Georg-August-Universität
Göttingen

Faculty of Economic Sciences
Chair of International Economic Policy
Prof. Dr. Krisztina Kis-Katos

Econometric analysis

Elías Cisneros 2020 Bachelor seminar

Agenda

- **Understanding empirical results**
  - Estimation equation
  - Estimation results
- **Reporting the results in your seminar paper**
  - Data and measurement
  - Estimation strategy
  - Results

Estimation equation

$$D_i = \alpha + \beta T_i + \mathbf{X}_i \gamma + \varepsilon_i$$

Outcom

$$\log(y_{wpit}) = \beta NumDistrictsInProv_{pit} + \mu_{wpi} + \eta_{wit} + \varepsilon_{wpit},$$

$$\delta_i = Y_i(1) - Y_i(0)$$

Estimation equation

$$D_i = \alpha + \beta T_i + \mathbf{X}_i \gamma + \varepsilon_i$$

Dependent / outcome variable

explanatory variables / influencing factors
consist out of variable of interest and control variables

Estimator of interest

What do the variables stand for?

$$D_i = \alpha + \beta T_i + \mathbf{X}_i \gamma + \varepsilon_i$$

D_i	is household / plot / farm i's forest loss
T_i	is a treatment intervention/ policy / exposure; dummy variable being abducted
X_i	is a vector of individual characteristics (age, no. of school years,…, plot size, property rights, crop type, elevation, travel distance, …)
α	intercept
ε_i	is a stochastic error

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

What is β ?

$$D_i = \alpha + \beta T_i + \mathbf{X}_i \gamma + \varepsilon_i$$

$H_0 : \beta = 0$? $H_A : \beta \neq 0$

D_i	is household / plot / farm i's forest loss
T_i	is a treatment intervention / policy / exposure; dummy variable being abducted
X_i	is a vector of individual characteristics (age, no. of school years,..., plot size, property rights, crop type, elvation, travel distance, ...)
α	intercept
ε_i	is a stochastic error

So what effect does β measure?

$$\frac{\partial D}{\partial T} = \beta$$

T ↑ (1 unit) → D ↑ by β units on average

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Typical structure of a table

Dependent variable

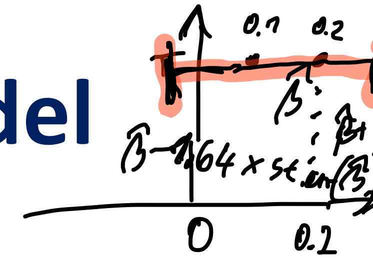
Explanatory variables

Coefficients
St.err

Addiotnal info
Fixed effects

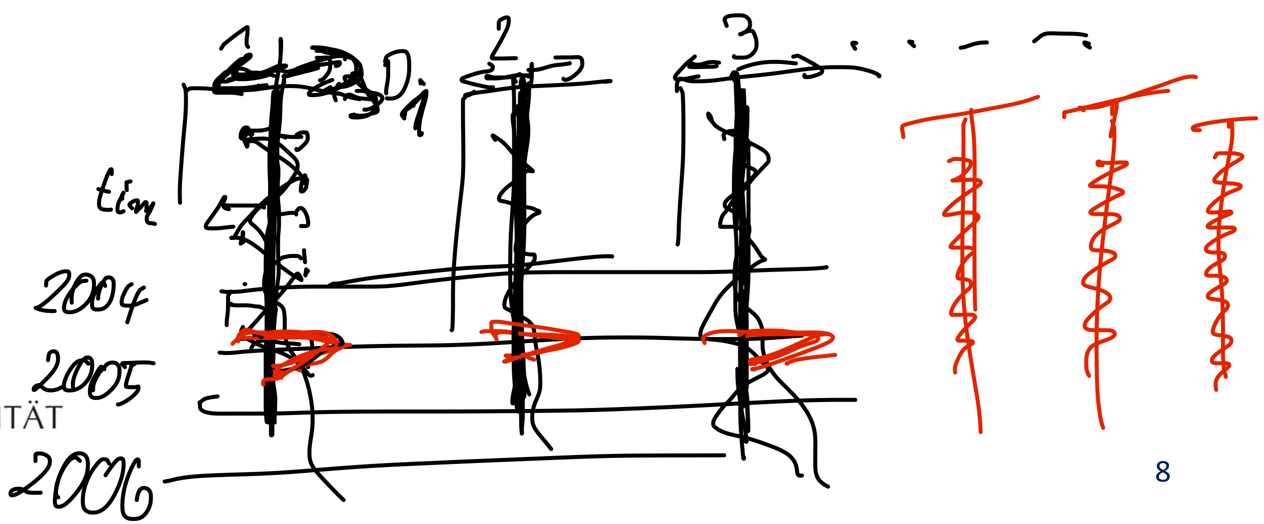
Dependent variable	Outcome 1	
	Model name 1	Model name 2
	(1)	(2)
T (treatment) variable name	0.100***	0.200
	(0.002)	(0.150)
X^1 control variable name 1	0.300	0.200
	(0.400)	(0.150)
X^2 control variable name 2	0.250	0.200
	(0.174)	(0.150)
Individual fixed effects	No	Yes
Further controls	No	Yes
Observations controls	150	148
Adj. R^2	0.02	0.40

Special cases: Fixed effects model

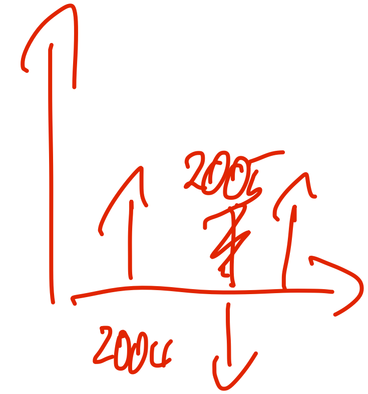


- Unobserved variables that we cannot control for
- Using fixed effects to absorb potential time-invariant effects
- estimation exploits only within variation

$$D_{it} = \beta T_{it} + \mathbf{X}_{it}\gamma + \lambda_i + \delta_t + \varepsilon_{it}$$



Alesina et al. (2019)



District splits, ethnic fragmentation and deforestation

In this subsection we provide a more rigorous test for the main prediction of our simple theoretical framework. In particular, we regress the log of deforestation (f) on the time-varying level of ethnic diversity (EF) while controlling for district-level fixed-effects u_i:

$$f_{ipt} = \beta EF_{ipt} + \gamma X_{ipt} + \delta ysplit_{ipt} + d_t + u_i + d_t \times v_p + \varepsilon_{ipt}, \tag{2}$$

where the coefficient β identifies the effect of a change in the index of ethnic fractionalization EF on the level of deforestation. District fixed effects control for time-invariant, district-specific characteristics. We also include a dummy for the year of splitting, $ysplit_{ipt}$. Robust standard errors are clustered at the district level. Since EF

Alesina et al. (2019)

$$f_{ipt} = \beta EF_{ipt} + \gamma X_{ipt} + \delta ysplit_{ipt} + d_t + u_i + d_t \times v_p + \varepsilon_{ipt},$$

Dependent variable

Explanatory variables

Coefficients
St.err

Addiotnal info
Fixed effects

TABLE 6
ETHNIC FRACTIONALIZATION AND DEFORESTATION

	(1)	(2)	(3)	(4)
EF	1.18**	1.428**		
	(0.484)	(0.532)		
EF (sum of L0 to L1)			1.711***	
			(0.529)	
EF (sum of L0 to L2)				2.379***
				(0.732)
Controls	No	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Province-by-year fixed effects	Yes	Yes	Yes	Yes
Observations	4044	3937	3937	3611
Districts that split	95	95	95	95
Total number of districts	337	331	331	331

Notes
Standard errors clustered at the district level in parentheses. The dependent variable is the log of square metres deforested. Each observation is a district based on 2000 district boundaries. Controls include a binary variable indicating the year of splitting, district-level GDP, population, government expenditure and expenditure on infrastructure. The coefficient of '*EF* (sum of L0 to L1)' is given by the cumulative sum of the contemporaneous and lagged effect. '*EF* (sum of L0 to L2)' includes also the second lag of *EF*.
*, **, *** indicate $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Excursion: Matching

Slides on matching

Rasolofoson et al. (2018)

Impacts of forests on children's diet in rural areas across 27 developing countries

At the population level, we can define the ATT (here, treatment is high exposure to forest)

obs. gonks Treated group

$$\begin{aligned} \text{ATT} &= E[Y_i(1) - Y_i(0)|D_i = 1] \\ &= E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] \end{aligned} \qquad (2)$$

where $E[.]$ is the expectation operator from probability theory and $|D_i = 1$ means conditional on the household being under high exposure

to forest (that is, being forest household). In other words, the ATT is the difference between the expected dietary diversity of forest households under high exposure to forest, $E[Y_i(1)|D_i = 1]$, and the expected dietary diversity of these same forest households were they under lack of exposure to forest, $E[Y_i(0)|D_i = 1]$. The former, which is the average dietary diversity of forest households, is observed. The latter, which is the dietary diversity of forest households, had they not been exposed to forest, is unobserved (the counterfactual).

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Rasolofoson et al. (2018)

Impacts of forests on children's diet in rural areas across 27 developing countries

The CMIA postulates that, conditional on comparable observed confounders, X, between forest and nonforest households, the expected dietary diversity of the nonforest households under lack of exposure to forest, $E[Y_i(0)|D_i = 0]$, represents the unobserved counterfactual average dietary diversity, $E[Y_i(0)|D_i = 1]$

$$E[Y_i(0)|D_i = 1, X] = E[Y_i(0)|D_i = 0, X] = E[Y_i(0)|X] \quad (3)$$

counterfactual

Observations of the control

Rasolofoson et al. (2018)

$$Y_i = \alpha + \beta D_i + \varepsilon_i$$

where $i \in$ set of matched observations based on X

Dependent variable

Explanatory variables

Coefficients

St.err (null hypo.)

Addiotnal info

Fixed effects

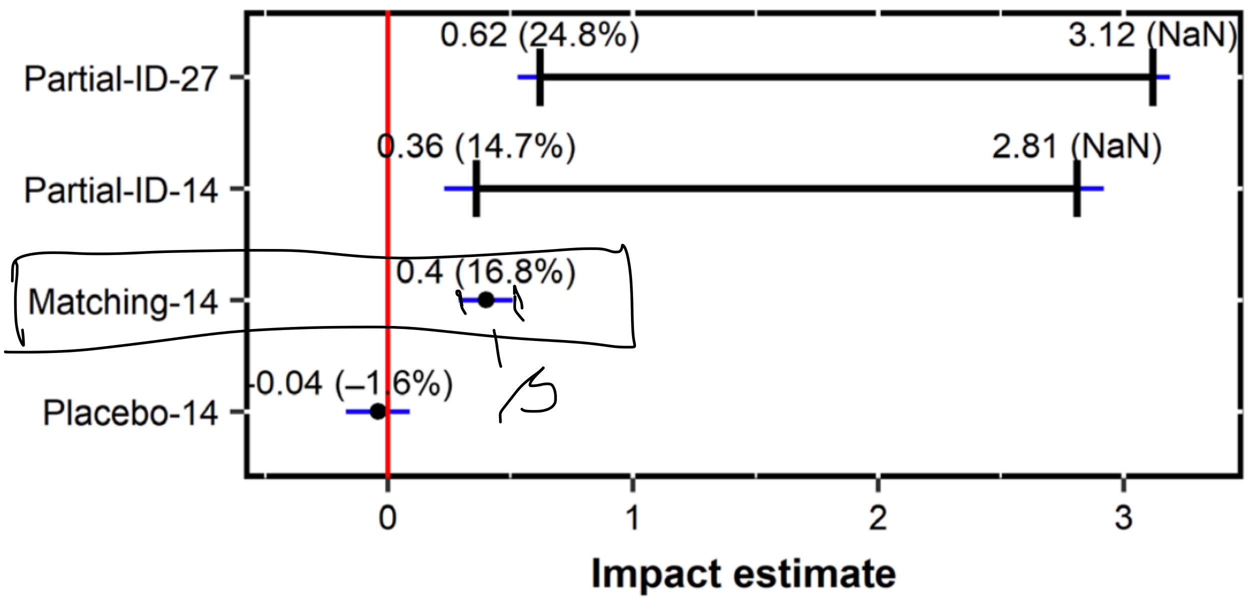


Fig. 3. Estimated impacts of forests on dietary diversity. Partial-ID-27, partial identification for 27 countries; Partial-ID-14, partial identification for 14 African countries; Matching-14, matching design for 14 African countries; Placebo-14, placebo test for the matching design for 14 African countries. Values in parentheses, impact expressed in percent of the average dietary diversity of nonforest households. NaN, not a number (undefined). Blue bars, 95% confidence intervals.

Alix-Garcia et al. (2015)

Only One Tree from Each Seed?
Environmental Effectiveness and Poverty Alleviation in Mexico's Payments for Ecosystem Services Program†

Alix-Garcia et al. (2015)

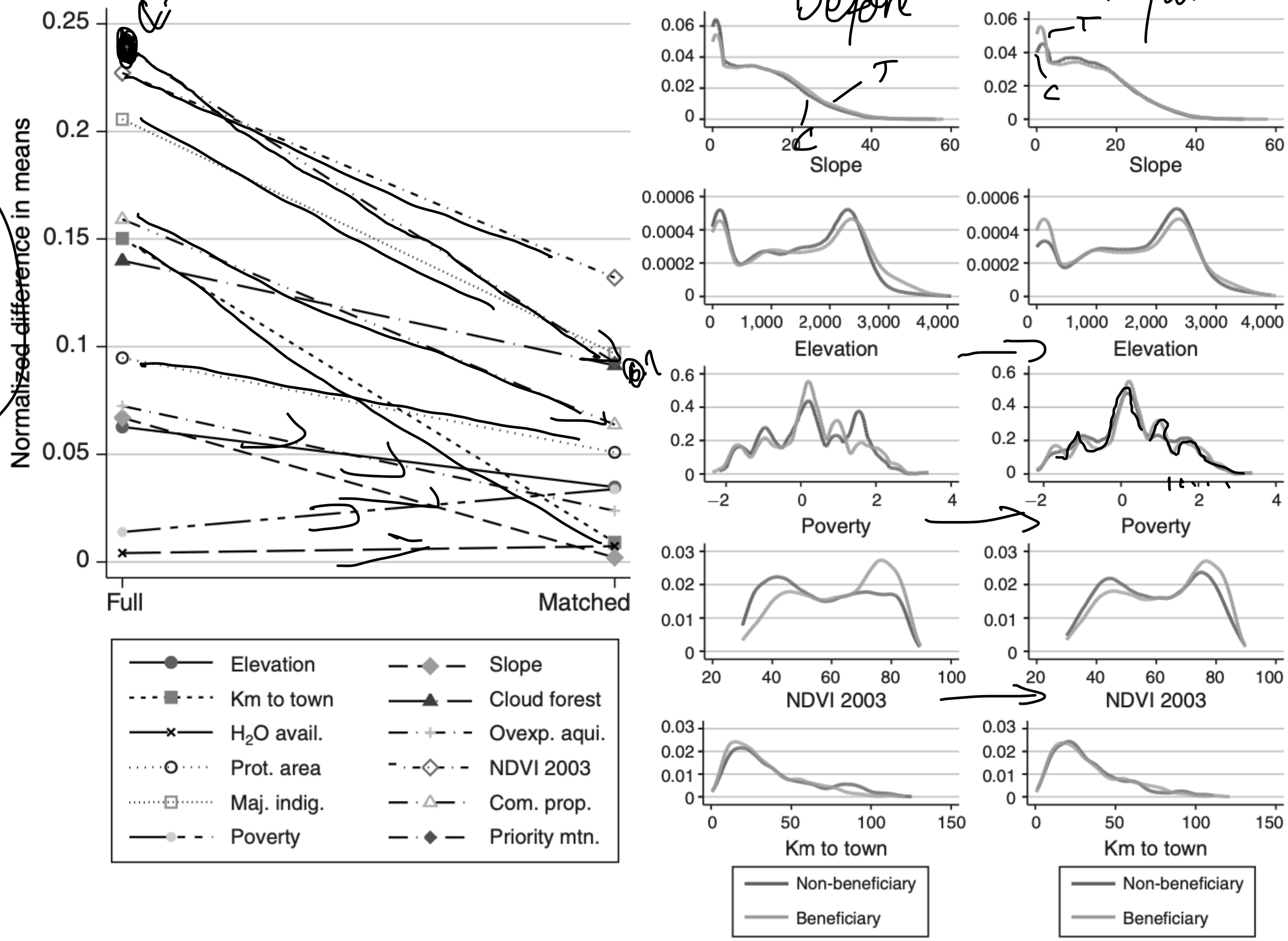


FIGURE 3. CHANGES IN NORMALIZED DIFFERENCES AND DISTRIBUTIONS AFTER MATCHING

Notes: Left side of figure shows normalized difference in means between beneficiary and non-beneficiary points before and after covariate matching. Right side of figure shows covariate distributions for beneficiary and non-beneficiary points before (left) and after (right) covariate matching.

Alix-Garcia et al. (2015)

We estimate panel regressions with point-level fixed effects on the matched subsample using the following specification:

$$(1) \quad MNDVI_{ipst} = \beta_0 + \beta_1 beneficiary_{it} + \delta' \mathbf{rainfall}_{it} + \alpha_{st} + \alpha_i + \varepsilon_{ipst},$$

where *MNDVI* is the mean dry season NDVI value for point i in parcel p, state s, and year t. The variable *beneficiary* is equal to 1 if the point was enrolled in the program in the previous year's cohort; β_1 is the average program impact.[20] To control for rainfall and hurricanes ($\mathbf{rainfall}_{it}$), we include the natural logarithms of dry season rainfall and of rainfall in the other months prior to the dry season. To control for hurricanes, we also include the standard deviation of rainfall across the year, and a dummy variable for being in the top tenth percentile of rainfall during the hurricane season (October/November). State-year fixed effects (α_{st}) control for possible economic shocks to states in each year and point-level fixed effects (α_i) control for unobservable fixed land characteristics. Standard errors are clustered at the parcel level to

Alix-Garcia et al. (2015)

Dependent variable

Explanatory variables

Coefficients
St.err (null hypo.)

Addiotnal info
Fixed effects

TABLE 4—IMPACTS OF PSAH ON NDVI

Dependent variable	Annual mean dry season NDVI (points data)					
	Change in levels (1)	Change in trend (2)	Change in levels, years in program (3)	Change in levels (4)	Change in trend (5)	Change in levels, years in program (6)
Beneficiary	0.1863*** (0.0721)		0.1396 (0.0880)	0.2455*** (0.0737)		0.1745** (0.0874)
Beneficiary × Time		0.0265** (0.0127)			0.0352*** (0.0125)	
Beneficiary × Years in program			0.0223 (0.0341)			0.0344 (0.0347)
Point FE	Yes	Yes	Yes			
Parcel FE				Yes	Yes	Yes
Observations total	196,164	196,164	196,164	174,368	174,368	174,368
Observations points	21,796	21,796	21,796	21,796	21,796	21,796
Observations parcels	3,495	3,495	3,495	3,495	3,495	3,495
R^2	0.49	0.49	0.49	0.68	0.68	0.68
Effect size (percent of five year trend)	40–51	28–36	54–69	39–66	28–47	55–93

Notes: Columns 1–6: Point or parcel-level fixed effects model (equation (1)). Columns 1–6 all include state-year fixed effects and rainfall controls. Robust standard errors clustered at the parcel level in parentheses. Dependent variable is mean dry season NDVI (ranges from 0 to 100). Regressions use data from the 21,796 points within program beneficiaries ($N = 17{,}307$) and matched rejected applicants ($N = 4{,}489$) (Table 2, columns 2 and 4). Regressions 1–3 use NDVI outcomes from 2003–2011. Regressions 3–6 use NDVI outcomes from 2004–2011 and include NDVI 2003 and other point-level covariates shown in Table 2 as controls. The effect sizes use counterfactual trends of NDVI loss: among matched controls, we find an average annual loss of NDVI between 2004 and 2011 of −0.0731 with point fixed effects and −0.0748 using parcel fixed effects. Using all initially forested points, we find a trend of −0.0935 with point fixed effects and −0.1250 with parcel fixed effects.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Reporting in the seminar paper

Data and measurement

- Explain the study setting
  - Geographical coverage
  - Time period
  - Event
- Describe the data sources
- Describe the main variables and how they are measured

Estimation strategy

- State the estimation equation
  - Reformulate if necessary
  - Formulate if only implicit (matching)
- State what the variables stand for
- State the estimation strategy (and why it is used)

Results

- Include one regression table
- State the different regression models and the magnitude and significance of the estimated effects of interest
- Interpret the effects
- (Verbally explain their other findings)

Further questions?

Student Question 1: Jayachandran

- Table 3
- Model
- Data
- Table 3 Result interpretation

- Table S9
- Model
- Result interpretation

- Table S3
- Model

Student Question 1: Jayachandran

Table 3. Effect of the PES program on tree cover. All regressions and means are weighted by the proportion of available tree-classification data for the observation. All columns include subcounty fixed effects and the four village-level baseline variables used to balance the randomization. Columns 2, 3, 5, and 6 also control for dummy variables for the date of the baseline satellite image, and columns 2 and 3 control for 1990 and 2010 area covered by photosynthetic vegetation within the village polygon and in aggregate in PFO land circles for the village; columns 5 and 6 control for 1990 and 2010 area covered by photosynthetic vegetation within the village polygon and in the PFO's land circle. Standard errors are heteroskedasticity-robust in columns 1 to 3 and clustered by village in columns 4 to 6. Significance: *P < 0.10, **P < 0.05, ***P < 0.01.

	Village boundaries			PFO-level land circles		
	ΔTree cover (ha)	ΔTree cover (ha)	ΔLog of tree cover	ΔTree cover (ha)	ΔTree cover (ha)	ΔIHS of tree cover
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment group	5.549*	5.478**	0.0521**	0.245**	0.267**	0.0447*
	[2.888]	[2.652]	[0.021]	[0.110]	[0.106]	[0.023]
Control group	–13.371	–13.371	–0.095	–0.349	–0.349	–0.073
Control variables	No	Yes	Yes	No	Yes	Yes
Observations	121	121	121	995	995	995

Student Question 1: Jayachandran

Table S9: **Heterogeneous effects of the PES program on tree cover.** Standard errors are clustered by village. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The outcome variable in all columns is Δ*Tree cover (ha)*, which is measured in hectares. All regressions are weighted by the proportion of available tree-classification data. All columns include subcounty fixed effects, the four village-level baseline variables used to balance the randomization, dummy variables for the baseline satellite date, and 1990 and 2010 photosynthetic vegetation in the PFO's land circle and in the village boundary. In column 8, predicted tree loss is the predicted value from the regression reported in Table S3, column 3.

	Heterogeneous treatment effects on ΔTree cover (ha) by:							
	Above-median tree cover in land circle (1)	% of land circle with tree cover (2)	Cut any trees in the last 3 years (3)	Cut trees to clear land for cultivation (4)	Cut trees for timber products (5)	Cut trees for emergency/lumpy expenses (6)	IHS of total revenue from cut trees (7)	Predicted change in tree cover (8)
Treat × Characteristic	0.472**	2.054**	0.424**	0.025	0.342**	0.410*	0.122***	-0.686**
	[0.204]	[0.929]	[0.163]	[0.162]	[0.172]	[0.227]	[0.044]	[0.292]
Treated	0.020	-0.174	-0.094	0.265**	0.016	0.140	-0.014	-0.008
	[0.072]	[0.145]	[0.124]	[0.118]	[0.128]	[0.095]	[0.079]	[0.088]
Characteristic	-0.577***	-2.659***	-0.329**	0.076	-0.335**	-0.413**	-0.105***	0.532
	[0.183]	[0.804]	[0.134]	[0.120]	[0.140]	[0.198]	[0.039]	[0.396]
Observations	995	995	993	995	995	995	993	994

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Student Question 2: Jayachandran

Table S3: **Correlates of program enrollment in treatment group.** Standard errors are clusterɩ by village. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All columns include subcoun fixed effects, and the first three columns include the four village-level baseline variables used to balan the randomization. Missing independent variables have been imputed with the sample mean, and indicat variables for having a missing value are included in the regression. IHS denotes the inverse hyperbolic si function.

	Enrolled (1)	Enrolled (2)	ΔTree cover (ha) (3)	Enrolled (4)
Household head's age	0.002 [0.001]		0.003 [0.003]	
Household head's years of education	0.004 [0.005]		0.001 [0.014]	
IHS of self-reported land area (ha)	0.055* [0.028]	0.059** [0.024]	-0.318** [0.135]	
Self-reported forest area (ha)	-0.004 [0.006]		-0.073 [0.054]	
Cut any trees in the last 3 years	0.036 [0.094]		0.038 [0.163]	
Cut trees to clear land for cultivation	0.028 [0.057]		0.046 [0.140]	
Cut trees for timber products	0.090 [0.076]		0.049 [0.178]	
Cut trees for emergency/lumpy expenses	-0.129*** [0.041]	-0.099** [0.040]	-0.340* [0.195]	
IHS of total revenue from cut trees	-0.010 [0.010]		-0.049 [0.029]	
Rented any part of land	-0.046 [0.067]		0.004 [0.175]	
Dispute with neighbor about land	0.051 [0.045]		-0.063 [0.109]	
Involved in any environmental program	-0.014 [0.079]		0.216* [0.125]	
Agree: Deforestation affects the community	0.032 [0.039]		-0.037 [0.086]	
Agree: Need to damage environ. to improve life	-0.239*** [0.075]	-0.200*** [0.068]	-0.403 [0.346]	
Tree cover in PFO land circle (ha)	-0.003** [0.001]	-0.003** [0.001]		
% change in vegetation in PFO land circle, 1990–2010	0.285 [0.348]		1.787** [0.886]	
Predicted change in tree cover				-0.024 [0.034]
Sample	Treatment group	Treatment group	Control group	Treatment group
R^2	0.140	0.115	0.280	0.067
Observations	564	564	486	564

GEORG-AUGUST-UNIVERS
GÖTTINGEN

Student Question 2: Burgess et al.

$$\mathbf{E}(deforest_{dit}) = \mu_{di}\exp(\beta PCOilandGas_{dit} + \gamma Numdistricts_{dit} + \eta_{it}),$$

TABLE VI

SUBSTITUTES OR COMPLEMENTS? EFFECTS OF DISTRICT-LEVEL OIL AND GAS REVENUES ON DEFORESTATION AS MEASURED WITH SATELLITE DATA

Variables	(1) All forest	(2) Production/ Conversion	(3) Conservation/ Protection
Panel A			
Oil and gas revenue per capita	−0.00316**	−0.00284*	−0.00597**
	(0.00160)	(0.00165)	(0.00252)
Observations	6464	3064	3400
Panel B: lags			
Oil and gas revenue per capita	−0.00492***	−0.00432**	−0.0113***
	(0.00186)	(0.00190)	(0.00257)
Lag 1	0.000652	8.87e-05	0.00561***
	(0.00103)	(0.00126)	(0.00113)
Lag 2	0.00112	0.00132	0.000731
	(0.00130)	(0.00151)	(0.00138)
Lag 3	0.00519***	0.00530***	0.00574
	(0.00163)	(0.00160)	(0.00372)
Sum of L0–L3	0.00205	0.00240	0.000768
	(0.00134)	(0.00154)	(0.00195)
Joint p	<0.001	<0.001	<0.001
Sum of L0–L3 = L0 effect *p*-value	<0.001	<0.001	<0.001
Observations	6464	3064	3400

Notes. The forest data set has been constructed from MODIS satellite images, as described in Section III.C. The Production and Conversion zones are those in which legal logging can take place, while the Conservation and Protection zones are those in which all logging is illegal. The dependent variable is the number of forest cells deforested in the district-zone-year. A unit of observation is a 1990-borders district-forest zone. The oil and gas revenue per capita variable reports the value of per capita revenue from oil and gas extraction at the district level in U.S. dollars. A unit of observation is a 2008-borders district-forest zone. In Panel B, we include the oil and gas revenue variable and three lags of the oil and gas revenue variable; the coefficient reported as sum of L0–L3 is the sum of the coefficients on the oil and gas revenue variable and the first three lags. *p*-values are reported for tests of joint significance of the contemporaneous and lagged oil and gas revenue variables (joint p) and a test of whether the sum of the coefficients on the contemporaneous oil and gas revenue variable and the first three lags is equal to the contemporaneous coefficient (sum of L0–L3 = L0). All regressions include district-by-forest zone and island-by-year fixed effects and the number of districts the 1990 district has split into by year *t* (and three lags of this variable in Panel B), where a district is counted as having split when it reports receiving its own oil and gas revenue. Robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

$$\mathbf{E}(deforest_{dit}) = \tag{8}$$
$$\mu_{di}\exp\left(\begin{array}{c}\beta PCOilandGas_{dit} + \delta PostElection_{dit} \\ +\phi PCOilandGas \times PostElection_{dit} + \gamma Numdistricts_{dit} + \eta_{it}\end{array}\right).$$

TABLE VII

EFFECTS OF DISTRICT-LEVEL OIL AND GAS REVENUES ON DEFORESTATION BEFORE AND AFTER DIRECT ELECTIONS

Variables	(1) All forest	(2) Production/ Conversion	(3) Conservation/ Protection
Oil and gas revenue per capita	−0.00523***	−0.00457***	−0.0122***
	(0.00143)	(0.00159)	(0.00174)
Postelection	0.0218	0.0240	0.0299
	(0.110)	(0.118)	(0.217)
Oil and gas × Postelection	0.00175*	0.00147	0.00517***
	(0.000989)	(0.000976)	(0.00180)
Oil + Oil * Postelection	−0.00348***	−0.00310**	−0.00698***
	(0.00129)	(0.00140)	(0.00134)
p-value	0.00128	0.0161	<0.001
Observations	6403	3037	3366

Notes. The forest data set has been constructed from MODIS satellite images, as described in Section III.C. The Production and Conversion zones are those in which legal logging can take place, while the Conservation and Protection zones are those in which all logging is illegal. The dependent variable is the number of forest cells deforested in the district-zone-year. A unit of observation is a 1990-borders district-forest zone. The oil and gas revenue per capita variable reports the value of per capita revenue from oil and gas extraction at the district level in U.S. dollars. A unit of observation is a 2008-borders district-forest zone. The postelection variable is a dummy capturing whether the new direct election for district heads has taken place. All regressions include district-by-forest zone and island-by-year fixed effects and the number of districts the 1990 district has split into by year t, where a district is counted as having split when it reports receiving its own oil and gas revenue. Robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

Student Question 2: Canaviere-Bacarreza and Hanauer (2013) in comparison with Hanauer and Canaviere-Bacarreza (2015)

Student Question 2:

Table 4. *Results from primary and ancillary analyses. For each method Y(T = 1) and Y(T = 0) represent the average observed and counterfactual measures of poverty for protected municipalities, respectively. Each treatment column is calculated as Y(T = 1) − Y(T = 0)*

Method	Poverty index			NBI		
	$Y(T=1)$	$Y(T=0)$	Treatment	$Y(T=1)$	$Y(T=0)$	Treatment
Naïve difference in means	−1.33	−0.451	−0.838***	76.18	86.29	−10.11***
	[56]	[252]	{0.014}	[53]	[242]	{0.005}
Regression on raw data	−1.33	−0.828	−0.502***	76.18	81.7	−5.52***
	[56]	[268]	(0.098)	[53]	[258]	(1.17)
Regression dropping marginal	−1.33	−0.795	−0.535***	76.18	81.79	−5.62***
	[56]	[252]	(0.099)	[53]	[242]	(1.2)
Post-match frequency weighted regression	−1.33	−0.836	−0.494***	76.18	78.81	−2.63
	[56]	[41]	(0.106)	[53]	[45]	(1.7)
Genetic matching	−1.33	−0.805	−0.525***	76.18	81.16	−4.99
	[56]	[56]	(0.142)	[53]	[53]	(3.67)
Genetic matching, calipers = 1sd	−1.07	−0.511	−0.56***	79.04	81.51	−2.47
	[49]	[49]	(0.147)	[47]	[47]	(1.55)

[Number of observations].
(Standard errors).
{*P*-value}.
*** Treatment effect estimates are different from zero at 1% levels.

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Student Question 2: Hanauer and Canaviere-Bacarreza (2015)

Table 2. Results from first stage matching for deforestation and poverty samples. Γ represents maximum Gamma at which estimates are still significant at 10% level according to sensitivity to unobserved heterogeneity test. See the electronic supplementary material for more detail. Asterisks (*, **, ***) represent significance at the 10%, 5% and 1% level, respectively.

	deforestation			socioeconomic		
estimator	protected $Y_1 \mid D = 1$	counterfactual $\hat{Y}_0 \mid D = 1$	*ATT*	protected $Y_1 \mid D = 1$	counterfactual $\hat{Y}_0 \mid D = 1$	*ATT*
matching	0.009	0.028	−0.019**	1.765	1.424	0.341*
	3780[a]	3782[a]	0.008[b]	106[a]	106[a]	0.194[b]
			$\Gamma = 4.41$			$\Gamma = 0$
post-match	0.009	0.032	−0.023***	1.765	1.324	0.441*
regression	3780[a]	1872[a]	0.009[c]	106[a]	86[a]	0.245[c]
potential controls	16 220			1110		

[a]Number of observations.
[b]Abadie and Imbens heteroskedasticity robust standard errors.
[c]Clustered standard errors.