Econometrics II

Fabian Waldinger (LMU Munich)

Lecture Structure

1. Recap from last lecture
2. Randomized Experiments

Recap from Last Lecture

- Violations of GM3:
  - Omitted variable bias:
  - Measurement error in X
  - Simultaneity:
    - x causes y but also
    - y causes x

  This lead to a correlation of a regressor with the error term, and hence a violation of GM3

Introductory Example and Notation

- Suppose we wanted to learn about the causal effect of health insurance on health outcomes
- Why would be a simple comparison of people with and without health insurance be problematic?

Example Health Insurance

	Husbands		
	Some HI (1)	No HI (2)	Difference (3)
			A. Health
Health index	4.01 [.93]	3.70 [1.01]	.31 (.03)
			B. Characteristic
Nonwhite	.16	.17	−.01 (.01)
Age	43.98	41.26	2.71 (.29)
Education	14.31	11.56	2.74 (.10)
Family size	3.50	3.98	−.47 (.05)
Employed	.92	.85	.07 (.01)
Family income	106,467	45,656	60,810 (1,355)
Sample size	8,114	1,281	

Notation

- Individual (or firm, or country...) i will either receive a treatment or not:

$$D_i = \begin{cases} 1 & \text{if i receives treatment} \\ 0 & \text{otherwise} \end{cases}$$

- Potential outcomes (depending on whether one receives the treatment or not)
  - Y_{0i} if individual i does not receive the treatment
  - Y_{1i} if individual i receives the treatment
- Treatment effect: $Y_{1i} - Y_{0i}$
- Note: only one of these will be observed for each individual

Health Insurance Example

- Going back to the health insurance example
- Suppose there are two students Khuzdar and Maria:

	Khuzdar	Maria
Potential outcome without insurance: Y_{0i}	3	5
Potential outcome with insurance: Y_{1i}	4	5
Treatment (insurance status): D_i	1	0
Actual health outcome: Y_i	4	5
Treatment effect: $Y_{1i} - Y_{0i}$	1	0

Health Insurance Example

- The table is just imaginary because for each individual we either observe Y_{0i} or Y_{1i}
- Suppose we take the observed data at face value (and do not think about selection) and compare observed insurance status and how it relates to health outcomes:
  - $Y_{Khuzdar} = 4$ with insurance
  - $Y_{Maria} = 5$ without insurance
- The difference is:

$$Y_{\text{Khuzdar}} - Y_{\text{Maria}} = -1$$

- We would think that buying health insurance is bad for your health. The problem is that this difference suffers from selection bias (i.e. omitted variable bias)

Health Insurance Example

$$Y_{\text{Khuzdar}} - Y_{\text{Maria}} = Y_{1,\text{Khuzdar}} - Y_{0,\text{Maria}}$$

- Add and substract $Y_{0,\text{Khuzdar}}$:

$$= \underbrace{\{Y_{1,\text{Khuzdar}} - Y_{0,\text{Khuzdar}}\}}_{\text{causal effect for Khuzdar: } \mathbf{1}} + \underbrace{\{Y_{0,\text{Khuzdar}} - Y_{0,\text{Maria}}\}}_{\text{Selection Bias: } \mathbf{-2}}$$

- The selection bias term reflects Khuzdar's relative frailty
- The selection bias can potentially affect all comparisons of people with and without a certain treatment

Selection Bias

- Suppose a certain treatment has the same treatment effect (κ) on everyone. The outcome if i receives treatment is:

$$Y_{1i} = \kappa + Y_{0i}$$

- The comparison of means between individuals with and without treatment can therefore be rewritten as:

$$\begin{aligned} & Avg_n[Y_{1i}|D_i = 1] - Avg_n[Y_{0i}|D_i = 1] \\ &= \{\kappa + Avg_n[Y_{0i}|D_i = 1]\} - Avg_n[Y_{0i}|D_i = 0] \\ &= \underbrace{\kappa}_{\text{Avg. causal effect}} + \underbrace{\{Avg_n[Y_{0i}|D_i = 1]\} - Avg_n[Y_{0i}|D_i = 0]}_{\text{Selection Bias}} \end{aligned}$$

Selection Bias

- How do we know that the difference in means by treatment status is contaminated by selection bias (omitted variable bias)?
- Y_{0i} is shorthand for everything about person i related to the outcome, other than treatment status
- In our healthcare example: all the differences between insured and non-insured individuals (see bottom half of table above)
- If the only source of selection bias is a set of differences in characteristics that we can observe and measure, selection bias is easy to fix:
  $\rightarrow$ just include these characteristics as controls in the regression model
- In most situations the problem is that people who differ in observables most likely also differ in unobservables

Randomized Experiments

- Conceptually easy ways to overcome selection bias are randomized experiments
- In our healthcare example we could randomly provide health insurance to some people but not others and then compare their future health outcomes
- To be able to compare means we have to randomize a large enough sample from a given population so that the treatment and control groups will be similar in their underlying characteristics (e.g. a have a similar proportion of men and women, similar age, and so on)
- By the law of large numbers (LLN) the sample average will converge to the population averages
  $\rightarrow$ the randomly assigned groups should be similar in every way, including in ways that we cannot observe

Random Assignment Eliminates Selection Bias

- Because randomly assigned treatment and control groups come from the same underlying population, they are the same in every way, including their expected Y_{0i}
- I.e. $E[Y_{0i}|D_i = 0]$ and $E[Y_{0i}|D_i = 1]$ are the same if treatment D_i is randomly assigned

$$
\begin{aligned}
&E[Y_i \,|\, D_i = 1] - E[Y_i \,|\, D_i = 0] \\
&= E[Y_{1i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= E[\kappa + Y_{0i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= \kappa + E[Y_{0i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= \kappa
\end{aligned}
$$

- The last step follows from random assignment
- Hence, random assignment eliminates selection bias

Selection Bias

- Suppose a certain treatment has the same treatment effect (κ) on everyone. The outcome if i receives treatment is:

$$Y_{1i} = \kappa + Y_{0i}$$

- The comparison of means between individuals with and without treatment can therefore be rewritten as:

$$\begin{aligned} & Avg_n[Y_{1i}|D_i = 1] - Avg_n[Y_{0i}|D_i = 0] \\ &= \{\kappa + Avg_n[Y_{0i}|D_i = 1]\} - Avg_n[Y_{0i}|D_i = 0] \\ &= \underbrace{\kappa}_{\text{Avg. causal effect}} + \underbrace{\{Avg_n[Y_{0i}|D_i = 1]\} - Avg_n[Y_{0i}|D_i = 0]}_{\text{Selection Bias}} \end{aligned}$$

Random Assignment Elimatates Selection Bias

- Because randomly assigned treatment and control groups come from the same underlying population, they are the same in every way, including their expected Y_{0i}
- I.e. $E[Y_{0i}|D_i = 0]$ and $E[Y_{0i}|D_i = 1]$ are the same if treatment D_i is randomly assigned

$$
\begin{aligned}
&E[Y_i \,|\, D_i = 1] - E[Y_i \,|\, D_i = 0] \\
&= E[Y_{1i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= E[\kappa + Y_{0i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= \kappa + E[Y_{0i} \,|\, D_i = 1] - E[Y_{0i} \,|\, D_i = 0] \\
&= \kappa
\end{aligned}
$$

- The last step follows from random assignment
- Hence, random assignment eliminates selection bias

Randomized Experiments –An Example

- One may think that is very difficult to randomly assign medical insurance, however, the RAND Health Insurance Experiment (HIE), however, did just that
- 3,958 people were randomly assigned to one of 14 insurance plans
- We can group the insurance plans into four broad categories:
  - "Catastrophic coverage:"
    subscribers have to pay almost all health care expenditures up to a fairly high cap
  - "Deductible plan:"
    subscribers have to pay health care up to a lower cap
  - "Coinsurance plan:"
    subscribers only have to pay part of their health care costs
  - "Free plan:"
    all health care expenditure is covered

Randomized Experiments –Balancing Tests

- After randomization it is common (good) practice to check whether the treatment and control groups indeed look similar on observables that are either fixed over time or are measured before the treatment
- These tests are sometimes call balancing tests
- These are often done with simple t-tests comparing means

Balancing Tests -HIE

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Demographic characteristics					
Female	.560	−.023	−.025	−.038	−.030
		(.016)	(.015)	(.015)	(.013)
Nonwhite	.172	−.019	−.027	−.028	−.025
		(.027)	(.025)	(.025)	(.022)
Age	32.4	.56	.97	.43	.64
	[12.9]	(.68)	(.65)	(.61)	(.54)
Education	12.1	−.16	−.06	−.26	−.17
	[2.9]	(.19)	(.19)	(.18)	(.16)
Family income	31,603	−2,104	970	−976	−654
	[18,148]	(1,384)	(1,389)	(1,345)	(1,181)
Hospitalized last year	.115	.004	−.002	.001	.001
		(.016)	(.015)	(.015)	(.013)
B. Baseline health variables					
General health index	70.9	−1.44	.21	−1.31	−.93
	[14.9]	(.95)	(.92)	(.87)	(.77)
Cholesterol (mg/dl)	207	−1.42	−1.93	−5.25	−3.19
	[40]	(2.99)	(2.76)	(2.70)	(2.29)
Systolic blood pressure (mm Hg)	122	2.32	.91	1.12	1.39
	[17]	(1.15)	(1.08)	(1.01)	(.90)
Mental health index	73.8	−.12	1.19	.89	.71
	[14.2]	(.83)	(.81)	(.77)	(.68)

Randomized Experiments – Balancing Tests

- With randomized assignment people with and without health insurance look much more similar than in observational data
- Of the 40 comparisons of means only two (for proportion female in columns 4 and 5) are significantly different from each other
- Note: if we were to do 100 independent comparisons of means we would expect to find 5 significant differences (at the 5% level) – this is just the probability of a type I error

Example HIE – Results on Health Care Use

- Groups with cheaper access to health care have higher health-care use

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance catastrophic (5)
A. Health-care use					
Face-to-face visits	2.78 [5.50]	.19 (.25)	.48 (.24)	1.66 (.25)	.90 (.20)
Outpatient expenses	248 [488]	42 (21)	60 (21)	169 (20)	101 (17)
Hospital admissions	.099 [.379]	.016 (.011)	.002 (.011)	.029 (.010)	.017 (.009)
Inpatient expenses	388 [2,308]	72 (69)	93 (73)	116 (60)	97 (53)
Total expenses	636 [2,535]	114 (79)	152 (85)	285 (72)	198 (63)

Example HIE – Results on Health Outcomes

- Groups with cheaper access to health care, however, did not show a marketed improvement in health outcomes:

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance catastrophic (5)
		B. Health outcomes			
General health index	68.5	−.87	.61	−.78	−.36
	[15.9]	(.96)	(.90)	(.87)	(.77)
Cholesterol (mg/dl)	203	.69	−2.31	−1.83	−1.32
	[42]	(2.57)	(2.47)	(2.39)	(2.08)
Systolic blood pressure (mm Hg)	122	1.17	−1.39	−.52	−.36
	[19]	(1.06)	(.99)	(.93)	(.85)
Mental health index	75.5	.45	1.07	.43	.64
	[14.8]	(.91)	(.87)	(.83)	(.75)
Number enrolled	759	881	1,022	1,295	3,198

External vs. Internal Validity

- To make correct decisions based on empirical evidence we want to understand causal relationships
- We distinguish between two types of validity
  - *Internal validity:* the results give strong evidence of causality
  - *External validity*: the results are generalizable to other contexts
- Experiments are very good for ensuring internal validity

HIE External Validity

- Because everyone was offered some cap in their healthcare expenditure we may not learn how much truly uninsured individuals would benefit from health insurance
- Today's uninsured in the United States are different from the HIE population. They are:
  - Younger
  - Less educated
  - Poorer
  - Less likely to be working
- See Mastering Metrics Chapter 1 for a more recent experiment on providing health insurance

Estimating Treatment Effects in Experiments

- To analyse whether a certain treatment affected the outcome, we can just compare sample means in the treatment and control groups
- In practice, it may be useful to analyse experimental data using regression analysis
- With a constant treatment effect we have:

$$Y_{1i} - Y_{0i} = \kappa$$

- The observed outcome can be written as:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})D_i$$

- Using the fact that $Y_{1i} - Y_{0i} = \kappa$ we can rewrite this as:

$$Y_i = Y_{0i} + \kappa D_i$$

Estimating Treatment Effects in Experiments

- Y_i will not only differ because of the treatment status but also for other reasons: -> add an error term ε_i

$$Y_i = Y_{0i} + \kappa D_i + \varepsilon_i$$

- This very much looks like a simple regression model:

$$Y_i = \beta_1 + \beta_2 D_i + \varepsilon_i$$

- Where:
  - D_i is the treatment dummy
  - β_1 will estimate the mean of Y in the control group
  - β_2 will estimate the treatment effect κ

Advantages of Analyzing Experiments with Regression

1. Conditional random assignment:
   Sometimes randomization is conditional on some observable variable (e.g. on being poor)
2. You can add additional control variables to increase precision: although the control variables should be uncorrelated with D_i they may have substantial explanatory power for Y_i and therefore lower the standard error of the regression

Example of Large Randomized Experiment: Tennessee Project STAR

- Krueger (1999) econometrically re-analyses a randomized experiment of the effect of class size on student achievement
- The project is known as Tennessee Student/Teacher Achievement Ratio (STAR) and was run in the 1980s
- 11,600 students and their teachers were randomly assigned to one of three groups
  1. Small classes (13-17 students)
  2. Regular classes (22-25 students)
  3. Regular classes (22-25 students) with a full time teachers aide
- After the assignment, the design called for students to remain in the same class type for four years
- Randomization occurred within schools

Regression in Krueger (1999)

- Krueger estimates the following econometric model:

$$Y_{ics} = \beta_0 + \beta_1 SMALL_{cs} + \beta_2 Reg/A_{cs} + \beta_3 X_{ics} + \alpha_s + \varepsilon_{isc}$$

  - Y_{ics} = percentile score
  - $SMALL_{cs}$ = Indicator whether student was assigned to a small class
  - Reg/A_{cs} = Indicator whether student was assigned to a regular class with aide
  - α_s = School FE; because random assignment occurred within schools

Regression Results: Kindergarten

Explanatory variable	OLS: actual class size (1)	(2)	(3)	(4)
			A. Kindergarten	
Small class	4.82	5.37	5.36	5.37
	(2.19)	(1.26)	(1.21)	(1.19)
Regular/aide class	.12	.29	.53	.31
	(2.23)	(1.13)	(1.09)	(1.07)
White/Asian (1 = yes	—	—	8.35	8.44
			(1.35)	(1.36)
Girl (1 = yes)	—	—	4.48	4.39
			(.63)	(.63)
Free lunch (1 = yes)	—	—	−13.15	−13.07
			(.77)	(.77)
White teacher	—	—	—	−.57
				(2.10)
Teacher experience	—	—	—	.26
				(.10)
Master's degree	—	—	—	−.51
				(1.06)

Regression Results: 1st Grade

	B. First grade			
Small class	8.57	8.43	7.91	7.40
	(1.97)	(1.21)	(1.17)	(1.18)
Regular/aide class	3.44	2.22	2.23	1.78
	(2.05)	(1.00)	(0.98)	(0.98)
White/Asian (1 = yes)	—	—	6.97	6.97
			(1.18)	(1.19)
Girl (1 = yes)	—	—	3.80	3.85
			(.56)	(.56)
Free lunch (1 = yes)	—	—	−13.49	−13.61
			(.87)	(.87)
White teacher	—	—	—	−4.28
				(1.96)
Male teacher	—	—	—	11.82
				(3.33)
Teacher experience	—	—	—	.05
				(0.06)
Master's degree	—	—	—	.48
				(1.07)
School fixed effects	No	Yes	Yes	Yes
R^2	.02	.24	.30	.30

Problem 1: Attrition

- A common problem in randomized experiments is non-random attrition
- If attrition was random and affected the treatment and control groups in the same way, the estimates would remain unbiased
- Here, attrition is probably non-random: especially good students from large classes may have enrolled in private schools creating a selection bias problem
- Krueger addresses this concern by imputing test scores (from their earlier grades) for all children who leave the sample and then re-estimates the model including students with imputed test scores

Regression Results Imputing Test Scores to Address Attrition

	Actual test data		Actual and imputed test data	
Grade	Coefficient on small class dum.	Sample size	Coefficient on small class dum.	Sample size
K	5.32	5900	5.32	5900
	(.76)		(.76)	
1	6.95	6632	6.30	8328
	(.74)		(.68)	
2	5.59	6282	5.64	9773
	(.76)		(.65)	
3	5.58	6339	5.49	10919
	(.79)		(.63)	

Problem 2: Non-Compliance

- Students changed classes after random assignment
- A common solution to this problem is to use initial assignment (here initial assignment to small or regular classes) as an instrument for actual assignment (more on Instrumental Variable methods in the coming lectures)
- Krueger reports reduced form results where he uses initial assignment and not current status as explanatory variable
- In Kindergarten OLS and reduced form are the same because students remained in their initial class for at least one year
- From grade 1 onwards OLS (column 1-4) and reduced form (columns 5-8) are different.

Statistics Non-Compliance

A. Kindergarten to first grade

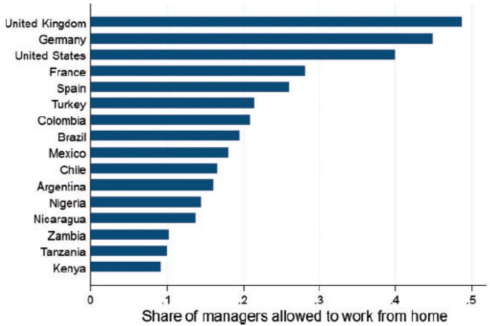
	First grade			
Kindergarten	Small	Regular	Reg/aide	All
Small	1292	60	48	1400
Regular	126	737	663	1526
Aide	122	761	706	1589
All	1540	1558	1417	4515

Non-Compliance Regression Results

Explanatory variable	OLS: actual class size				Reduced form: initial class size			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
B. First grade								
Small class	8.57	8.43	7.91	7.40	7.54	7.17	6.79	6.37
	(1.97)	(1.21)	(1.17)	(1.18)	(1.76)	(1.14)	(1.10)	(1.11)
Regular/aide class	3.44	2.22	2.23	1.78	1.92	1.69	1.64	1.48
	(2.05)	(1.00)	(0.98)	(0.98)	(1.12)	(0.80)	(0.76)	(0.76)
White/Asian (1 = yes)	—	—	6.97	6.97	—	—	6.86	6.85
			(1.18)	(1.19)			(1.18)	(1.18)
Girl (1 = yes)	—	—	3.80	3.85	—	—	3.76	3.82
			(.56)	(.56)			(.56)	(.56)
Free lunch (1 = yes)	—	—	−13.49	−13.61	—	—	−13.65	−13.77
			(.87)	(.87)			(.88)	(.87)
White teacher	—	—	—	−4.28	—	—	—	−4.40
				(1.96)				(1.97)
Male teacher	—	—	—	11.82	—	—	—	13.06
				(3.33)				(3.38)
Teacher experience	—	—	—	.05	—	—	—	.06
				(0.06)				(.06)
Master's degree	—	—	—	.48	—	—	—	.63
				(1.07)				(1.09)
School fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R^2	.02	.24	.30	.30	.01	.23	.29	.30

Example 2: Working From Home

- Working from home is becoming more important:



Selection Bias if we Used Observational Data

- Why can't we simply compare outcomes (e.g. productivity, promotion prospects,...) of individuals who work from home (WFH) to those who do not?
- Those who selected into working from home may be more or less productive to start with: classical selection bias
- Ctrip (a leading travel agency in China) decided to run an experiment to understand the causal effect on WFH
- Teamed up with economists at Stanford University (Nick Bloom, James Liang, John Roberts, Zhichung Jenny Ying) to run the experiment

Ctrip

- Ctrip is a leading travel agency in China
- Is quoted on the NASDAQ
- Worth about $5 billion at the time of the experiment

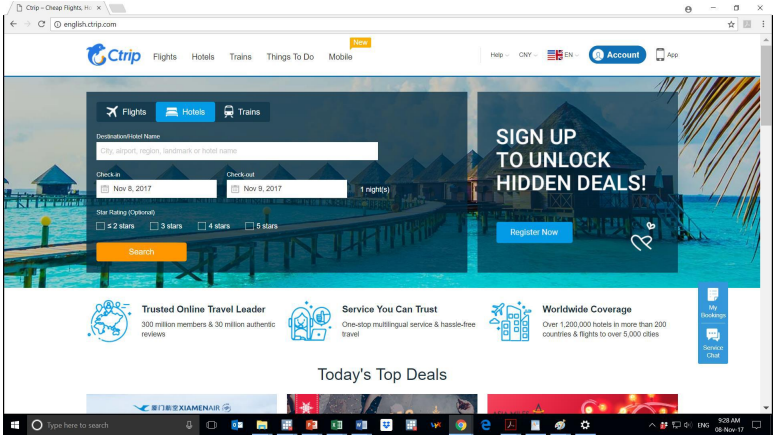

Headquarters in Shanghai


Main Lobby


Call Center Floor


Team Leader Monitoring Performance

Ctrip - English Website



Experimental Design

- Shanghai call centre workers were asked whether they wanted to change their work arrangements from
  - 5 days a week in the office to
  - 4 days at home and 1 day in the office
- 994 workers were asked whether they wanted to work from home; 503 volunteered for the experiment
- Why not simply compare the 503 individuals to the rest?
- Among the volunteers only those who had worked 6 months with the company, had broadband internet, and an independent workspace at home were allowed to participate (249 individuals)
- The 249 individuals were randomly allocated to a treatment (WFH) and control group (continued to work in the office)

Working From Home

Treatment groups were determined by a lottery

Working at home

Working at home

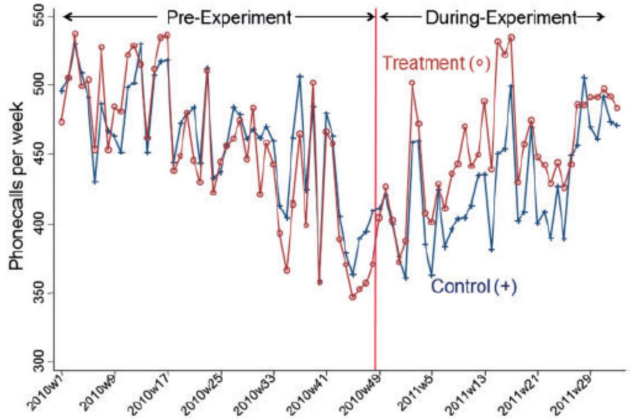
Working at home

Who Volunteers To WFH?

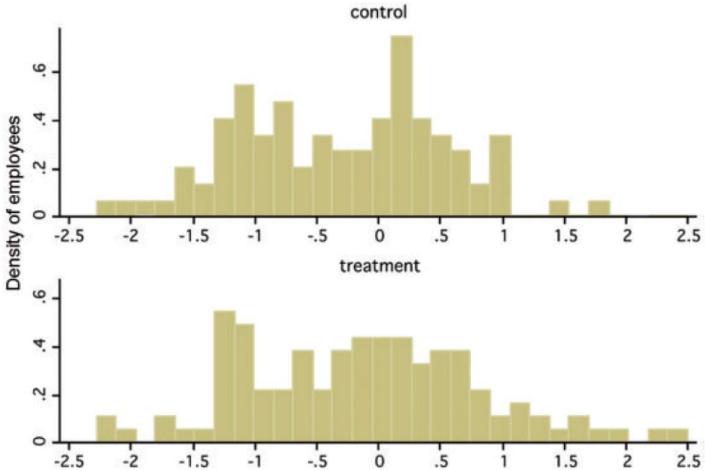
Dependent variable: volunteer to work from home	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Sample mean
Children	0.123**		0.054	0.075	0.081		0.084	0.08
	(0.056)		(0.083)	(0.083)	(0.083)		(0.084)	
Married[a]		0.095**	0.012	0.054	0.052		0.057	0.15
		(0.044)	(0.065)	(0.066)	(0.066)		(0.068)	
Daily commute (minutes[a])			0.062**	0.062**	0.071**		0.072**	80.6
			(0.030)	(0.031)	(0.032)		(0.0032)	
Own bedroom			0.095***	0.088**	0.089**		0.089**	0.60
			(0.035)	(0.035)	(0.036)		(0.037)	
Tertiary education and above				−0.080**	−0.088***		−0.086**	0.42
				(0.033)	(0.033)		(0.034)	
Tenure (months[a])				−0.268***	−0.415***		−0.401***	25.0
				(0.080)	(0.110)		(0.117)	
Gross wage (¥1,000)					0.048**	−0.019	0.048**	2.86
					(0.024)	(0.017)	(0.024)	
Age							−0.002	23.2
							(0.007)	
Male							0.010	0.32
							(0.036)	
Number of employees	994	994	994	994	994	994	994	994

Notes. The regressions are all probits at the individual level of the decision to work from home. Marginal effects calculated at the mean are reported. The total sample covers all Ctrip employees in

Productivity Over Time



Histograms of Productivity During Experiment



How to Evaluate the Experiment?

- The simplest way to evaluate the experiment would be to compare productivity differences between the treatment and control group during the experiment:

$$Productivity_i = \beta_1 + \beta_2 Treatment_i + \varepsilon_i$$

- Because they can measure productivity in every week, the authors can also estimate the following regression:

$$Productivity_{it} = \beta_1 + \beta_2 Treatment_{it} + \beta_3 Week1_t + \beta_4 Week2_t + \beta_5 Week3_t + ... + \varepsilon_i$$

- where $Week1$ is an indicator variable that is 1 if the observation comes from week1 and 0 for all other weeks and so on

Regression Results

	(2)
Dependent variable	Overall performance
Period	During experiment
Dependent normalizati	z-score
$Treatment_i$	0.184**
	(0.086)
Number of employees	249
Number of time period:	37
Individual fixed effects	No
Observations	7,476

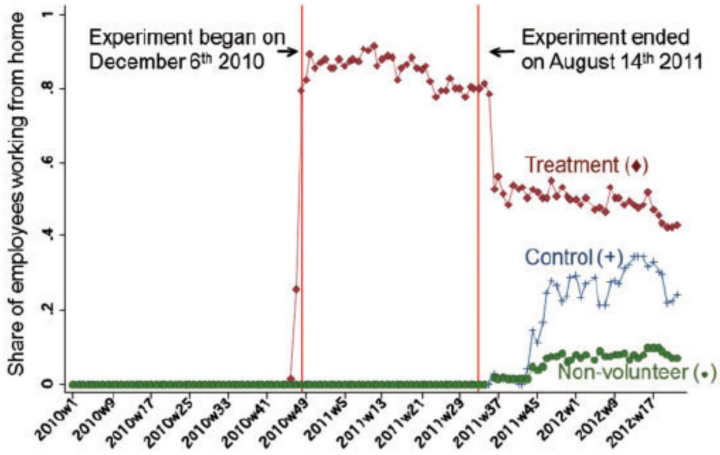
What Did the Firm Do After Seeing the Results?

- The experimental results indicate that productivity increased by about 13% in the treatment group
- The control group did not do worse than call centre workers in another location (rules out reduced motivation in the control group)
- WFH caused higher productivity and lower costs for the company (because office space is expensive in Shanghai)
  $\rightarrow$ After seeing the results the company rolled out voluntary WFH to the whole company

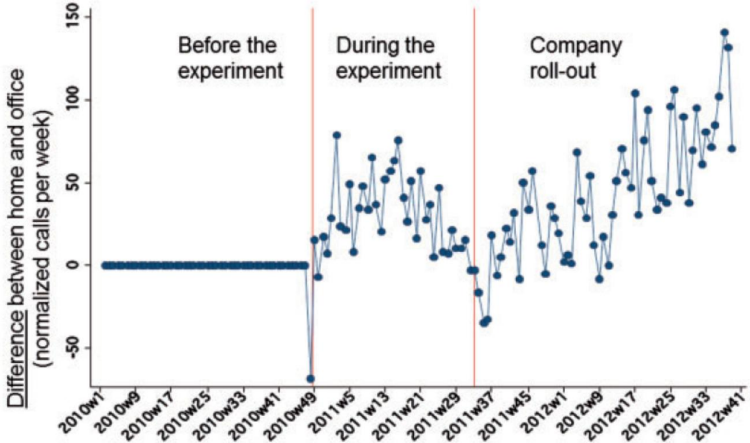
What Did the Firm Do After Seeing the Results?

- Some people in the treatment group decided to return to the office and some in the control group decided to work from home
- Because people (at least partly) understand whether they are more productive if they work from home, the ones who do not perform well sort back into working in the office and vice versa

Working From Home Treatment - Control



Selection Into Preferred Mode of Working After the Experiment Further Increased Performance

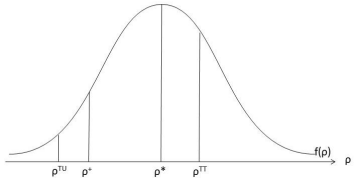


1

Potential Problems when Running Experiments

1. Randomization Bias

   Can occur if treatment effects are heterogeneous. The experimental sample may be different from the population of interest because of randomization. People selecting to take part in the randomized trial may have different returns compared to the population average



- $\rho_i = Y_{1i} - Y_{0i}$ is treatment effect of individual i
- ρ^* is the average treatment effect
- ρ^+ is the cutoff value above which people participate in the experiment
- ρ^{TT} is the treatment effect on the treated which is measured in the experiment
- ρ^{UT} is the treatment effect on the untreated which is not measured as those people would not

Potential Problems when Running Experiment

1. Supply Side Changes
   - If programmes are scaled up the supply side implementing the treatment may be different
   - In the trial phase the supply side may be more motivated than during the large scale roll-out of a programme
2. Attrition
   - Attrition rates (i.e. leaving the sample between the baseline and the follow-up surveys) may be different in treatment and control groups
   - The estimated treatment effect may therefore be biased

Potential Problems when Running Experiment

1. "Hawthorne" Effects
   - People behave differently because they are part of an experiment
   - If they operate differently on treatment and control groups they may introduce biases
   - If people from the control group behave differently these effects are sometimes called "John Henry" effects
2. Substitution Bias
   - Control group members may seek substitutes for treatment
   - This would bias estimated treatment effects downwards
   - Can also occur if the experiment frees up resources that can now be concentrated on the control group