2024 Capacity Development on

IMPACT EVALUATION

April 26, 2024 | 8:00 AM

Pasig City

Instrumental Variable

Aleli D. Kraft

Professor, UP School of Economics

PRESENTATION OUTLINE

Topic 1

Topic 2

Topic 3

Topic 4

TITLE

CHAPTER BREAK

Subtitle

Outline

- Introduction
- Uses of IV
- IV and Selection Bias
- 2SLS approach
- IV in RCTs
- Impacts with imperfect compliance
- LATE
- Randomized promotion design
- Instruments, validity and sources

Introduction

- Instrumental Variables (IV) is an estimation method that involves finding a variable (or instrument) that is highly correlated with program placement or participation but that is not correlated with unobserved characteristics affecting outcomes

Uses of IV in IE

- **IV methods can be used to address the problem of noncompliers in randomized controlled trials (RCTs).**
- Imperfect compliance may be due to:
  - Inability of program administrators ensure full compliance with treatment assignment
  - Universal coverage of program
  - Voluntary program enrollment

- Estimating average treatment effects (ATE) is not possible
  - No counterfactuals for “never takers” and “always takers”
  - Program impact on people who are never takers or always takers cannot be estimated
- Intuition for the IV approach
  - use the treatment assignment to obtain an estimate of the proportion of compliers in the treatment group
  - uses this estimated proportion to adjust the Intent to Treat (ITT) estimates to obtain an estimate of the Local Average Treatment Effects (LATE)

Uses of IV in IE

- **To avoid problems of selection bias in non-experimental settings**
- Selection bias → difference between the outcomes of the participants and the nonparticipant may reflects other differences (not just the program)
- How IV estimates can reduce selection bias
  - Variation in program participation (T) in the eligible population comes from two sources:
    - Random factors that have nothing to do with program impacts
    - Nonrandom selectivity (endogeneity in program placement)
  - Instrumental variables, which can be denoted by Z, extract the random variation from T (by regressing T on Z), and use only this extracted random part of T to estimate the program's impacts.

IV for selection bias

- Suppose this equation (1) to estimate impact

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

- Without selection bias, (i.e., non-random treatment) , β is an unbiased estimate of the treatment
- Selection bias may result because unobserved characteristics in the error term will contain variables that correlate with the treatment dummy, T. → $cov(T, \varepsilon) \neq 0$.
- This biases the estimate of the program effect β

IV for selection bias

- IV aims to clean up the correlation between T and ε
- Therefore, isolate the variation in T that is uncorrelated with ε
- Instrumental variable, Z, has to satisfy these two conditions:
  - Correlated with T: $cov(Z, T) \neq 0$
  - Uncorrelated with ε : $cov(Z, \varepsilon) = 0$
- Exclusion criterion → Instrument Z, affects selection into the program but is not correlated with factors affecting the outcomes

IV for selection bias: 2SLS approach

- 1st stage regression: Regress the treatment on the instrument, Z, the other covariates in eq. 1 and a disturbance term μ_i.

$$T_i = \gamma Z_i + \varphi X_i + \mu_i \quad (2)$$

- Predicted treatment from this regression: $\hat{T}$, reflects the part of the treatment that is affected only by Z and thus embodies only exogenous variation in the treatment.
- 2nd stage: Predicted T is substituted for treatment in equation 1:

$$Y_i = \alpha X_i + \beta(\hat{\gamma} Z_i + \hat{\varphi} X_i + \mu_i) \quad + \varepsilon_i. \quad (3)$$

- $\hat{\beta}_{IV}$ is the IV estimate of the program impact

Is treatment endogeneous?

- Using the Wu-Hausman test:
  - Regress T on Z and X and obtain the residual $\hat{\mu}_i$.
  - Regress Y on X, Z and $\hat{\mu}$.
- If coefficient of $\hat{\mu}$ is statistically significant from zero, unobserved characteristics jointly affecting the treatment T and outcomes Y are significant, and the null that T is exogenous is rejected.

IV in RCTs

- Two applications of IV for IE in randomized designs
  - Extension of randomized assignment when there is imperfect compliance
  - Use to design randomized promotion of treatment

Types of IE impact estimates

- Non-full compliance of treatment group – IE can estimate the effect of offering a program or effect of participating
- *Intention-to-treat* (ITT) - Estimated impact Δ when comparing groups to which the program has been randomly offered (in the treatment group) or not (the comparison group)
  - Weighted average of the outcomes of the participants and nonparticipants in the treatment group compared with the average outcome of the comparison group.
- *Treatment-on-the-treated* (TOT) - impact of the program for the group of individuals who are offered the program and who actually participate. This estimated impact is called the
- The ITT and the TOT will be the same when there is full compliance.

Types of IE impact estimates

- Imperfect compliance
- Case 1: Units assigned to treatment choose not to enroll or are left untreated
  - TOT – impact for individuals who are assigned to treatment who show up
- Case 2: Units assigned to comparison manage to participate
  - No counterfactual for some units in treatment group since units in comparison group were treated.

Impact options for imperfect compliance

- ITT – compares those who were intended to treat to intended to not treat
  - Non-compliance on the comparison side results in biased estimates of impacts

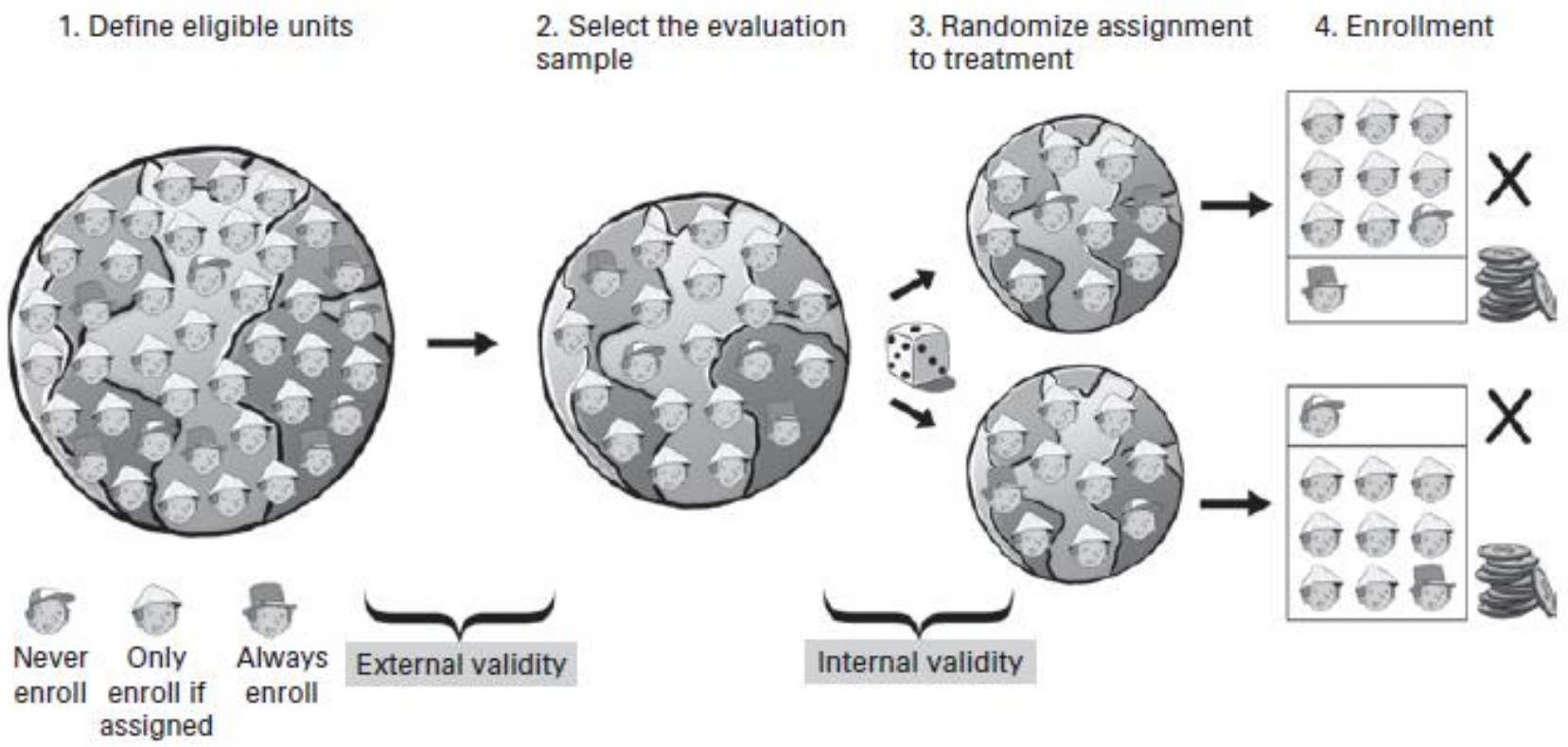
- Local average treatment effect (LATE) – program effects for the subgroup of compliers, i.e., units in the treatment group who enrolled in the program who would not have enrolled had they been in the treatment
  - Applies in cases when there is non-compliance in treatment, comparison or both

Estimating LATE

- Types of individuals by assignment and final take-up
  - *Enroll if assigned*. Individuals who comply with their assignment.
  - *Never*. Individuals who never enroll in or take up the program irrespective of assignment.
  - *Always*. Individuals who will find a way to enroll in the program or take it up even if they are assigned to a comparison group.
- Non-compliers – "Nevers" when assigned to treatment or "Always" when assigned to comparison

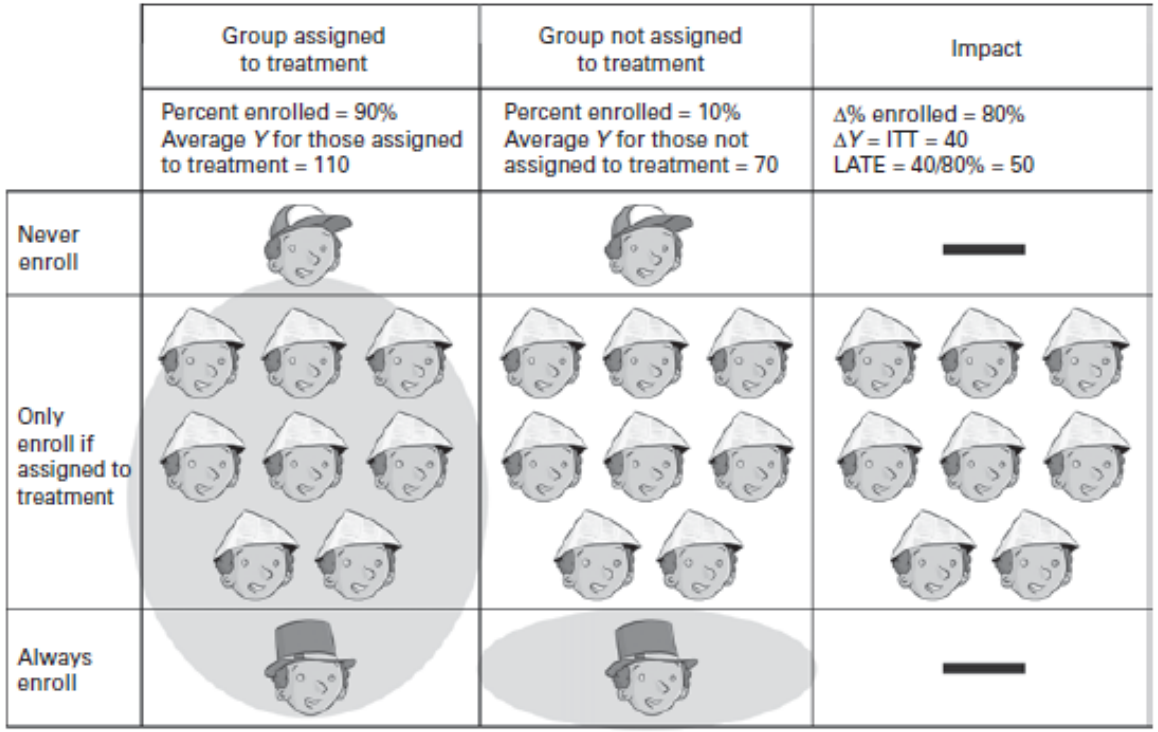
Assignment and final take-up

Figure 5.1 Randomized Assignment with Imperfect Compliance



LATE with imperfect compliance

Figure 5.2 Estimating the Local Average Treatment Effect under Randomized Assignment with Imperfect Compliance



Note: Δ = causal impact; Y = outcome. The intention-to-treat (ITT) estimate is obtained by comparing outcomes for those assigned to the treatment group with those assigned to the comparison group, irrespective of actual enrollment. The local average treatment effect (LATE) estimate provides the impact of the program on those who enroll only if assigned to the program (*Enroll-if-assigned*). The LATE estimate does not provide the impact of the program on those who never enroll (the *Nevers*) or on those who always enroll (the *Always*).

LATE

- Randomized assignment serves as an Instrumental Variable.
  - Predicts actual enrollment but is not correlated with other characteristics of the units that may be related to outcomes
  - Randomized assignment can be used as an instrument to predict final enrollment.
- This IV allows recovery of the LATE from ITT estimates for the *enroll-if-assigned type of units*.

LATE

- An easier way to estimate LATE, which is simply to use Z as an instrumental variable for T, that is:

$$LATE = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[T|Z=1] - E[T|Z=0]}$$

LATE

- Underlying assumption of monotonicity
  - Increase in Z from $Z = z$ to $Z = z'$ leads some to participate but no one to drop out of the program.
- Participation depends on certain values of the instruments such that:
  - $P(T = 1|Z = z)$ and $P(T = 1|Z = z')$ are the probabilities of participating when Z=z and Z=z'
  - Can be interpreted as the propensity scores for participation based on the instruments.

LATE

- The LATE, $\beta_{IV,LATE}$ can be written as:

$$\beta_{IV,LATE} = \frac{E(Y|P(Z) = P(z)) - E(Y|P(Z) = P(z'))}{P(z) - P(z')}$$

- Using linear IV methods.
  - 1st stage: estimate program participation T as a function of the instruments Z to obtain the propensity score: $\hat{P}(Z) = \hat{P}(T = 1|Z)$
  - 2nd stage: regress the outcome on the propensity score $\hat{P}(Z)$: $Y_i = [T_i.Y_i(1) + (1 - T_i).Y_i(0)]$
- Estimated program effect $\hat{\beta}_{IV}$ → is the average change in outcomes Y from a change in the estimated propensity score of participating $\hat{P}(Z)$, holding other observed covariates X fixed.

LATE vs. ATE

- Estimate of the LATE applies only to a specific subset of the population, the complier type.
- Individuals who comply (enroll if assigned types) with the program are different from those who do not comply (Nevers and Always)
- LATE estimate provides the impact for a particular sub-group of the population, i.e., only the enroll-if-assigned types.
- The impact we find through LATE does not apply to the Nevers or Always types

Random assignment as a valid IV

- A valid IV must satisfy two basic conditions.
    - Exogeneity. IV should not be correlated with the characteristics of the treatment and comparison groups. Achieved by randomly assignment of treatment.
    - Relevance. The IV must affect participation rates in the treatment and comparison groups differently. Participation is higher in the treatment group compared with the comparison group.

Randomized promotion design

- Randomized promotion (or encouragement design) – IV method that randomly assigns a promotion or encouragement to participate in a program
- Randomized promotion serves as an IV
- "Compliers" – those who enroll if "encouraged"
- Promotion is randomly assigned instead of treatment.
- Promotion should not influence outcomes
- LATE can be estimated

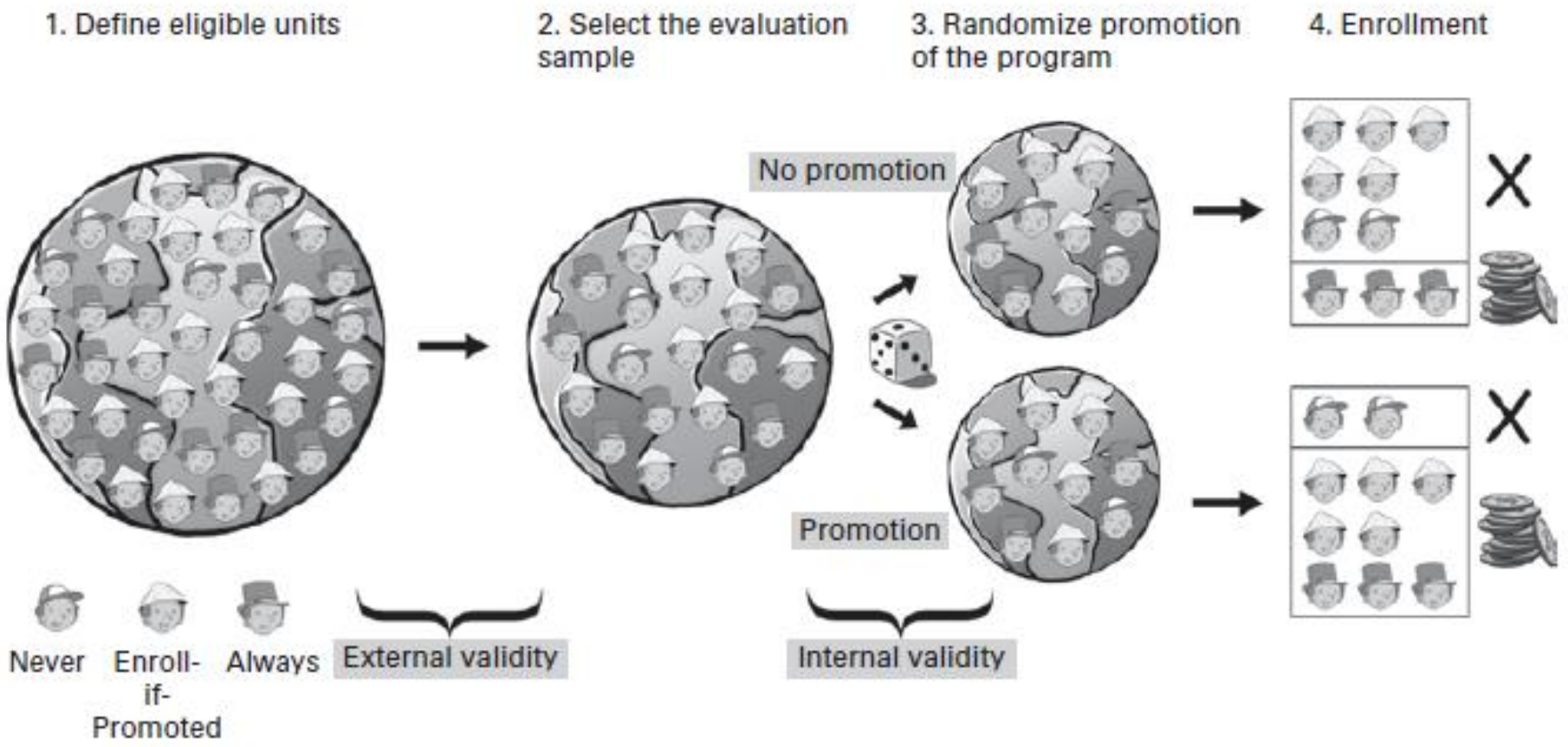
Randomized promotion design

- Conditions for the randomized promotion approach to produce valid estimates of program impact

    - Average characteristics of the two groups must be statistically equivalent.
    - Promotion should not directly affect the outcome of interest.
    - Promotion must substantially change the enrollment rates in the promoted group relative to the unpromoted group.

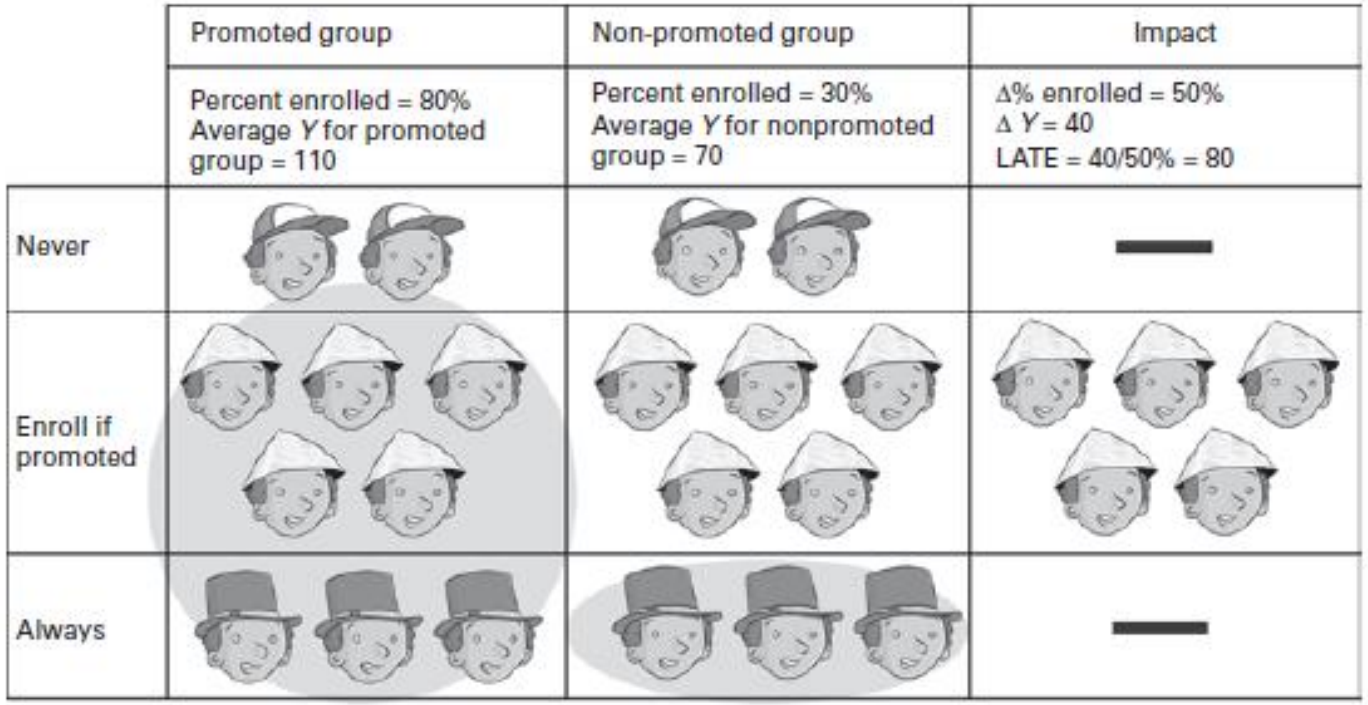
Randomized promotion process

Figure 5.3 Randomized Promotion



Estimating impact

Figure 5.4 Estimating the Local Average Treatment Effect under Randomized Promotion



Note: Δ = causal impact; Y = outcome. Characters that appear against the shaded background are those who enroll.

Limitations of random promotion

- Depends on effectivity of promotion strategy on increasing enrollment
- Only estimates LATE

Checklist: random promotion

- Compare the baseline characteristics of both groups – should be balanced
- Compare the program take-up rates in the promoted and non-promoted subsamples – should be higher in promoted subsamples
- Promotion campaign should not directly affect outcomes?

Instruments

- Drawback of IV - potential difficulty of finding a good instrument, esp. in non-experimental settings
- Estimates of program impact will be biased if $cov(Z, \varepsilon) \neq 0$
- With weak correlation of T with Z, the standard error of the IV estimate is likely to increase because the predicted outcome will be measured less precisely.
  - Asymptotically: $\hat{\beta}_{IV} = \beta + cov(Z, \varepsilon)/cov(Z, T)$,
  - The lower is $cov(Z, T)$ the greater the asymptotic bias of the $\hat{\beta}_{IV}\ from\ \beta$

Instruments

- 'Invalid" instruments – when instrument is correlated with the disturbance term of the outcome of interest
  - Yield biased and inconsistent IV estimator

- Weak instrument – instrument that is weakly correlated with the endogenous variable that it will not overcome bias of OLS and will yield misleading estimates of statistical significance

Instruments

- Testing for weak instruments
  - One cannot test for whether a specific instrument satisfies the exclusion restriction
  - Test for overidentifying restrictions with multiple instruments,
  - Steps:
    - Estimate the outcome equation by 2SLS, and obtain the residuals $\hat{\varepsilon}_i$
    - Regress $\hat{\varepsilon}_i$ on X and Z. Obtain the R^2
    - Use the null hypothesis that all the instrumental variables are uncorrelated with the residuals, $nR^2 \sim \chi_q^2$
      - q is the number of instrumental variables from outside the model minus the total number of endogenous explanatory variables.
      - If nR^2 > critical value, then the null hypothesis is rejected, and one can conclude that at least one of the instrumental variables is not exogenous.

Instruments: Sources

- Factors underlying program targeting and take-up can help in finding appropriate instruments
  - Randomization as a source of IVs
  - Other sources of instruments
    - Geographic variation
    - Political characteristics
    - Correlation of program with other policies
    - Exogenous shocks affecting program placement
    - Program design such as eligibility rules or the nature of treatment

Validity of instruments

- Methods to support validity of instruments:
  - Testing for over-identifying restrictions
  - Counter anticipated arguments why instrument is valid
  - Take care of omitted variables
  - Compare results from alternative instruments
  - Use intuition to suggest instruments' validity

Validity of instruments

- With over-identification
  - Include remaining potential instruments among the explanators rather than as instruments
  - Failure to reject the null hypothesis that the remaining instruments have zero coefficients in the 2nd stage would support the validity of those extra variables as instruments.
  - Key is knowing that an exactly identifying subset of the instruments are valid
- Preclude links between instruments and disturbances
  - Attempt to anticipate and test possible arguments why instruments may be invalid

Validity of instruments

- Omitted variables
  - IV estimator is biased if an omitted variable that belongs in the model is correlated with the X's or the Z's (instruments).
- Use alternative instruments
  - Getting similar results from alternative instruments enhances the credibility of IV estimates
- Use and check intuition
  - Intuitive argument for why an instrument is valid is better than no argument at all
  - Intuition can be checked by running reduced form regression with IV as the explanatory variable and the outcome variable or participation variable as dependent variables
  - IV should have coefficients that are different from zero with signs that support the identification story – dangerous if the sign is at odds with the identification story.

IV Applications

Capuno, Joseph, Aleli D. Kraft and Owen O'Donnell (2021) "Effectiveness of clinic-based cardiovascular disease prevention: A randomized encouragement design experiment in the Philippines" *Social Science and Medicine* 283

Introduction

- Improving primary prevention of CVDs in LMICs often rely on opportunistic screening of risk factors at health clinics.
  - WHO Essential Package of Noncommunicable Disease (NCD) Interventions for Primary Health Care on Low Resource Settings (PEN) sets out protocols for clinics to routinely assess adult patients for CVD risks
  - No evidence on the extent to which a visit to a clinic routinely following the PEN protocol is effective in raising exposure to CVD prevention.
- Observational studies
  - show often deficient PEN implementation, comparing outcomes at PEN and non-PEN clinics
  - Without inducing random exposure to PEN clinic, these studies were not capable of identifying effects of visiting such a clinic
- Dissemination of clinical guidelines did not always lead to their implementation and had inconsistent effects on outcomes.

Study Objective

- To establish by how much a visit to a clinic that should be screening opportunistically as per the PEN protocol raised exposure to primary prevention of CVD and consequently reduced risk factors.
  - Used conditional offer of a money-prize lottery ticket to induce random variations in visits to public health clinics responsible for operating the PEN.
  - Estimate effects of a check-up visit in exposure to CVD prevention processes (tests, diagnoses, medication, and medical advice), health behaviors, CVD risk factors and predicted CVD risk.

CVD in the Philippines

- Heavy CVD burden in the Philippines (
  - main cause of death in the Philippines
  - Between 1990-2017, age-standardized CVD death rate rose by 15% in the Philippines while it fell by 20% in Southeast Asia on average
  - Philippines ranks 28th worldwide for premature, avertable NCD mortality
- Philippines authorized its version of PEN in 2012 based on the WHO protocol for the integrated management of diabetes and hypertension, including risk assessment, risk screening, referral to facilities, prediction of global CVD risk score and based on score and individual risk factors, medication of (hypertension, diabetes and dyslipedemia) and lifestyle counselling
- DOH - financial and operational responsibility of providing PhilPEN technical assistance and for training clinic staff in the operation of the protocol, procure and distribute maintenance medicines
- LGUs - were obliged to adopt PhilPEN in their clinics, which operate Hypertension and Diabetes Clubs

Study design

- Evaluation approach
  - Determining the effects of visits to public health clinics operating PhilPEN on CVD prevention required comparison between individuals who visited such clinics and other individuals who had not.
  - Without random variation in clinic attendance, health and other differences between these two groups would have confounded the effects of a visit.
  - Used a random offer of a lottery ticket conditional on presentation for a check-up at a public health clinic responsible for operating PhilPEN .
  - Induced random variation in visits to such clinics used to identify the effects of those visits on exposure to CVD prevention processes, as well as health behaviours and CVD risk factors.

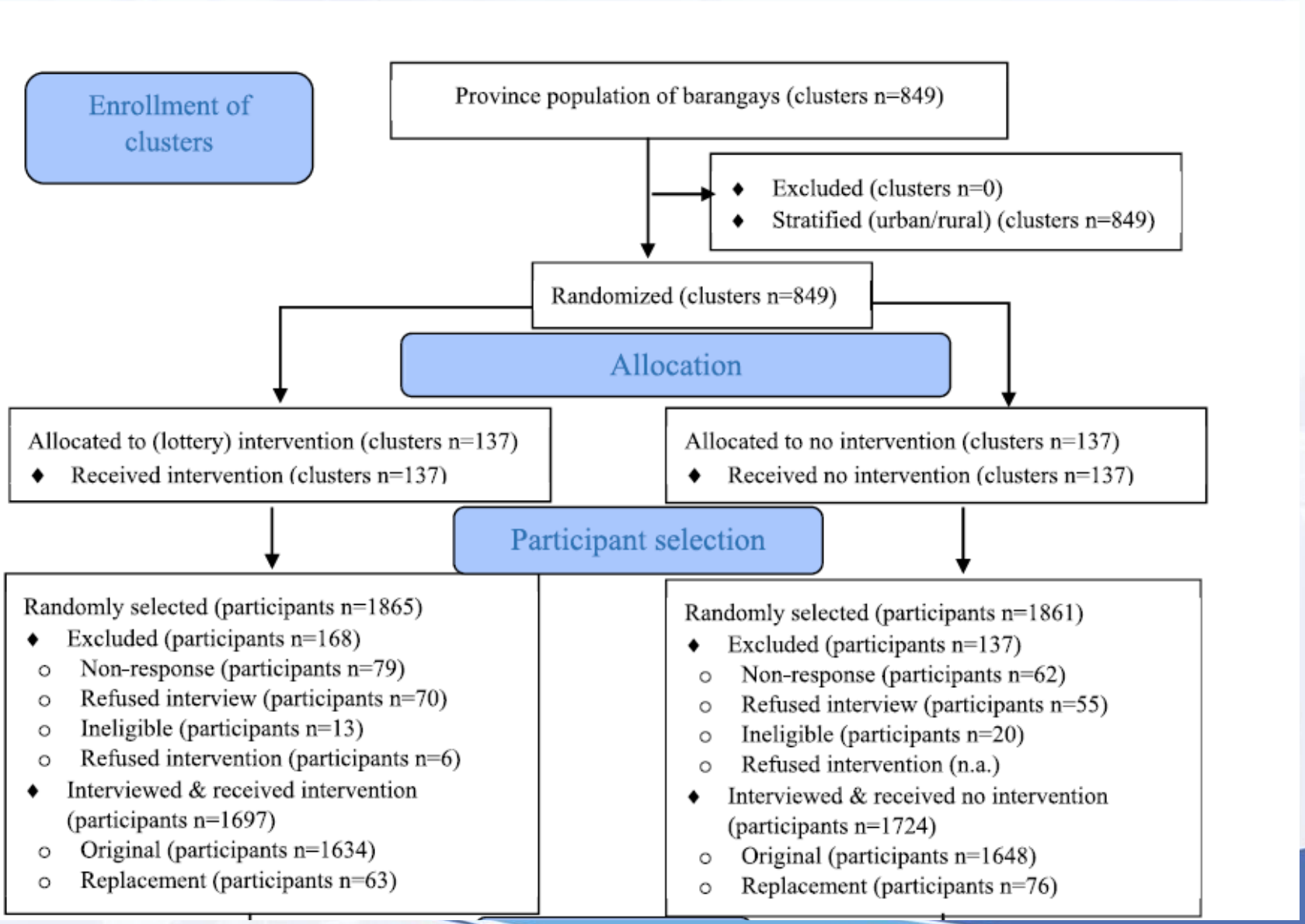
Nueva Ecija Cardiovascular Risk Experiment (NECVaRE)

- Study site: Nueva Ecija
  - Relatively high prevalence of major CVD risks in Central Luzon vs. National (DOST-FNRI 2015)
    - High blood pressure 27.4% vs.23.9%
  - Poverty – 20.7% poverty rate in 2015

Sources:
Philippine Statistics Authority
DOST-Food and Nutritional Research Institute

Nueva Ecija Cardiovascular Risk Experiment (NECVaRE)

- Timeline:
  - January to May 2018: Baseline survey
  - October to November 2018: Follow-up survey
- Sample inclusion criteria:
  - Aged 40 to 70 years old
  - Not diagnosed with diabetes or heart disease
  - Have never had a heart attack or stroke
  - Not currently taking medication for hypertension
  - No health condition preventing measurement of height, weight or blood pressure


Enrollment of clusters
Province population of barangays (clusters n=849)
Excluded (clusters n=0)
Stratified (urban/rural) (clusters n=849)
Randomized (clusters n=849)
Allocation
Allocated to (lottery) intervention (clusters n=137)
Received intervention (clusters n=137)
Allocated to no intervention (clusters n=137)
Received no intervention (clusters n=137)
Participant selection
Randomly selected (participants n=1865)
Excluded (participants n=168)
Non-response (participants n=79)
Refused interview (participants n=70)
Ineligible (participants n=13)
Refused intervention (participants n=6)
Interviewed & received intervention (participants n=1697)
Original (participants n=1634)
Replacement (participants n=63)
Randomly selected (participants n=1861)
Excluded (participants n=137)
Non-response (participants n=62)
Refused interview (participants n=55)
Ineligible (participants n=20)
Refused intervention (n.a.)
Interviewed & received no intervention (participants n=1724)
Original (participants n=1648)
Replacement (participants n=76)

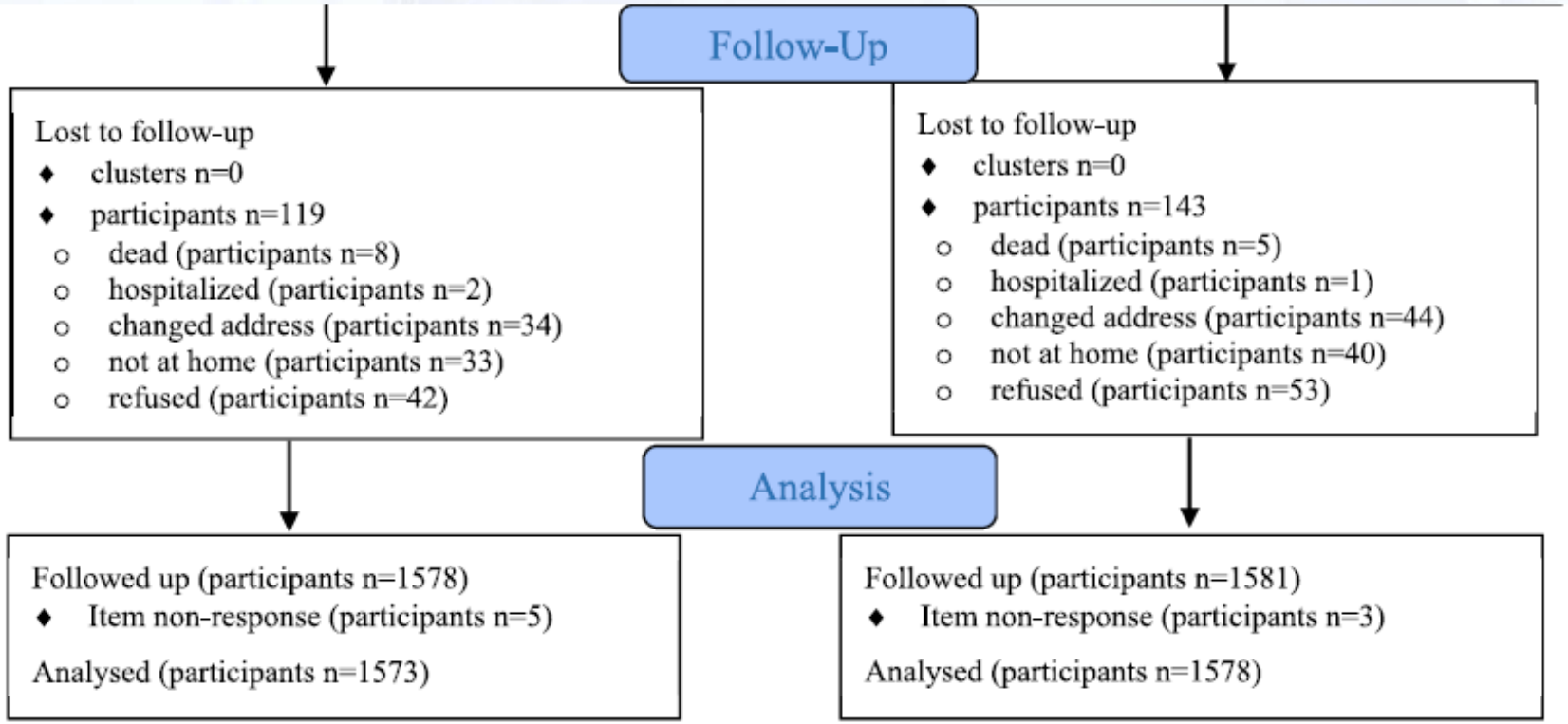
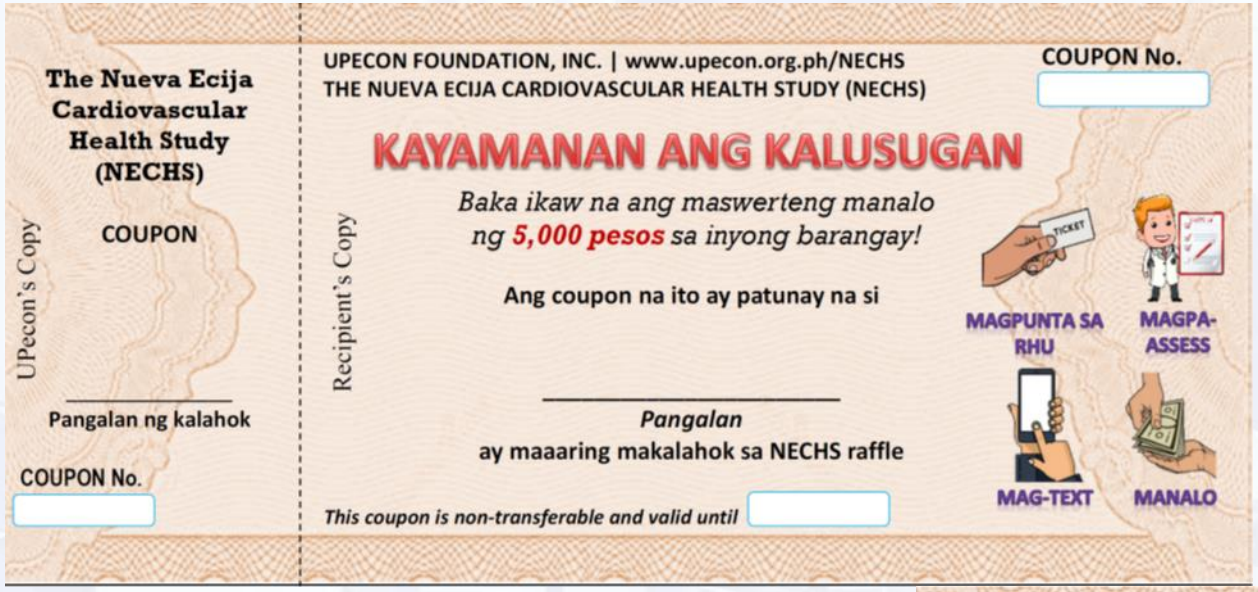


Fig. 1. Participant flow through cluster-randomized experiment. *Notes:* The 849 barangays in Nueva Ecija province were stratified by urban/rural and randomly drawn into the lottery intervention (137) or the control (137). "Replacement participants" were randomly selected to replace others with measurement errors that invalidated the calculated global CVD risk score at baseline. Households/participants that did not respond, refused or were ineligible at baseline were replaced with randomly selected others. All those lost due to item non-response at endline had measured blood pressure above 180 SBP or 120 DBP. Following the study's ethical protocol, the interview was stopped after such a reading and the participant was instructed to seek immediate medical attention. At baseline, any such cases were replaced by another respondent randomly selected from the same household or barangay.

Intervention

- Baseline interview – inform participants that they could go for a check-up at a health clinic where a doctor or nurse could examine them, perform tests and prescribe medication
- offer all participants a coupon in each intervention barangay a coupon giving entitlement to enter a lottery with a money prize conditional on going for a check-up at the public health clinic that served the barangay.
- Instructed them to ask for a medical assessment at the clinic where the coupon would be exchanged for a lottery ticket.
- A designated health worker recorded the participants' details, including age, carried out the medical assessment or referred the participant for assessment and issued the lottery ticket.
- Health workers were asked to deal with study participants as they would any patient presenting at the clinic for medical assessment.
- On average, the lottery induced around 7 participants per barangay to visit a clinic within a 6 week period

Lottery coupon



The Nueva Ecija Cardiovascular Health Study (NECHS)

COUPON

UPecon's Copy

_____
Pangalan ng kalahok

COUPON No.

UPECON FOUNDATION, INC. | www.upecon.org.ph/NECHS
THE NUEVA ECIJA CARDIOVASCULAR HEALTH STUDY (NECHS)

COUPON No.

KAYAMANAN ANG KALUSUGAN

Recipient's Copy

*Baka ikaw na ang maswerteng manalo ng **5,000 pesos** sa inyong barangay!*

Ang coupon na ito ay patunay na si

_____
Pangalan
ay maaaring makalahok sa NECHS raffle

This coupon is non-transferable and valid until

MAGPUNTA SA RHU

MAGPA-ASSESS

MAG-TEXT

MANALO

PAANO MAKASALI SA RAFFLE

1 Dalhin itong coupon at isang valid ID na may litrato sa RHU _____ hanggang

2 Humiling ng medical assessment sa RHU.

3 Pagkatapos ng assessment, humingi ng raffle ticket.

4 Basahin ang instruction sa raffle ticket para makasali sa raffle.

Kung may mga katanungan, tumawag sa:

NECHS Research Team
UP Campus, Diliman, Quezon City
Hotline Nos. :
0929-424-5515 (Smart)
0915-527-8801 (Globe)
Website: www.upecon.org.ph/NECHS

Verified by:

_____ _____
RHU Point Person Date

(1) Pumapayag akong sumali sa raffle.
(2) Pinapayagan ko ang UPecon na ibigay ang aking pangalan at tirahan sa RHU
_____.

_____
Name & signature of the respondent

Verified by:

_____
Enumerator/Date

Lottery ticket

Outcomes

Table 1
Outcomes.

Outcome		Definition and measurement
Outcomes are measured for all participants and are self-reported (except S-Zz) at follow-up		
Measurement		
A	Blood pressure (BP)	Had BP measured by medic since baseline
B	Blood sugar or cholesterol	Had blood sugar or cholesterol measured by medic since baseline
Diagnosis		
C	Hypertension	Ever diagnosed as hypertensive by medic
D	Undiagnosed hypertension	Measured systolic BP $\geq$ 140 and/or diastolic BP $\geq$ 90 and never been diagnosed as hypertensive by medic
E	Diabetes, dyslipidaemia or heart disease	Ever diagnosed with diabetes, dyslipidaemia or heart disease by medic
Medication		
F	Hypertension	Taken prescribed medication for hypertension since baseline
G	Diabetes or dyslipidaemia	Taken prescribed medication for diabetes or dyslipidaemia since baseline
Medical advice		Since baseline, received advice from medic to:
H	Quit smoking	quit smoking
I	Less alcohol	drink less alcohol
J	Less salt and fat	eat less salty and fatty foods
K	More fruit, veg. & pulses	eat more fruit, vegetables & pulses
L	Lose weight	lose weight
M	More exercise	do more physical exercise
Health behaviour		
N	Smoker	Currently smokes tobacco
O	Heavy episodic drinker	Drunk at least 4 (female)/5 (male) units alcohol on $\geq$1 occasion in last 30 days
P	Fruit & vegetables	(# days in 7 eats fruit + # days in 7 eats vegetables)/2
Q	No salt	Never adds salt or salty sauce to food
R	Physically active	Physically active for at least 30 min on typical day
CVD risk factors		
S	Systolic BP (SBP)	Average of last two (out of three) measurements on single occasion
T	Hypertension	SBP $\geq$140 and/or diastolic BP $\leq$ 90
U	Body mass index (BMI)	Measured weight (kg)/measured height $(m)^2$
V	Overweight	BMI $\geq$ 25
W	Waist circumference	Measured in cm
X	Central obesity	Waist circumference $\geq$ 80 cm (female)/90 cm (male)
Global CVD risk		
Y	CVD risk score	Percentage chance of heart attack or stroke within 10 years predicted from age, sex, SBP, BMI and smoking using office-based Globorisk
Z	Elevated CVD risk	CVD risk score $\geq$ 10%
Zz	High CVD risk	CVD risk score $\geq$ 20%

Notes: Outcomes in categories *Measurement, Diagnosis, Medication, Medical advice* and *Health behaviour* were reported by the participant in the follow-up survey. *CVD risk factors* were measured in that survey. *Global CVD risk* was predicted from measured risk factors (and reported smoking). Blood pressure and anthropometry were measured during the survey. "Medic" refers to a doctor or health worker.

Identification and estimation

- Identifying the variation on each outcome that was associated only with public health clinic visits that were induced by the random lottery offer.
- Estimated models with the following structure:

$$(1)\ Y_{1i}^{*} = \rho VISIT_i + \lambda Y_{01} + \mathbf{X_{0i}}\boldsymbol{\beta} + \varepsilon_i$$

$$(2)\ VISIT_i^{*} = \gamma LOTTERY_i + \theta Y_{0i} + \mathbf{X_{0i}}\boldsymbol{\delta} + \mu_i$$

- Where
  - Y_{1i}^{*} is a latent outcome at endline that corresponds to an observed continuous outcome ($Y_{1i} = Y_{1i}^{*}$ or determines an observed binary outcome by its sign $Y_{1i} = 1\ [Y_{1i}^{*} > 0]$
  - $VISIT_i^{*}$ is a binary indicator of having visited a public health clinic between the baseline and the endline and one month before the endline. This is determined by (2).
  - $LOTTERY_i$ is an indicator of belonging to the randomly selected intervention group that was offered a lottery ticket conditional on visiting the clinic for check-up.
- Controlled for baseline values of the outcomes and baseline covariates

Estimation

- For continuous outcomes,
  - (1) is estimated by two-stage least squares (2SLS) with visit instrumented by lottery.
  - Resulting estimate of ρ is the local average treatment effect (LATE) of a clinic visit – the effect on those who would be induced by the lottery offer to visit a clinic for a check-up.
- For binary outcomes,
  - assumed that the error terms to be jointly normally distributed,
  - estimated a bivariate probit model in which (1) and (2) are the linear latent indices and obtained the average marginal effect of VISIT on the probability of a positive outcome - corresponds to the average treatment effect (ATE)

Inference

- Deal with multiple comparisons problems by:
- Aggregate outcomes into three summary indices; clinic centered outcomes (a-c and e-m), health behavior outcomes (o-r) and global 10 year CVD risk score (y).
  - Each of the first two indices was a weighted average of the effect sizes of the outcomes entering that index
  - Effect of the clinic visit on each summary index was estimate by 2SLS
- Estimate effects on the separate outcomes and to correct for multiple comparisons through control of the False Discovery Rate (FDR) using the Benjamini two-stage procedure to produce sharpened q-values

Table 2
Balance of analytical sample on baseline outcomes.

		Baseline mean [SD]		H0: (1)=(2)	Normalized
		Control	Intervention	p-value	difference
		(1)	(2)	(3)	(4)
Measurement					
a	Blood pressure (BP)	0.805	0.803	0.929	0.005
b	Blood sugar or cholesterol	0.267	0.310	0.043	−0.096
Diagnosis					
c	Hypertension	0.036	0.045	0.190	−0.046
d	Undiagnosed hypertension	0.245	0.263	0.277	−0.041
e	Diabetes, dyslipidaemia or heart disease	0.021	0.029	0.157	−0.053
Medication					
f	Hypertension	0.025	0.028	0.575	−0.020
g	Diabetes or dyslipidaemia	0.015	0.018	0.417	−0.030
Medical advice					
h	Quit smoking	0.089	0.076	0.235	0.045
i	Less alcohol	0.056	0.044	0.130	0.055
j	Less salt and fat	0.221	0.226	0.776	−0.012
k	More fruit, veg. & pulses	0.304	0.322	0.395	−0.040
l	Lose weight	0.080	0.091	0.373	−0.037
m	More exercise	0.054	0.067	0.168	−0.054
Health behaviour					
n	Smoker	0.271	0.272	0.933	−0.003
o	Drinker	0.249	0.254	0.814	−0.011
p	Fruit & vegetables	4.05 [2.37]	3.98 [2.44]	0.416	0.043
q	No salt	0.272	0.319	0.069	−0.102
r	Physically active	0.613	0.673	0.075	−0.125
CVD risk factors					
s	Systolic BP (SBP)	125.2 [20.8]	125.1 [20.3]	0.934	0.003
t	Hypertension	0.264	0.284	0.235	−0.046
u	Body mass index (BMI)	23.0 [5.11]	23.3 [4.94]	0.165	−0.061
v	Overweight	0.291	0.315	0.213	−0.052
w	Waist circumference	85.9 [12.5]	85.9 [12.0]	0.943	−0.003
x	Central obesity	0.497	0.510	0.540	−0.025
Global CVD risk					
y	CVD risk	10.98 [7.79]	10.98 [8.20]	0.999	−0.000
z	Elevated CVD risk	0.471	0.451	0.270	0.040
zz	High CVD risk	0.117	0.125	0.524	−0.023
n clusters		137	137		
n participants		1578	1573		
F-statistic joint significance				1.463	(n = 3151)

Notes: Columns (1) and (2) give means in Control group (not offered lottery ticket) and Intervention group (offered lottery ticket) at baseline in sample used to produce estimates, i.e. after attrition and item non-response at endline. Standard deviations [SD] of continuous variables in brackets. Column (3) gives p-values from t-tests of equal means. Column (4) gives difference in means divided by the square root of the sum of the group variances. F-statistic is for test of joint significance of all the outcomes used to explain the intervention indicator. Variable definitions in Table 1. For any outcome that was defined for the period between baseline and endline, this table shows the mean at baseline for ever having experienced the outcome.

Table 3
Balance of analytical sample on baseline covariates.

	Baseline mean [SD]		H_0: (1)=(2)	Normalizeddifference
	Control	Intervention	p-value	
	(1)	(2)	(3)	(4)
Male	0.342	0.319	0.251	0.049
Age (years)	52.4 [8.88]	52.2 [8.67]	0.529	0.024
Married or cohabiting	0.810	0.790	0.178	0.051
Urban	0.270	0.257	0.815	0.028
Worked in last 7 days	0.570	0.582	0.643	−0.023
Education, highest level				
<elementary	0.143	0.124	0.215	0.055
Elementary	0.297	0.266	0.110	0.070
some high school	0.187	0.167	0.196	0.052
high school graduate	0.261	0.305	0.029	−0.098
College	0.112	0.138	0.106	−0.078
Wealth quintile group				
Poorest	0.215	0.193	0.300	0.055
2nd poorest	0.206	0.200	0.750	0.014
Middle	0.205	0.196	0.634	0.021
2nd richest	0.197	0.194	0.859	0.008
Richest	0.177	0.216	0.085	−0.099
Arthritis, rheumatism, osteoporosis	0.183	0.170	0.483	0.035
Angina, cancer, lung or neurological disease	0.082	0.090	0.456	−0.028
SF-20 health-related quality of life	84.7 [17.2]	85.2 [17.1]	0.446	−0.047
Relative with hypertension, cholesterol or diabetes	0.676	0.668	0.687	0.018
Public health insurance (PhilHealth) cover	0.681	0.680	0.962	0.002
# visits to public health clinic last 6 months	0.25 [1.04]	0.18 [0.68]	0.014	0.114
Outpatient visit last 30 days	0.081	0.077	0.726	0.016
Inpatient admission last 12 months	0.014	0.022	0.121	−0.058
n clusters	137	137		
n participants	1578	1573		
F-statistic joint significance			1.391 (N = 3151)	

Notes: Columns (1)–(4) as in notes to Table 2. F-statistic is for test of joint significance of all the covariates used to explain the intervention indicator. *Wealth quintile groups* formed from first principal component of house ownership, materials, size, amenities and state of repair, water source, sanitation, household durables, e.g. television, car, etc., ownership of assets (houses, land, agricultural, business, financial), receipt of remittances and conditional cash transfer. *Arthritis* etc., refers to reported diagnosis of any of these conditions. *Angina* etc. Refers to symptoms of angina identified from the Rose Questionnaire or diagnosis of cancer, lung disease, asthma, COPD, neurological or psychiatric disorder/disease. *SF-20* is the mean of the scores on the six components of SF-20 (Stewart et al., 1988; Hays et al., 1995). *Relative* is a parent or sibling. A public health clinic is a rural health unit, city health center or barangay health post. Number (#) of visits to a public health clinic is censored at 4.

Table 4
Effects of check-up clinic visit on aggregated outcomes.

	Clinic-centred	Health behaviour	Global CVD risk
	(1)	(2)	(3)
LATE	0.1562	0.0337	0.0055
(SE)	(0.0508)	(0.0648)	(0.2567)
Naïve p-value	0.0021	0.6038	0.9829
FWER p-value	0.0218	0.7500	0.9840
First stage	0.4732	0.4741	0.4746
(SE)	(0.0216)	(0.0216)	(0.0215)
F-statistic	489.4	490.1	493.2
n clusters	274	274	274
n participants	3151	3151	3151

Notes: First row gives 2SLS estimates of effects of a clinic visit. Column (1) and (2) are for weighted averages of effect sizes of outcomes a-c + e-m and o-r in Table 1, respectively. Control group means of these outcomes are zero by construction. Column (3) is for predicted 10-year CVD risk. Control group mean of this outcome is 10.9. FWER p-value is adjusted for multiple testing using the Romano and Wolf (2005, 2016) implemented in Stata® using rwolf (Clarke et al., 2019) with 10,000 bootstrap replications. Naïve p-value is unadjusted. First stage is OLS estimate of effect of lottery offer on probability of a clinic visit. F-statistic is the robust Kleibergen-Paap (2006) Wald rk F statistic of instrument strength. Standard errors in parentheses are adjusted for clustering at the level of randomization. All models include the baseline values of the outcome and the covariates listed in Table 3, except that age and sex are controlled through sex-specific 5-year age group indicators.

Table 5
Effects of a check-up clinic visit on specific outcomes.

		Effect	(SE)	Naïve p-value	FDR q-value	Control mean	Wald IV strength
		(1)	(2)	(3)	(4)	(5)	(6)
Measurement							
a	Blood pressure (BP)	0.1699	(0.0597)	0.0048	0.0640	0.4233	404.9
b	Blood sugar or cholesterol	−0.0028	(0.0243)	0.9091	1.0000	0.1071	402.0
Diagnosis							
b	Hypertension	0.0354	(0.0220)	0.0846	0.2420	0.0488	403.5
d	Undiagnosed hypertension	−0.0108	(0.0295)	0.7146	1.0000	0.2376	404.9
e	Diabetes, dyslipidaemia or heart disease	−0.0122	(0.0126)	0.3341	0.6970	0.0425	403.3
Medication							
f	Hypertension	0.0221	(0.0171)	0.1727	0.3640	0.0336	407.3
g	Diabetes or dyslipidaemia	−0.0065	(0.0093)	0.4737	0.9010	0.0247	403.7
Medical advice							
h	Quit smoking	0.0357	(0.0177)	0.0182	0.1170	0.0241	404.5
i	Less alcohol	0.0173	(0.0115)	0.1121	0.2900	0.0184	404.9
j	Less salt and fat	0.0684	(0.0315)	0.0243	0.1260	0.1071	402.7
k	More fruit, veg. & pulses	0.1280	(0.0408)	0.0008	0.0220	0.1229	401.5
l	Lose weight	0.0464	(0.0203)	0.0119	0.1060	0.0336	404.3
m	More exercise	0.0287	(0.0163)	0.0610	0.2250	0.0368	402.6
Health behaviour							
n	Smoker	0.0019	(0.0179)	0.9172	1.0000	0.2643	407.6
o	Heavy episodic drinker	0.0066	(0.0249)	0.7895	1.0000	0.2028	406.7
p	Fruit & vegetables	−0.0322	(0.1587)	0.8394	1.0000	4.3403	492.5
q	No salt	0.0850	(0.0508)	0.0866	0.2420	0.2218	405.2
r	Physically active	0.0599	(0.0757)	0.4311	0.8590	0.5311	402.4
CVD risk factors							
s	Systolic BP (SBP)	−0.4381	(0.9676)	0.6507	1.0000	123.99	491.0
t	Hypertension	0.0036	(0.0292)	0.9028	1.0000	0.2592	404.0
u	Body mass index (BMI)	−0.0393	(0.1030)	0.7029	1.0000	22.977	488.5
v	Overweight	0.0089	(0.0186)	0.6314	1.0000	0.2858	403.3
w	Waist circumference	−0.0839	(0.4755)	0.8599	1.0000	84.938	491.8
x	Central obesity	0.0106	(0.0203)	0.6056	1.0000	0.4708	405.8
Global CVD risk							
z	Elevated CVD risk	−0.0225	(0.0245)	0.3559	0.6970	0.4683	404.3
zz	High CVD risk	0.0349	(0.0187)	0.0449	0.1870	0.1096	403.6

Notes: Column (1) gives bivariate probit estimates of ATEs of a clinic visit on binary outcomes (a-r, t, v, x-zz) and 2SLS estimates of LATEs of a clinic visit on continuous outcomes (s, u, w). In both cases, a clinic visit was (effectively) instrumented with the randomized conditional offer of a lottery ticket. Column (2) gives standard errors adjusted for clustering at the level of randomization of the instrument. All models also included the baseline values of the outcome and the covariates listed in Table 3, except that age and sex are controlled through sex-specific 5-year age group dummies. Column (3) gives unadjusted p-values. Column (4) gives Benjamini et al. (2006) sharpened q-values that adjusted for multiple testing by controlling the False Discovery Rate (Anderson, 2008). Column (5) gives the mean of the outcome in the control group at endline. Column (6) gives the Wald test statistic for instrument strength. For binary outcomes, this is $\sim\chi^2(1)$. For continuous outcomes, it is the robust Kleibergen-Paap (2006) Wald rk F statistic. For all models, sample size is 276 clusters and 3151 individuals. See Appendix E, Table E3 for estimates of the effect of the lottery offer on clinic visit probability.

Discussion

- Aggregate outcome - Going to a public health clinic for a check-up increased exposure to CVD prevention, increase weighted average of effect sizes if clinic-centered outcomes by .16 of a std. deviation
- Results suggest that visiting a clinic for a check-up did raise the likelihood of receiving some kind of advice on healthy living.
- But did not translate to any consequent change in smoking, drinking, diet and exercise.
- Possible explanation – receipt of medical advice while large relative to low base levels, were small in the context of highly prevalent unhealthy behaviors
- Lack of effect on health behaviours and a muted effect on the use of hypertensives, explain why there were no effects on physiological risk factors and the predicted CVD risk.

Limitations

- Encouragement design effectivity to identify effects of a check-up clinic visit in CVD prevention outcomes dependent on those outcomes not varying with respect to the receipt of the lottery offer conditional on clinic attendance
  - Overestimate effects - if the Cardiovascular Health Study label and the fact the we instructed participants to ask for an assessment may have led clinic staff to apply PhilPEN protocol more diligently
  - Underestimate effects if clinic staff were more dismissive of patients who they think are just interested in the lottery.
- Study's encouragement design could potentially limit the extent to which its results generalize.
  - Effects estimated are local to the population who would be induced by a lottery to visit a public health clinic for a check-up.
  - If the effects of these compliers differ from the non-compliers, then the study did not estimate the average effects in the population from which the sample was drawn.

Summary

- Study demonstrated that going for a check-up at a public health clinic responsible for conducting opportunistic CVD risk screening using a widely adopted protocol increased the exposure of Filipinos to CVD prevention processes only modestly.

- Issuing clinics in the Philippines with well-founded protocols for CVD risk screening and management were not sufficient to ensure that predominantly poor individuals got more of the diagnoses and medications that are effective in preventing CVDs.

Thomas, D.R., Harish, S.P., Kennedy, R. and Urpelainen, J., (2020). "The effects of rural electrification in India: An instrumental variable approach at the household level" ***Journal of Development Economics,*** **146, p.102520.**

IE of rural electrification

- Selection problem - those with connections may be qualitatively different from those without
- Some have used propensity score matching, IV or DID approach in an attempt to address this selection problem and produce causal estimates.
- Not all evidence is conclusive about the positive effects of electrification
- None of the studies have attempted to estimate the causal impact of legal grid connections.

Conceptual Framework

- Legal electrification has positive effects through 3 channels
  - Electrification increases the hours and quality of lighting for HH.
  - Electrification gives the HH the ability to use more appliances.
  - Electrification can improve the health of HH, as electric lights and appliances have less pollutants than other sources.
- Therefore, expect positive effects on expenditures, household activity and appliance usage and ownership.

Research Design

- IV design - exploiting a law in the state of upper Pradesh → rural households differential access to legal electricity connection based solely on the distance from of a household from a power pole. – within 40 m of a power pole
- Sampled households within 20-35 m, and within 45-60 m. of a pole.
  - Did not survey HH that do not fall within the 35-45 of a power pole because it increases the likelihood of error on the part of enumerators and state regulators.
  - By creating a buffer zone, increase the proportion of compliers in the sample.
- Conducted an original survey in early 2018 of 686 HH across 120 habitations with a total of 154 electricity poles. Survey was conducted after announcement of the scheme but prior to its implementation

Outcome variables

Table 1
Summary of outcome variables, and the hypothesized directions of their relationships with legal electrification.

Family	Outcome	Description	Hypothesis
Expenditure	Total expenditure	Total household expenditure in a typical month measured in rupees.	⇑
	Expenditure on food	Expenditure on food by a household in a typical month measured in rupees.	⇑
	Expenditure on education	Expenditure on education by a household in a typical month measured in rupees.	⇑
	Expenditure on kerosene	Expenditure on kerosene by a household in a typical month measured in rupees.	⇓
Household Activity	Child activity	Number of hours children spend at home in a given day.	⇑
	Adult activity	Number of hours adults spend at home in a given day.	⇑
Appliances	Number of appliances	The number of appliances owned by a household.	⇑
	Appliance use	The number of daily hours using appliances by a household.	⇑

Estimation Strategy

- Use distance from power poles as an instrument which induces a discontinuous change in the probability of having a legal connection to grid electricity
  - Distance takes on 1 when they are within 20-35 of a pole, 0 when they are within 45-60 m
- Opt for local randomization design – HH very close to cutoff are investigated vs. RDD since HH directly at border are not measured
- Treatment variable = Legal electrification
  - value of 1 if the HH is connected to the grid legally, 0 if there is an illegal connection or no connection.
- Two stage least squares IV strategy:
  - First stage: estimate the effect of the instrument on the endogenous treatment variable:

$$\widehat{Legal}_{hpv} = \beta Distance + X_h + \mu_p + \varepsilon_{hpv}$$

  - 2nd stage

$$Outcome_{hpv} = \beta \widehat{Legal}_{hpv} + X_h + \mu_p + \varepsilon_{hpv}$$

Estimation strategy

- Deterministic monotonicity – the level of treatment taken is a monotonic increasing function of the level of the instrument variable.
- No interference – treatment assignment of one HH does not affect the behavior of another household.
  - Anticipate no spillover effects of outcomes
- Exclusion criterion - IV variables has no direct effect on outcome. Given that the instrument is based on a discontinuity, exclusion criterion holds.

For identification

- Balance in pre-treatment covariates of HH at both sides of discontinuity

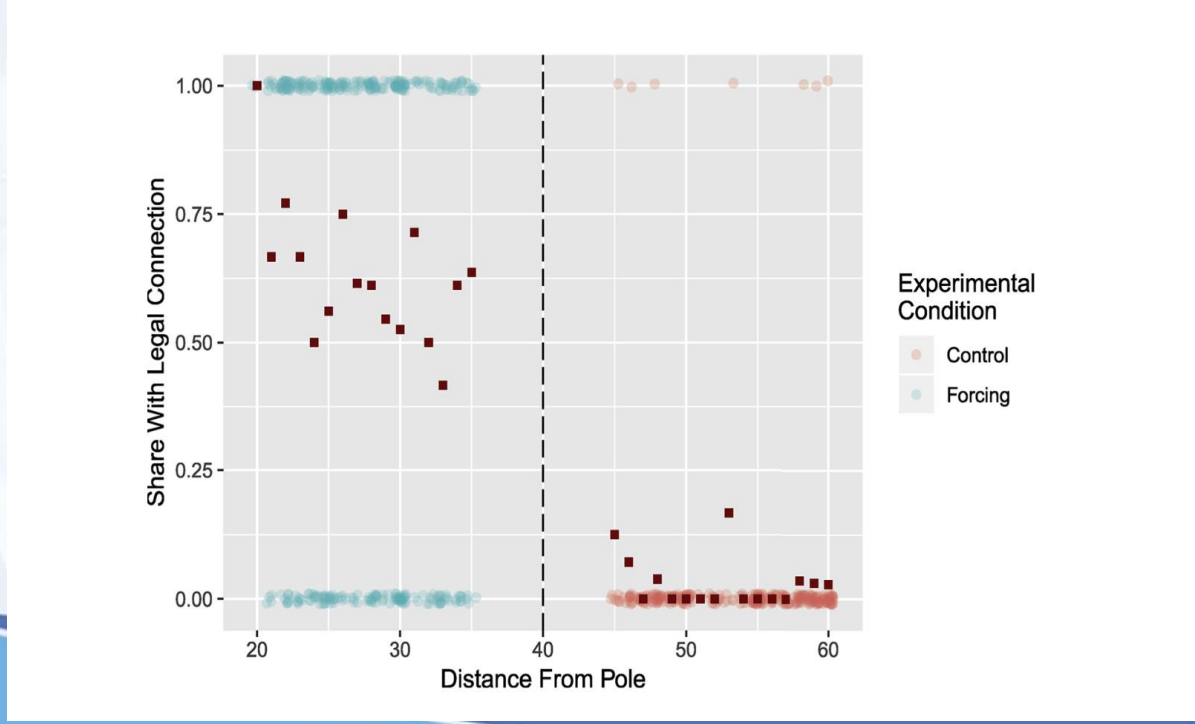
Table 3

This table depicts the balance between our pre-treatment covariates on either side of the discontinuity. Overall, balance is achieved: through t-tests, we are unable to reject the null hypothesis that the true difference in means is equal to 0 for any of the covariates except Age of Home. This lends support to the validity of our identification strategy. Covariates below the horizontal line were not included in the pre-analysis plan.

Covariate	Control Mean	Treatment Mean	p.value
Female Household Head	0.231	0.250	0.561
Birthplace	0.769	0.759	0.752
Age	42.020	41.950	0.940
Hindu	0.713	0.706	0.839
Scheduled Caste/Tribe	0.406	0.419	0.747
Education Level	0.316	0.311	0.894
Age of Home	13.890	16.840	0.005

For identification

- Instrument must not be weak.
  - Wald test F-statistic is 435.32 at p<.01
  - Reject null hypothesis that instrument is weak



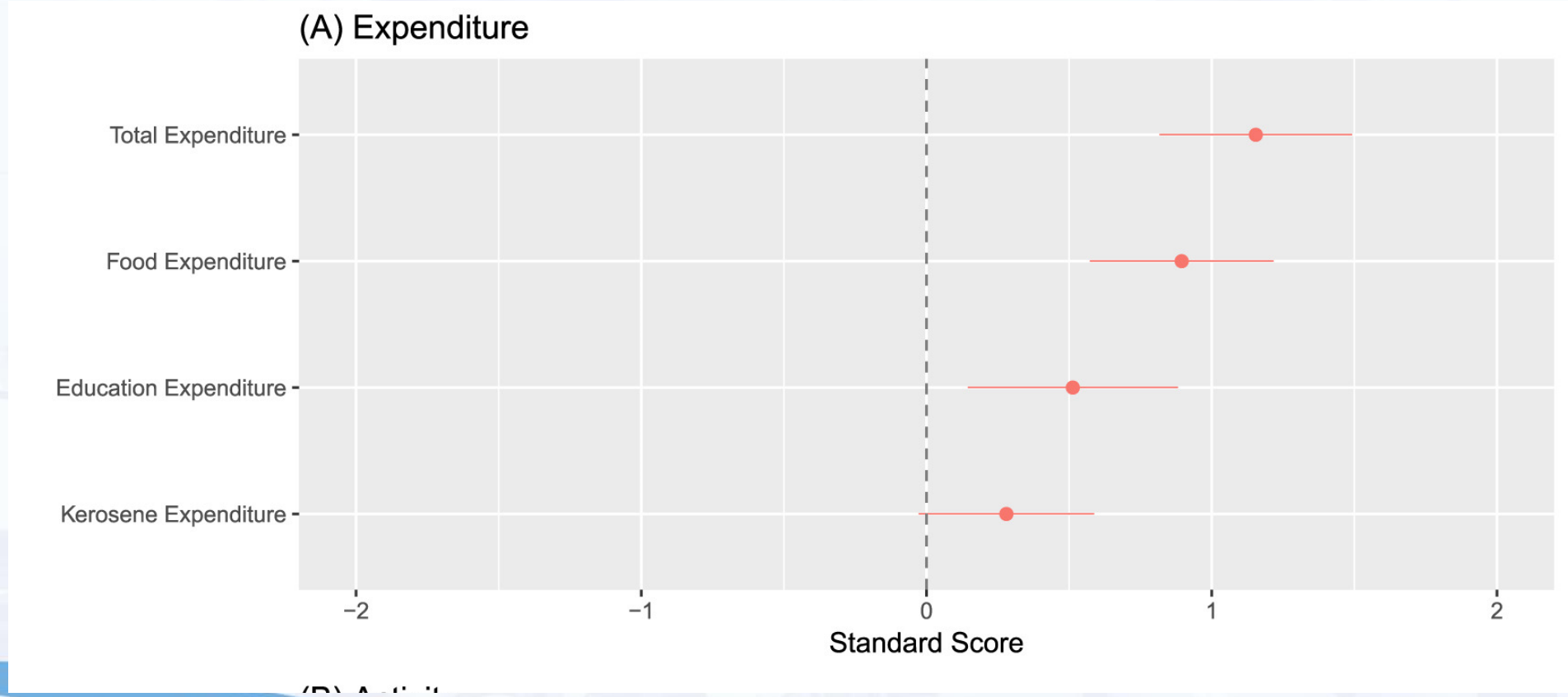

(A) Expenditure
Total Expenditure
Food Expenditure
Education Expenditure
Kerosene Expenditure
−2
−1
0
1
2
Standard Score

Table 4

This table shows the unscaled relationship between legal electrification and expenditures. We test the relationship further accounting for outliers in section A3.3. p-values reported are corrected within outcome families using the Benjamini-Hochberg method.

	Dependent variable:			
	Total Expenditure	Food Expenditure	Education Expenditure	Kerosene Expenditure
	(1)	(2)	(3)	(4)
Legal Electrification	4509.000***	3023.000***	655.000***	17.560*
	(672.500)	(555.800)	(239.900)	(9.849)
Constant	6865.000***	6031.000***	242.200	65.890***
	(1217.000)	(1028.000)	(413.700)	(15.420)
Control Variables	Yes	Yes	Yes	Yes
N Poles	152	152	152	152
Observations	685	686	686	686
Adjusted R^2	0.135	0.075	0.084	0.092

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

Standard errors clustered at pole level. Pole fixed effects included.

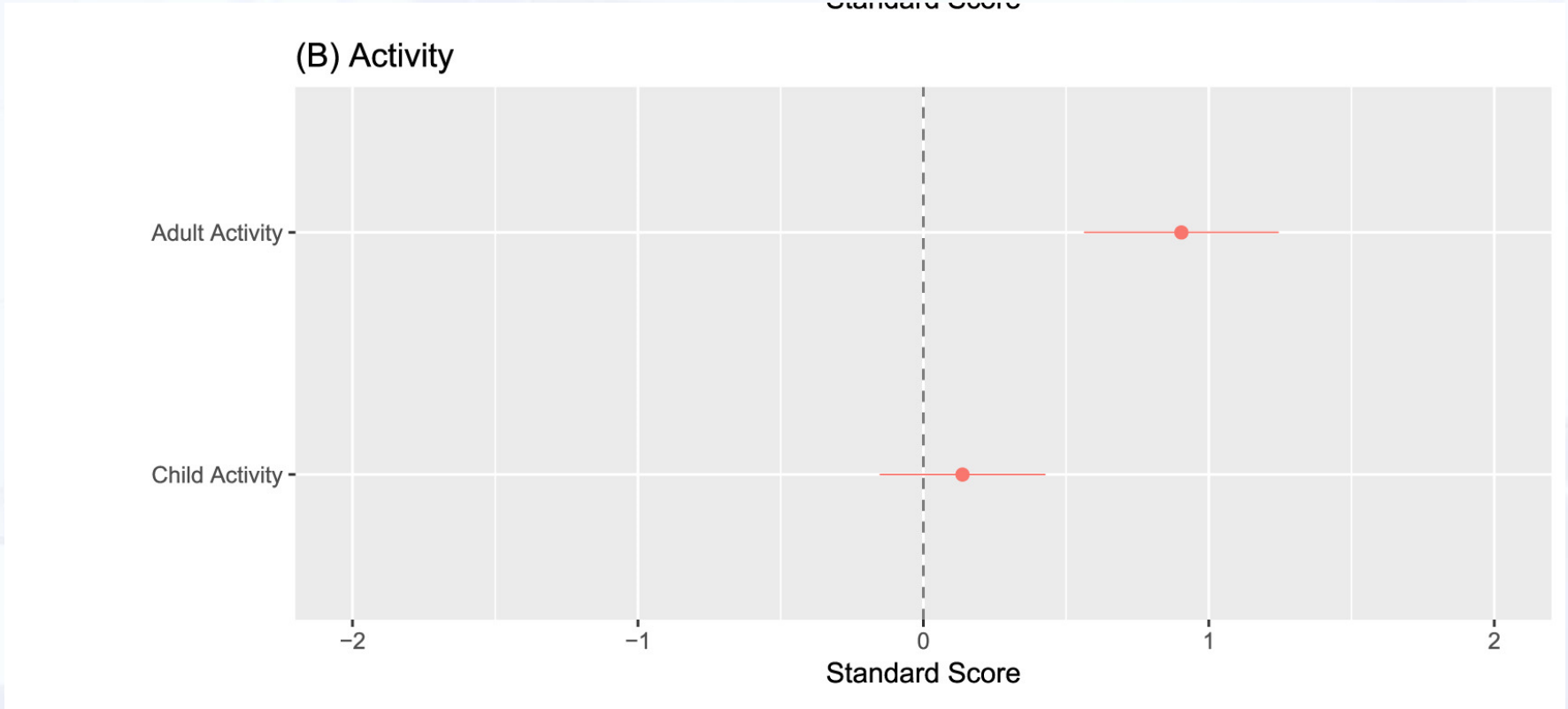

(B) Activity
Adult Activity
Child Activity
−2
−1
0
1
2
Standard Score

Table 5

This table shows the relationship between legal electrification and household activity. We find that legal electrification increases adult activities in the household but has no effect on the activities of children. p-values reported are corrected within outcome families using the Benjamini-Hochberg method.

	Dependent variable:	
	Adult Activity (1)	Child Activity (2)
Legal Electrification	0.640***	0.122
	(0.123)	(0.131)
Constant	3.789***	2.402***
	(0.217)	(0.246)
Control Variables	Yes	Yes
N Poles	152	152
Observations	686	686
Adjusted R^2	0.365	0.172

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

Standard errors clustered at pole level. Pole fixed effects included.

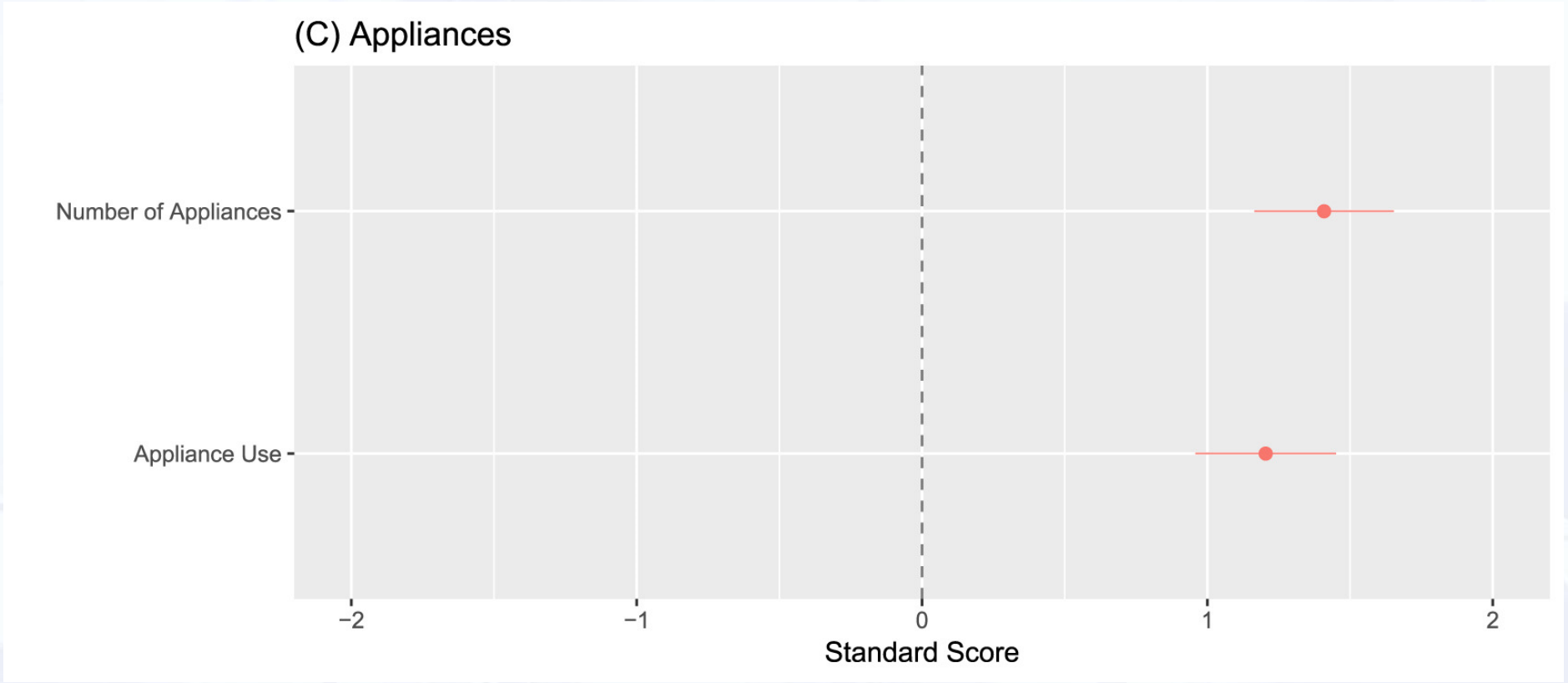

(C) Appliances
Number of Appliances
Appliance Use
−2
−1
0
1
2
Standard Score

Table 7

This depicts the relationship between legal electrification and appliances. Legal electrification has a clear, positive effect on appliance ownership and usage. p-values reported are corrected within outcome families using the Benjamini-Hochberg method.

	Dependent variable:	
	Number of Appliances (1)	Appliance Use (2)
Legal Electrification	3.760***	11.120***
	(0.333)	(1.159)
Constant	2.038***	3.550
	(0.686)	(2.258)
Control Variables	Yes	Yes
N Poles	152	152
Observations	686	686
Adjusted R^2	0.461	0.354

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

Standard errors clustered at pole level. Pole fixed effects included.

Miyazaki, T., (2018) "Examining the relationship between municipal consolidation and cost reduction: an instrumental variable approach" *Applied Economics*, 50(10), pp.1108-1121.

Background

- Local governments have undergone comprehensive consolidation
- Some advantages of consolidation –
  - Efficient and effective delivery of public services
  - Main purpose is to receive benefits associated with economies of scale
- Disadvantages of consolidation
  - May lessen competition among local governments
- Consequently, empirical works obtain inconclusive findings in terms of cost-reduction effects of local government consolidation
- Article explores cost-reduction effects of local government consolidation using data from Japanese municipalities
  - Advantage of study – examines cost efficiencies arising from consolidation using panel data on a large number of consolidation cases, data set contains relatively balanced treatment and control groups

Estimation strategy

- Endogeneity problem in the examination of consolidation decision and local costs
  - Municipalities that incur inefficiently large costs in the provision of public services are likely to consolidate.
  - Some municipalities consolidated to overcome financial difficulties - → reverse causality
  - Realization of economies of scale is a primary reason that local governments seek consolidation
  - Cost-inefficient governments therefore are more likely to seek consolidation
- Contribution of article – deal with endogeneity by exploiting policy changes in grant allocation for municipalities
  - Employ an IV approach, with the IV being the reduction in unconditional grants to small municipalities in 2002.
  - The IVs seem to fulfill the exclusion criterion - governmental policy reform could correlate closely with consolidations yet remain exogenous.

Endogeneity

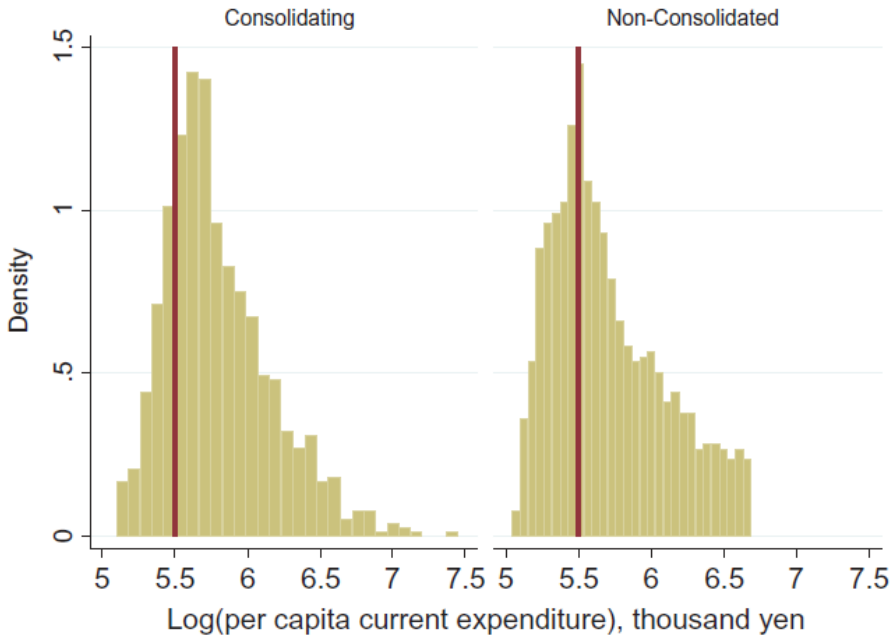


Figure 2. Log of per capita current expenditure for consolidating and non-consolidated municipalities, 2000.

Note: 'Consolidating' refers to municipalities that consolidated during 2001–2010; 'Non-Consolidated' refers to those that did not consolidate during this period. The solid vertical lines denote logged per capita current expenditure of 5.5.

Empirical model

- Econometric model is specified as:

$$LEXP_{it} = \beta_1 TREAT_{it} + \beta_2 TREND_{it} + x_{it}\beta_x + \tau_t + c_i + \varepsilon_{it}$$

  - $i = 1, \ldots N$
  - $t = 2000, 2005, 2010$
  - LEXP = log of current expenditure per capita
  - TREAT = dummy variable that takes on a value of one for consolidated municipalities in the post consolidation period. Treatment group consisted of municipalities that consolidated during the sample period
  - TREND= Number of years since consolidation is included in some regressions to capture that consolidation impact may depend on the years elapsed since consolidation
  - X is a vector of control variables (logs of population, population density, per capita taxable income, proportion of residents aged 65 and over, proportion of foreign residents)
  - c and tau represent individual and time effects

IV

- IV for municipality consolidation decision - changes in the law on unconditional grants
  - “correction coefficients” - an index used to allocate unconditional grants to conform to actual municipal fiscal needs
  - Framework for the correction was modified in 1998 and 2002.
  - Reduction in unconditional grants arising from the reform of the correction coefficients has induced consolidation
  - Employ three categorical variables that take the value one for non-consolidated municipalities with populations of <1000, 1000-4000 and 4000-8000, respectively for the years after 2002.

Table 4. Summary statistics for the treatment and control groups, 2000, 2005 and 2010.

	Treatment Group					Control Group				
	Year 2000	Year 2005	Year 2010	Diff.	Diff.	Year 2000	Year 2005	Year 2010	Diff.	Diff.
Variable	(1)	(2)	(3)	(2) − (1)	(3) − (2)	(4)	(5)	(6)	(5) − (4)	(6) − (5)
Per capita current expenditure	252.32	276.20	306.80	23.89	30.59	240.54	247.25	274.77	6.71	27.52
	(58.2)	(68.8)	(67.4)			(72.7)	(68.9)	(71.6)		
Population	98.50	86.68	81.60	−11.82	−5.08	41.41	42.66	46.07	1.26	3.40
	(130.8)	(119.4)	(108.7)			(69.3)	(72.0)	(75.4)		
Population density	0.41	0.40	0.375	−0.01	−0.02	1.02	1.06	1.15	0.04	0.09
	(0.71)	(0.69)	(0.66)			(1.85)	(1.92)	(2.02)		
Per capita taxable income	3387.3	3165.7	2870.8	−221.54	−294.94	3643.3	3437.6	3139.4	−205.76	−298.13
	(364.9)	(354.1)	(327.4)			(515.4)	(495.8)	(460.8)		
Proportion of population	18.88	21.72	24.73	2.84	3.01	16.39	19.36	22.59	2.97	3.23
Aged 65 years or over	(4.30)	(4.50)	(4.23)			(4.68)	(4.57)	(4.29)		
Proportion of foreigners	0.80	0.98	0.98	0.18	−0.01	0.86	1.05	1.07	0.19	0.02
	(0.66)	(0.86)	(0.79)			(0.76)	(0.92)	(0.88)		
Observations	443	567	564			1157	1110	1024		

The table reports summary statistics for the treatment and control groups. The upper values in each row indicate the means, and the values within parentheses are the standard deviations. Pre-consolidation values are calculated by summing the values of the relevant consolidation partners.

Table 5. Estimation of the cost-reduction effect of consolidation.

	No trend		Linear trend		Quadratic trend	
	FE model	FEIV model	FE model	FEIV model, baseline	FE model	FEIV model
	(1)	(2)	(3)	(4)	(5)	(6)
TREAT	0.035***	0.075***	0.061***	0.225***	0.054***	0.200***
	(0.005)	(0.016)	(0.005)	(0.077)	(0.005)	(0.066)
TREND			−0.010***	−0.051**		
			(0.001)	(0.025)		
$(TREND)^2$					−0.001***	−0.008**
					(0.000)	(0.004)
LPOP	−0.087***	−0.089***	−0.086***	−0.080**	−0.089***	−0.091***
	(0.026)	(0.022)	(0.026)	(0.032)	(0.026)	(0.033)
LPOPDEN	−0.144***	−0.123***	−0.166***	−0.223***	−0.160***	−0.208***
	(0.018)	(0.020)	(0.018)	(0.055)	(0.018)	(0.050)
LTAXINC	0.072	−0.075	0.079	−0.099*	0.079	−0.095*
	(0.069)	(0.049)	(0.069)	(0.056)	(0.069)	(0.056)
POP65	−0.003***	0.001	−0.004***	−0.002	−0.004***	−0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
FOREIGNER	0.005*	0.006**	0.004	0.002	0.004	0.003
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Hausman test		131.379		425.121		24.549
p-Value		[0.000]		[0.000]		[0.006]
Partial *F*-statistics		238.990				
Partial *F*-statistics 1				168.022		168.022
Partial *F*-statistics 2				61.562		53.668
Hansen *J*-statistics		0.505		5.728		5.760
p-Value		[0.777]		[0.220]		[0.218]
Adjusted R^2	0.595		0.606		0.604	
Centred R^2		0.556		0.358		0.329
Observations	4865	4865	4865	4865	4865	4865

The table reports estimation results of the cost-reduction effect of consolidation. Robust standard errors and *p*-values are, respectively, in parentheses and brackets. *, ** and *** denote significance at 10%, 5% and 1%, respectively. FE: fixed-effect; FEIV: fixed-effect instrument variable. Every estimation includes year dummies for 2005 and 2010. The Hausman test examines endogeneity between the FE and FEIV models. Partial *F*-statistics are used to test whether instruments correlate with the endogeneity variable in the first-stage regression of the IV (2SLS) estimation. Partial *F*-statistics 1 (2) represents the *F*-statistics of the first-step regression for *TREAT* (*TREND*). Hansen *J*-statistics are employed to test correlation between instrumental variables and errors.

Results

- Absence of trend-effect in post-consolidation municipal costs could lead to specification errors
- Endogeneity between consolidation decision and size of expenditure causes underestimation of the consolidation treatment and trend.
    - Treatment effects are small in the FE regressions relative to the FEIV and the trend effects are less negative in the former than in the latter.

References

- Capuno, J. , A.D. Kraft and O. O'Donnell (2021) "Effectiveness of clinic-based cardiovascular disease prevention: A randomized encouragement design experiment in the Philippines" *Social Science and Medicine* 283
- Gertler P, Martinez S, Premand P, Rawlings L, Vermeersch C (2016). *Impact Evaluation in Practice,* 2nd Edition. Washington, DC: Inter-American Development Bank and the World Bank
- Glewwe P, Todd P. (2022) *Impact evaluation in international development: theory, methods, and practice.* World Bank Publications; 2022
- Khandker, S., G. Koolwal, and H Samad. (2010). *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, DC, The World Bank.
- Miyazaki, T., (2018) "Examining the relationship between municipal consolidation and cost reduction: an instrumental variable approach" *Applied Economics*, 50(10), pp.1108-1121.
- Thomas, D.R., Harish, S.P., Kennedy, R. and Urpelainen, J., (2020). "The effects of rural electrification in India: An instrumental variable approach at the household level" *Journal of Development Economics,* 146, p.102520