	11.81
	(0.597)	(0.234)	(0.875)	(0.863)	(5.251)	(0.669)
Out of Sample Prediction	26.16	18.886	15.08	12.956	20.86	21.57
	(0.059)	(0.253)	(0.635)	(0.520)	(3.793)	(0.456)

Distribution of Days worked

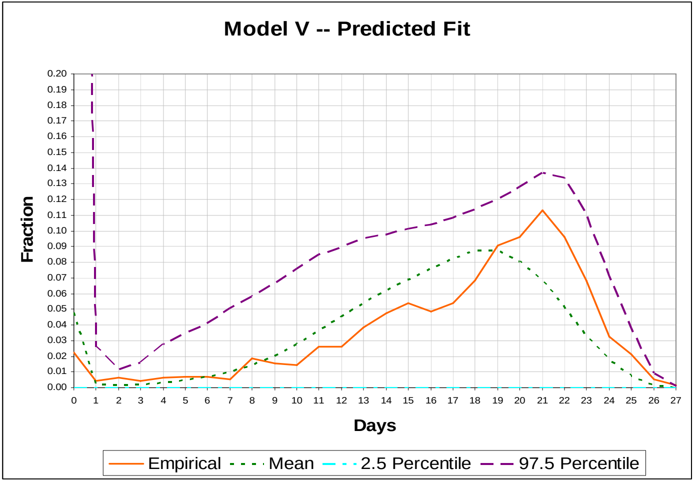


Figure 6B: CounterFactual Fit From Model V

Two out of sample tests

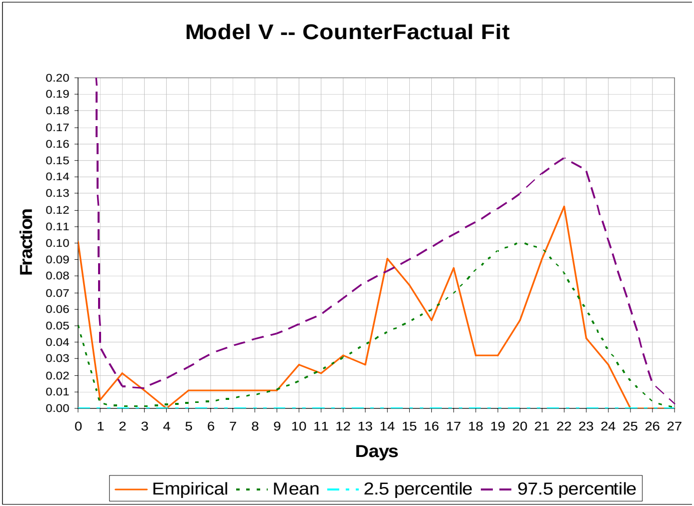
- Prediction of the number of days worked under no incentives
  - Model predicts that teachers would work 52% of the time in control group
  - In fact they work 58%
  - Predicted difference treatment vs control is 26%, vs 21% in reality
- The impact of a change in rule.
  - Seva Mandir changed rule after experiment was over (and model was estimated!)
  - New rule: Rs 700 for 12 days of work. Increment of Rs 70 after the 13th day
  - Model does well too.
- Note that all the alternative models do rather poorly in these counterfactuals.

Results from the structural Model

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
Heterogeneity	None	FE	None	RC	RC	RC
$\in_{\text{Bonus}}$	3.52	1.687	6.225	10.08	0.306	0.370
	(1.550)	(0.098)	(0.634)	(1.249)	(0.038)	(0.029)
$\in_{\text{bonus\_cutoff}}$	-75.49	-16.04	-50.22	-63.11	-1.29	-1.78
	(6.506)	(1.264)	(2.612)	(3.395)	(0.479)	(0.449)
Predicted Days Worked	20.50	19.00	15.30	12.15	20.23	21.36
	(0.031)	(0.062)	(0.058)	(0.102)	(3.512)	(0.373)
Days Worked BONUS=0	1.60	6.02	1.29	1.318	13.55	11.81
	(0.597)	(0.234)	(0.875)	(0.863)	(5.251)	(0.669)
Out of Sample Prediction	26.16	18.886	15.08	12.956	20.86	21.57
	(0.059)	(0.253)	(0.635)	(0.520)	(3.793)	(0.456)

Distribution of Days worked under new rule



Results from the structural model: Lessons

- A nice set up where we can corroborate assumptions of structural model.
- Other example: Todd and Wolpin (AER). They estimate a structural model in the control group and then validate it by predicting the Treatment Control difference.
- Model incorporating both serial correlation and heterogeneity does well, other models do poorly
- It seems that entire effect of program was through financial incentives.