Toulouse School of Economics

TSE 16th January 2019

Applied Econometrics for Development: Experiments II

Ana GAZMURI

Paul SEABRIGHT

The Cohen-Dupas bednets study

- The question: does subsidizing insecticide-treated anti-malarial bednets (ITNs) affect the take-up of such nets and their use by those who have acquired them?

- The background: many types of health intervention depend for their effectiveness not just on being administered to users but also on actions taken users (eg hanging and dehanging nets).

- ITNs are known to reduce malarial infection substantially both among users and among non-users in the vicinity of a concentration of users – malarial mortality can be reduced by over 20%

- But they require effort to use properly and use is low (estimated at 23% of children, 27% of pregnant women)

Why should subsidy levels affect usage?

- The price charged to users is likely to affect *take-up* (acquisition of nets), for familiar demand curve reasons

- But it could also affect the proportion of acquirers who use the nets
  - 1) Screening of users according to need (remember that nets have some value for non-health uses)
  - 2) Signaling effect of price on beliefs about value of the nets
  - 3) Sunk-cost effects

- Experimenters rule out type 2 effects on a priori grounds (widespread knowledge about ITNs effectiveness), and try to distinguish between 1) and 3) by use of a post-take-up surprise lottery

Experimental design

- 20 clinics chosen for study in Western Kenya
  - 4 as control group
  - 5 to provide ITNs free to pregnant women on first visit
  - 5 to provide them for 10 Ksh (15 US cents)
  - 3 to provide them for 20 Ksh (30 c)
  - 3 to provide them for 40 Ksh (60 c)
- Lotteries performed on random days in clinics with positive prices
- Eligible women were given an anemia test to measure objective need; total of 545 women. Sample of 246 women visited at home to establish ITN usage; 226 (92%) agreed to be interviewed

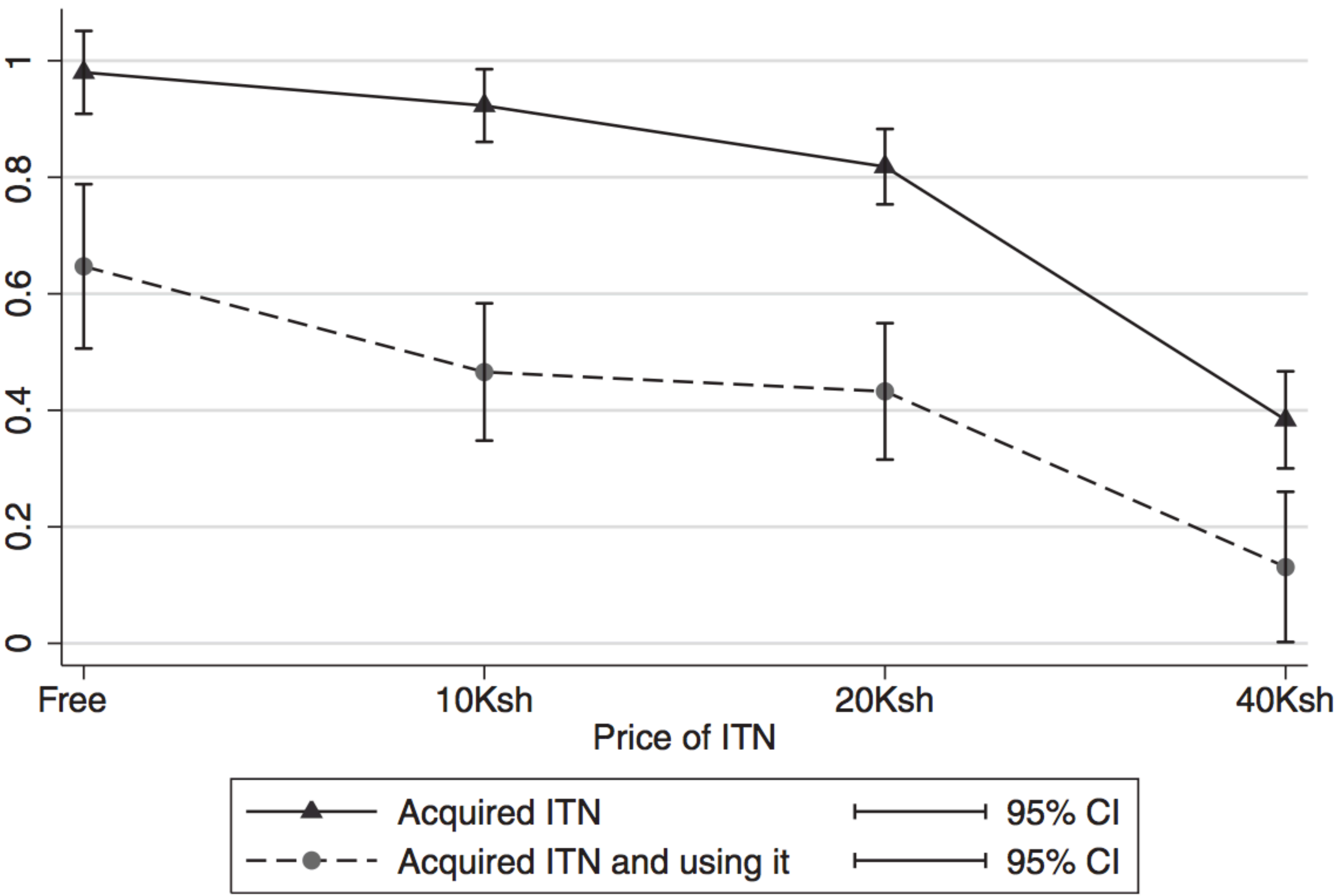

FIGURE I

Ownership vs. Effective Coverage

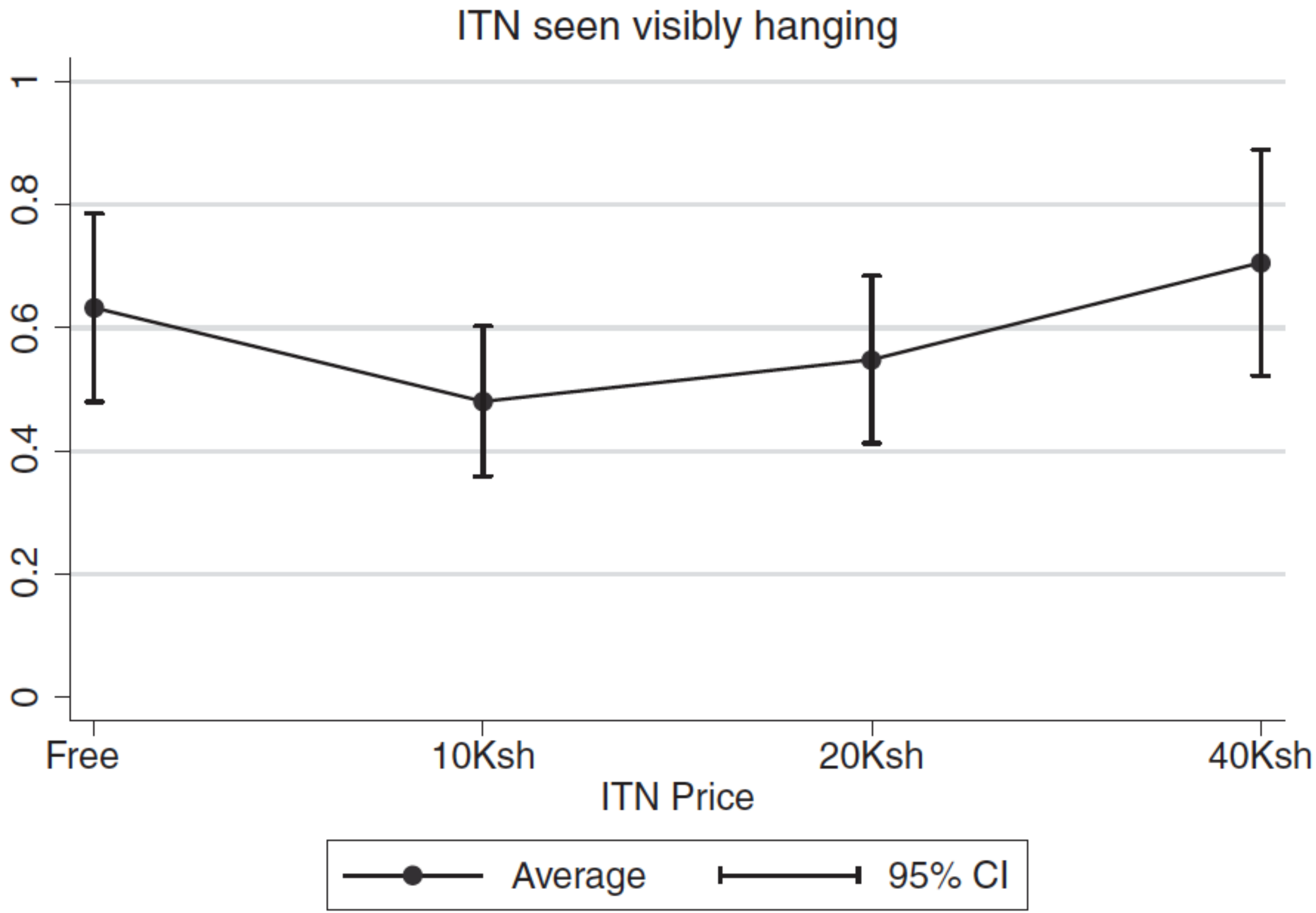

FIGURE II
Program ITN Usage Rates (Conditional on Uptake) by ITN Price

Econometric and statistical issues

- In regressing individual-level outcomes on clinic-level characteristics, need to bear in mind that observations within each clinic are not independent; need to adjust standard errors for clustering

- An alternative to parametric estimation is provided by Fisher's exact P-values under randomization-based inference (see Imbens & Wooldridge 2009).

- The idea behind randomization-based inference is that it tests the hypothesis that the treatment effect is precisely zero, and examines the likelihood that, under this null, the observed outcomes could have been the result purely of assignment of clinics to treatments

Randomization-based inference: an analogy

- Suppose I want to test the null hypothesis that a purported cure for the common cold has no effect on whether people catch colds or not

- I assign people at random the cure or a placebo, and estimate the average difference in frequency of colds between the two groups

- Under the null, I therefore assume that the reasons why some people have got colds consist entirely of other individual circumstances

- If I know the randomization algorithm for the assignment (eg just toss a coin), and if I know how many people were destined to catch cold anyway, I can calculate the exact probability that my random assignment would have yielded at least as many people catching colds as actually did so. That is my exact p-value.

Results (clinic-level data)

- A price of 20 Ksh was accompanied by a fall of 18 %age points of prenatal clients acquiring bednets, with an exact p-value of 0.036, and 27 %age points of usage (p-value 0.143)

- A price of 40 Ksh was accompanied by a fall of 58 %age points of prenatal clients acquiring bednets, with an exact p-value of 0.018, and 54 %age points of usage (p-value 0.54)

- In regression analysis, each 10 Ksh increase in price was accompanied by a fall of 8 nets in weekly sales (mean of 41); significant at 10% without controls, 1% with controls

Results (individual-level data)

- In regression analysis, each 10 Ksh increase in price was accompanied by a fall of 15 % points in purchase probability from a mean of 0.81 (significant at under 1%).
- For subsample of women making first pre-natal visit, effect is around 18 %age points.
- No significant regression effects on usage found, conditional on ownership
- Unconditional regression effects are negative and highly significant
- "Pure" psychological effects of price on usage are insignificant (though standard errors are high)

Results (selection)

- Comparison of cdfs of haemoglobin levels suggest that women who purchase ITNS are not more likely to be anemic than average prenatal women in the area – but they are more likely to be anemic than women who received free ITNs

- This suggests there is some selection effect, though this does not offset the overall effect of a positive price on take-up

- Authors estimate that "effective coverage of the anemic population is thus 60% lower under cost-sharing"

Overall conclusions

- Authors conduct a cost-effectiveness simulation based on assumptions about the size of the externality. They conclude “The general conclusion of this cost-effectiveness exercise is thus that cost-sharing is at best marginally more cost-effective than free distribution, but free distribution leads to many more lives saved“.

- There do not seem to be effects of prices on usage; there are some seleciton effects about which the authors say little but would be interesting to explore further

- How easily can we generalize from these results to other kinds of intervention?

The Muralidharan-Sundararaman teacher study

- Unintended side effects
- The Muralidharan-Sundararaman study
- The experiment
- The results
- Conclusions

Unintended side-effects (I)

- These are a variant of the measurement error problem, this time relating to difficulties about the definition or measurement of *Y.*

- Suppose the intervention occurs through two distinct channels, and the true causal process is something like this:

$$(5) \qquad Y_i = \alpha + \beta^1 . W_i + \beta^2 . W_i + \gamma . Z_i + \varepsilon_i$$

- However, the researcher in fact estimates only one component of the outcome, namely

$$(6) \qquad Y_i^2 = \alpha^2 + \beta^1 . W_i + \gamma . Z_i^2 + \varepsilon_i^2$$

Unintended side-effects (II)

- Then evidently the researcher's estimate will be an imprecise approximation to the true treatment effect. Will it also be biased?
- If the true value of the omitted coefficient is genuinely unknown, then there will not necessarily be bias. But there are circumstances under which we may expect it to be positive, or to be negative
- One particular circumstance in which it will be negative is if the intervention causes subjects, or others acting for them or with them, to alter the effort they allocate to two different tasks. Suppose the true process involves efforts that are a function of the intervention

$$(7) \qquad Y_i = \alpha + \gamma^1 t^1(W_i) + \gamma^2 t^2(W_i) + \gamma.Z_i + \varepsilon_i$$

Unintended side-effects (III)

- Suppose that the intervention cannot incentivize t^1 and t^2 directly but must do so by a noisy incentive mechanism, offering a reward

$$B_i = g^1 t^1 + g^2 t^2 + \eta_i \qquad (8)$$

where $g^1 > \beta^1 ; g^2 < \beta^2$

- Then, if the efforts are costly, this will lead to a relative re-allocation of effort by the subjects towards t^1 and away from t^2 – and if the costs are high enough, can lead to an absolute reduction in effort on t^2

- In special cases (modeled by Holmstrom and Milgrom 1991) it may even be better to give no incentive at all, because the effort on t^2 at zero incentives may outweigh the effect at any positive incentive

The Muralidharan-Sundararaman study

- Teaching (like learning) is well-known as a multi-tasking activity
- Some skills are much easier to test than others; both across subjects and within subjects - concern about “teaching to the test”
- M-S investigate this by an intervention to reward teachers for test performance, and using two methods to illuminate multi-task effects
  - Testing skills where teacher effort cost is higher
  - Testing skills that are not part of the incentive payment
- In both cases they find no evidence of adverse effects
- Also test group versus individual incentives

The Muralidharan-Sundararaman set-up

- 500 primary schools in rural Andhra Pradesh, India, chosen using a geographically stratified population-weighted random sample
- 100 schools given individual bonuses, 100 given group bonuses
- 100 given extra teacher, 100 given extra block grant
- Children given baseline math and reading tests, then tested after one year and after two years; Bonus calculated as lump sum X (% improvement in testscores – 5%)
- Tests distinguished between repeat/nonrepeat questions, basic versus conceptual skills, and incentive versus nonincentive subjects

TABLE 3
IMPACT OF INCENTIVES ON STUDENT TEST SCORES
Dependent Variable: Normalized End-of-Year Test Score

	Year 1 on Year 0		Year 2 on Year 0	
	(1)	(2)	(3)	(4)
	A. Combined (Math and Language)			
Normalized lagged test score	.503***	.498***	.452***	.446***
	(.013)	(.013)	(.015)	(.015)
Incentive school	.149***	.165***	.219***	.224***
	(.042)	(.042)	(.047)	(.048)
School and household controls	No	Yes	No	Yes
Observations	42,145	37,617	29,760	24,665
R^2	.31	.34	.24	.28
	B. Math			
Normalized lagged test score	.492***	.491***	.414***	.408***
	(.016)	(.016)	(.022)	(.022)
Incentive school	.180***	.196***	.273***	.280***
	(.049)	(.049)	(.055)	(.056)
School and household controls	No	Yes	No	Yes
Observations	20,946	18,700	14,797	12,255
R^2	.30	.33	.25	.28
	C. Telugu (Language)			
Normalized lagged test score	.52***	.510***	.49***	.481***
	(.014)	(.014)	(.014)	(.014)
Incentive school	.118***	.134***	.166***	.168***
	(.040)	(.039)	(.045)	(.044)
School and household controls	No	Yes	No	Yes
Observations	21,199	18,917	14,963	12,410
R^2	.33	.36	.26	.30

TABLE 4
IMPACT OF INCENTIVES BY REPEAT AND NONREPEAT QUESTIONS
Dependent Variable: Percentage Score

	COMBINED		MATH		TELUGU	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Percentage score on nonrepeat questions	.335***	.328***	.256***	.257***	.414***	.397***
	(.007)	(.007)	(.007)	(.008)	(.008)	(.007)
Percentage score on repeat questions	.352***	.42***	.252***	.386***	.452***	.468***
	(.006)	(.005)	(.007)	(.006)	(.007)	(.007)
Incremental score in incentive schools for nonrepeats	.030***	.039***	.033***	.046***	.027***	.033***
	(.009)	(.009)	(.009)	(.010)	(.010)	(.010)
Incremental score in incentive schools for repeats	.043***	.043***	.042***	.044***	.043***	.041***
	(.011)	(.011)	(.013)	(.012)	(.011)	(.013)
Test for equality of treatment effect for repeat and nonrepeat questions (*F*-statistic, *p*-value)	.141	.584	.374	.766	.076	.354
Observations	62,872	54,972	31,225	29,594	31,647	25,378
R^2	.24	.18	.26	.23	.29	.18

TABLE 5
IMPACT OF INCENTIVES BY MULTIPLE CHOICE AND NON-MULTIPLE-CHOICE QUESTIONS
Dependent Variable: Percentage Score

	COMBINED		MATH		TELUGU	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Percentage score on non-multiple-choice questions	.311*** (.007)	.311*** (.007)	.258*** (.007)	.278*** (.008)	.364*** (.008)	.344*** (.008)
Percentage score on multiple-choice questions	.379*** (.004)	.391*** (.004)	.227*** (.005)	.284*** (.004)	.529*** (.005)	.497*** (.005)
Incremental score on non-multiple-choice questions in incentive schools	.028*** (.009)	.037*** (.010)	.032*** (.010)	.047*** (.010)	.023** (.010)	.027** (.011)
Incremental score on multiple-choice questions in incentive schools	.034*** (.009)	.042*** (.009)	.034*** (.009)	.041*** (.009)	.034*** (.011)	.042*** (.009)
Test for equality of treatment effect for multiple-choice questions and non-multiple-choice questions (*F*-statistic *p*-value)	.168	.282	.671	.341	.119	.025
Observations	84,290	59,520	41,892	29,594	42,398	29,926
R^2	.197	.187	.213	.178	.302	.289

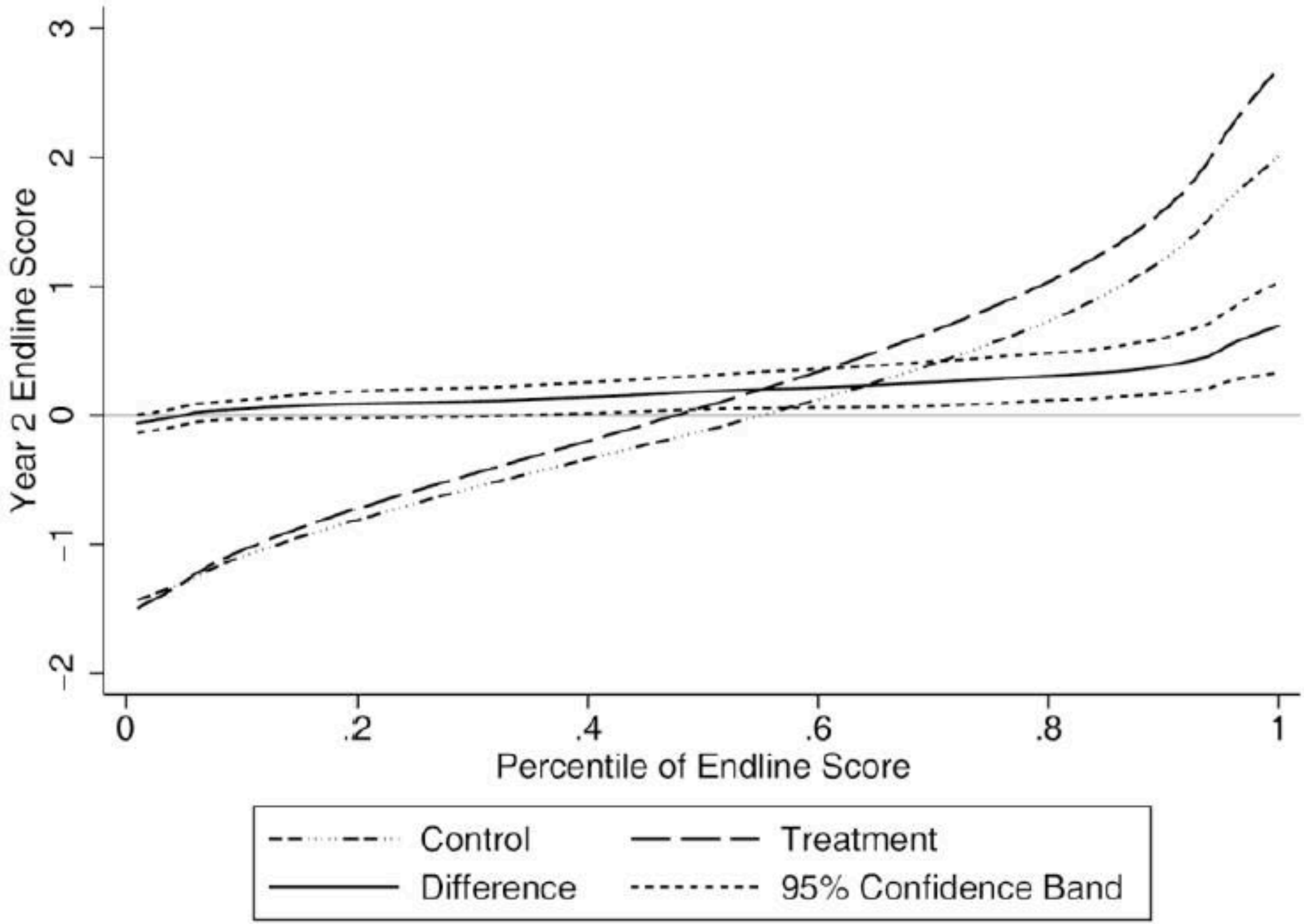

FIG. 2.—Quantile treatment effects of the performance pay program on student test scores.

TABLE 7
IMPACT OF INCENTIVES ON NONINCENTIVE SUBJECTS
Dependent Variable: Normalized End Line Score

	Year 1		Year 2	
	Science	Social Studies	Science	Social Studies
	A. Reduced-Form Impact			
Normalized baseline math score	.215***	.224***	.156***	.167***
	(.019)	(.018)	(.023)	(.024)
Normalized baseline language score	.209***	.289***	.212***	.189***
	(.019)	(.019)	(.023)	(.024)
Incentive school	.112**	.141***	.113**	.18***
	(.052)	(.048)	(.044)	(.050)
Observations	11,786	11,786	9,143	9,143
R^2	.26	.31	.19	.18

	B. Mechanism of Impact			
Normalized math predicted score	.382***	.340***	.274***	.330***
	(.032)	(.027)	(.041)	(.044)
Normalized Telugu predicted score	.298***	.487***	.429***	.360***
	(.028)	(.026)	(.036)	(.036)
Normalized math residual score	.319***	.276***	.232***	.247***
	(.025)	(.024)	(.032)	(.035)
Normalized Telugu residual score	.343***	.425***	.399***	.341***
	(.024)	(.025)	(.032)	(.036)
Incentive school	−.01	.011	−.054*	.009
	(.031)	(.027)	(.030)	(.033)
Incentive school × normalized math residual score	.048	.045	−.007	.014
	(.035)	(.031)	(.038)	(.042)
Incentive school × normalized Telugu residual score	−.006	.024	.058	.099**
	(.029)	(.031)	(.039)	(.043)
Test for equality math and Telugu residuals	.548	.001	.002	.128
Observations	11,228	11,228	8,949	8,949
R^2	.48	.54	.41	.39

TABLE 8
GROUP VERSUS INDIVIDUAL INCENTIVES
Dependent Variable: Normalized End-of-Year Test Score

	YEAR 1 ON YEAR 0			YEAR 2 ON YEAR 0		
	Combined (1)	Math (2)	Telugu (3)	Combined (4)	Math (5)	Telugu (6)
Individual incentive school	.156***	.184***	.130***	.283***	.329***	.239***
	(.050)	(.059)	(.045)	(.058)	(.067)	(.054)
Group incentive school	.141***	.175***	.107**	.154***	.216***	.092*
	(.050)	(.057)	(.047)	(.057)	(.068)	(.052)
F-statistic *p*-value (testing group incentive school = individual incentive school)	.765	.889	.610	.057	.160	.016
Observations	42,145	20,946	21,199	29,760	14,797	14,963
R^2	.31	.299	.332	.25	.25	.26

TABLE 10

IMPACT OF INPUTS VERSUS INCENTIVES ON LEARNING OUTCOMES

Dependent Variable: Normalized End-of-Year Test Score

	YEAR 1 ON YEAR 0			YEAR 2 ON YEAR 0		
	Combined (1)	Math (2)	Language (3)	Combined (4)	Math (5)	Language (6)
Normalized lagged score	.512***	.494***	.536***	.458***	.416***	.499***
	(.010)	(.012)	(.011)	(.012)	(.016)	(.012)
Incentives	.15***	.179***	.121***	.218***	.272***	.164***
	(.041)	(.048)	(.039)	(.049)	(.057)	(.046)
Inputs	.102***	.117***	.086**	.085*	.089*	.08*
	(.038)	(.042)	(.037)	(.046)	(.052)	(.044)
F-statistic *p*-value (inputs = incentives)	.178	.135	.298	.003	.000	.044
Observations	69,157	34,376	34,781	49,503	24,628	24,875
R^2	.30	.29	.32	.225	.226	.239

NOTE.—These regressions pool data from all 500 schools in the study: group and individual incentive treatments are pooled together as incentives, and the extra contract teacher and block grant treatments are pooled together as inputs. All regressions include mandal (subdistrict) fixed effects and standard errors clustered at the school level.

* Significant at 10 percent.

** Significant at 5 percent.

*** Significant at 1 percent.

Conclusions

- Studying unintended side effects is hard but may be important
- This study uses several ingenious methods to look out for these
- Of course you can’t investigate these unless you can measure them somehow
- And theory (eg about multi-tasking in incentive problems) can be helpful in knowing where to look.
- The positive results of this study should not be used to imply that unintended side-effects of teacher bonuses are never important

Toulouse School of Economics

TSE 16th January 2019

Applied Econometrics for Development: Experiments II

Ana GAZMURI

Paul SEABRIGHT