Detecting Land-Use Change and On-Farm Investments at the Plot Scale

Luke Sanford, Ran Goldblatt, Jennifer Burney*

University of California, San Diego

Paper Number: *GC33D-1107*

Background

Intensification of agricultural production, particularly for the world's smallholders, is critical for food security and economic development, but also for avoiding land degradation, land use change, and subsequent land use emissions. Public, Private, and NGO development entities around the globe are pursuing technological, institutional, and educational interventions to help farmers boost production on our increasingly stressed land base. Impact evaluation of such projects is thus very important, but rigorous impact evaluations of these kinds of interventions are expensive and difficult:

- Many interventions are not implemented with an intent to evaluate, and thus lack sufficiently defined control units.
- Implementers may not be able to afford rigorous evaluation, as it is time-, money-, and labor- intensive.
- Most rigorous evaluations only provide a snapshot of changes before and after an intervention, but no further information over the multi-year time frames relevant for measuring longer-run changes at both the intensive and extensive margins.
- Even within a proper evaluation, outcomes like changes in management and agricultural productivity are hard to measure well from the ground.

Can we use satellite data to better measure outcomes and evaluate projects?

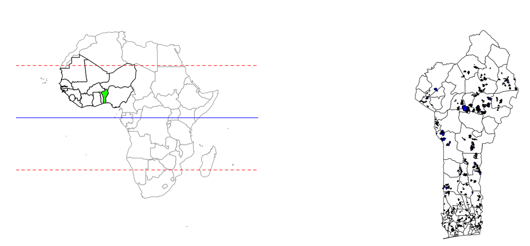
Here we use a land titling project conducted by the World Bank in Benin, West Africa, as a test case for using remote-sensing to measure land use change at the plot scale, and to conduct impact evaluations of projects that may not have been designed for rigorous evaluation of land use changes.

Land Titling in Benin

World Bank Plans Fonciers Ruraux Project:
Does improved land security lead to higher productivity?

Hypothesis: Titling → Security → Investment → Higher Value of Production

- Land demarcated and titled in 300 villages
- 70,000 "treated" plots
- Intervention from 2009-2011
- Follow-up survey in 2011



Control villages were selected and surveyed, but not geo-referenced!
Of 24 Hypotheses tested, two self-reported outcomes were significant at $\alpha = .05$:

1. Tree planting (1.7 % more likely)
2. Perennial Crops (2.6% more likely)

But none are significant after multiple hypothesis correction.

A better strategy?

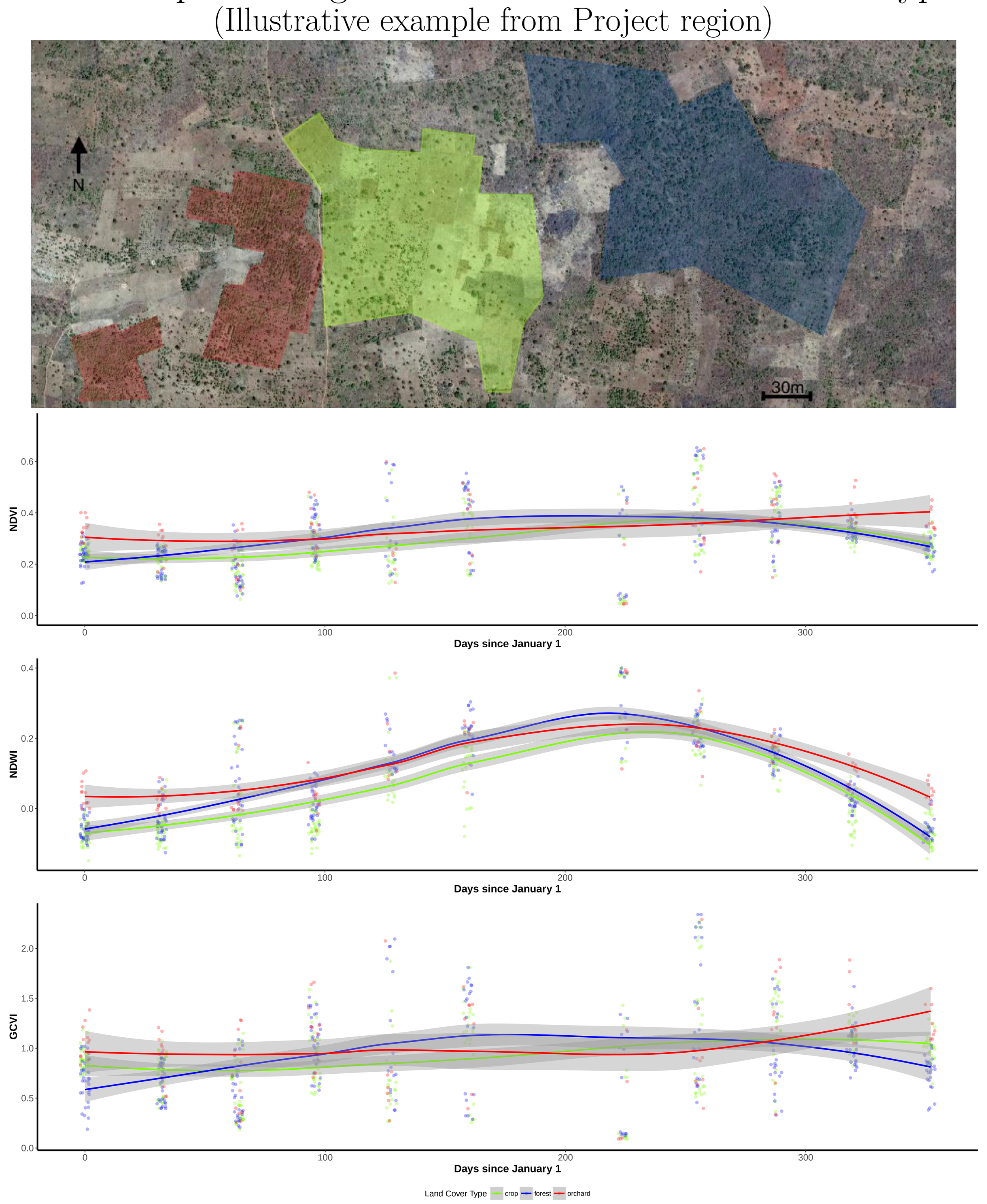
Test the impact of the program more effectively using Remote Sensing to:

1. **Better measure outcomes**: Use Landsat top-of-atmosphere reflectances to (a) define spectral signatures for land use types, and (b) extend time series for whole region to many years before and after the intervention to look for longer-run changes.
2. **Generate a quasi-experimental evaluation framework**: Use a Geospatial Synthetic Control method to generate a set of "control" units that look like PFR plots. Compare "treated" and "control" units in the pre- and post-treatment periods (difference-in-differences).

Use Google Earth Engine for data aggregation and time series extraction, R for analysis.

Remote Sensing of Crop Type

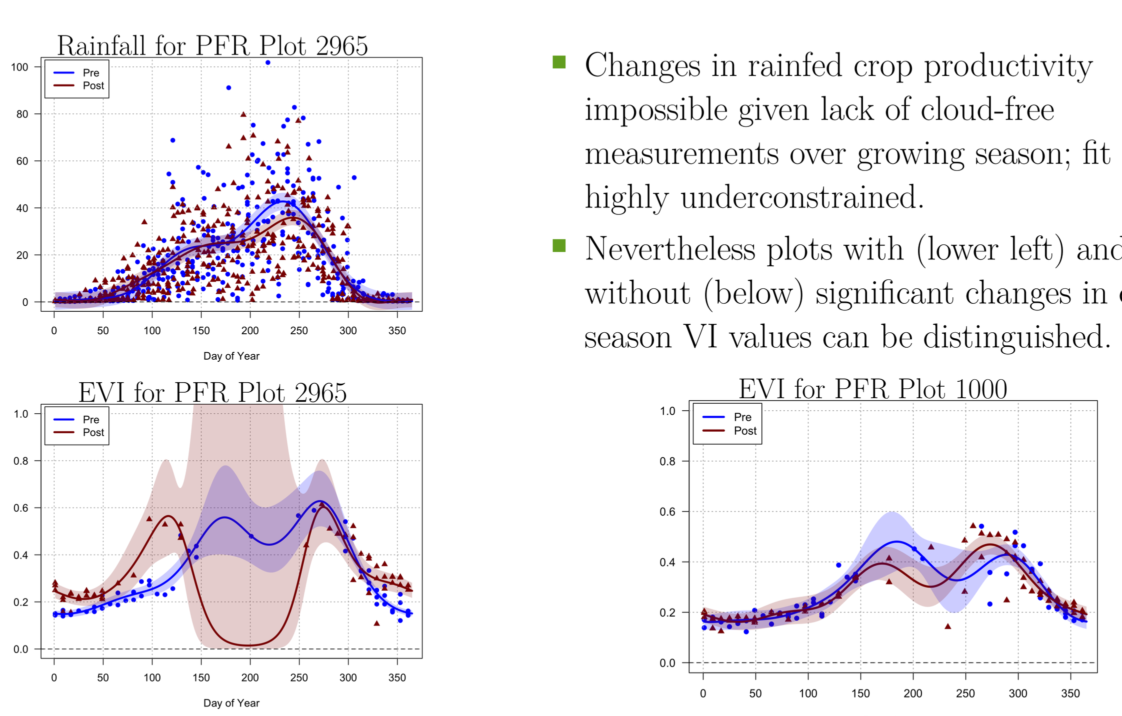
Seasonal Spectral Signatures of Different Land Cover Types:
(Illustrative example from Project region)



We can successfully distinguish planted orchards from other cropland in vegetation indices (VIs), so should be able to detect this type of investment through dry season signatures:

- Orchards stay greener through the dry season
- Forests respond more quickly to new rainfall

Remote Sensing of Land Use Changes

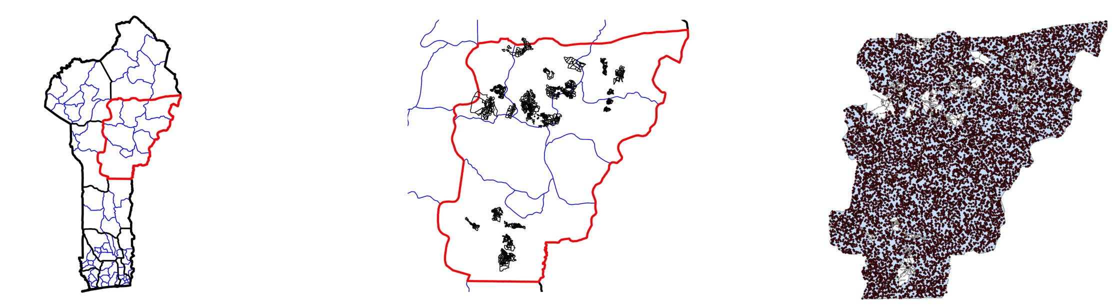


- Changes in rainfed crop productivity impossible given lack of cloud-free measurements over growing season; fit highly underconstrained.
- Nevertheless plots with (lower left) and without (below) significant changes in dry season VI values can be distinguished.

Synthetic Controls

The Main Idea: Want to know the *counterfactual*, or what *would* have happened to a plot had it *not* been titled. Use principal component analysis to generate a well-matched hypothetical "control" plot for every ("treated") PFR plot. Then compare these two groups before and after program implementation to estimate impacts.

Focus: Borgou Department



Borgou | PFR Plots | Non-Program Sample

Methods Overview:

- Sample points randomly from all non-program land areas in the region.
- For each treated (PFR) plot, generate a weighted average of non-program points that most closely resembles the PFR plot, pre-implementation.
- This weighted average of non-program points is the "synthetic control"

Methods Details:

We want to prevent over-fitting and escape the curse of dimensionality but still generate "parallel trends" between treated and synthetic control units:

1. Conduct factor analysis of control units Nc in pre-treatment period $T1$ and post-treatment period $T2$ (3 factors)

$$\underbrace{\begin{bmatrix} y_{1,1} & \cdots & y_{1,T} \\ \vdots & \ddots & \vdots \\ y_{Nc,1} & \cdots & y_{Nc,T} \end{bmatrix}}_{Nc\times T1+T2} = \underbrace{\begin{bmatrix} l_{1,1} & l_{1,2} & l_{1,3} \\ \vdots & \vdots & \vdots \\ l_{Nc,1} & l_{Nc,2} & l_{Nc,3} \end{bmatrix}}_{Nc\times K} \times \underbrace{\begin{bmatrix} f_{1,1} & \cdots & f_{1,T} \\ f_{2,1} & \cdots & f_{2,T} \\ f_{3,1} & \cdots & f_{3,T} \end{bmatrix}}_{K\times T} + \underset{Nc\times T}{\varepsilon}$$

2. Generate factor loadings for treatment data that minimize ε in pre-treatment period:

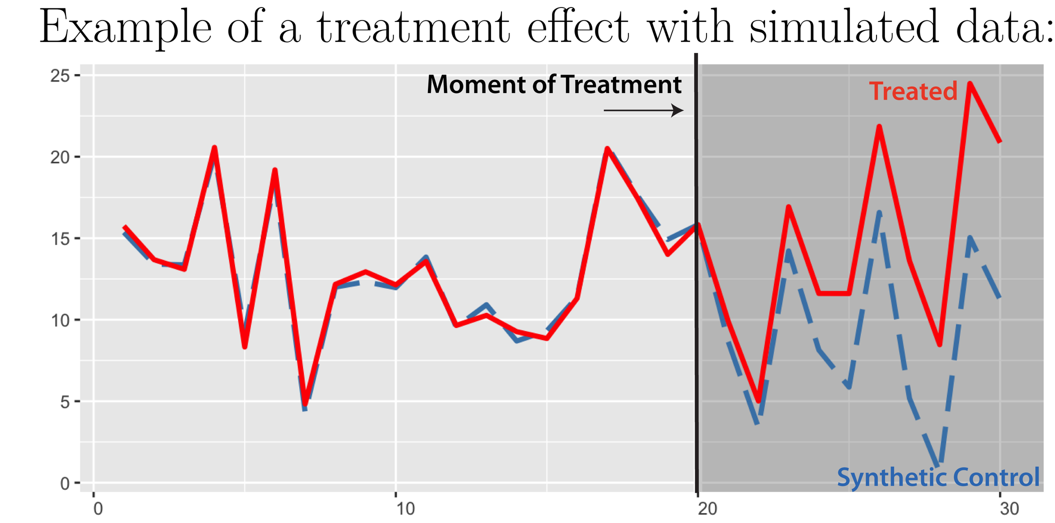
$$\begin{bmatrix} l_{1,1} & l_{1,2} & l_{1,3} \\ \vdots & \vdots & \vdots \\ l_{Nt,1} & l_{Nt,2} & l_{Nt,3} \end{bmatrix}$$

3. Infer untreated counterfactuals using these time-varying factors and unit-varying loadings:

$$\underbrace{\begin{bmatrix} y(0)_{1,1} & \cdots & y(0)_{1,T2} \\ \vdots & \ddots & \vdots \\ y(0)_{Nt,1} & \cdots & y(0)_{Nt,T2} \end{bmatrix}}_{Nt\times T2} = \underbrace{\begin{bmatrix} l_{1,1} & l_{1,2} & l_{1,3} \\ \vdots & \vdots & \vdots \\ l_{Nt,1} & l_{Nt,2} & l_{Nt,3} \end{bmatrix}}_{Nt\times K} \times \underbrace{\begin{bmatrix} f_{1,1} & \cdots & f_{1,T2} \\ f_{2,1} & \cdots & f_{2,T2} \\ f_{3,1} & \cdots & f_{3,T2} \end{bmatrix}}_{K\times T2} + \underset{Nt\times T2}{\varepsilon}$$

4. Uncertainty is generated using Monte-Carlo simulation.
5. Treatment effect: difference between $y_{i,t}$ (observed value for treated unit) and $y(0)_{i,t}$ (synthetic control, or inferred counterfactual, for treated unit).

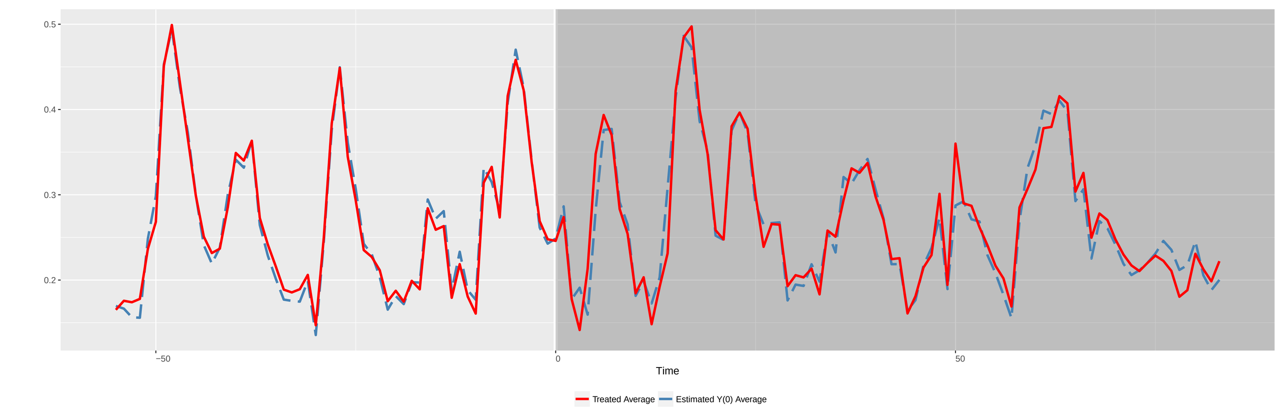
Example of a treatment effect with simulated data:



Good match in pre-treatment period; treated outcome deviates from synthetic control after intervention.

Results

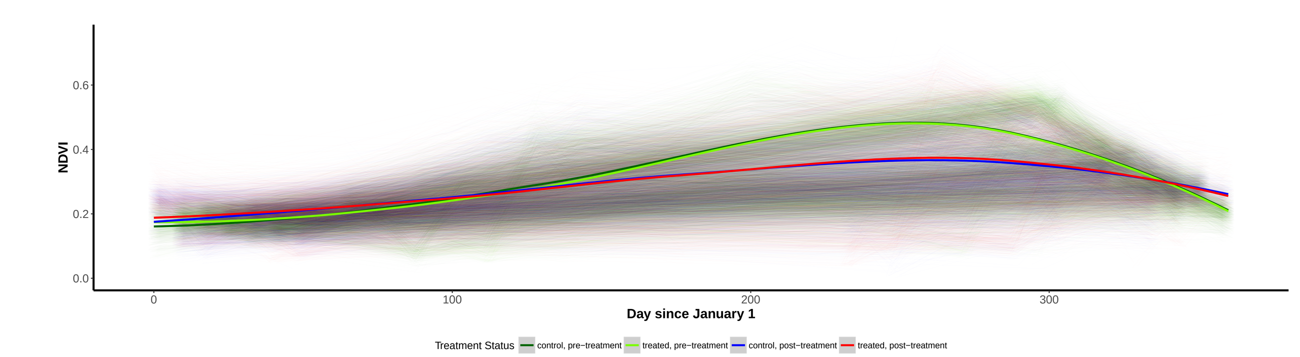
1. Successful creation of geospatial synthetic controls:



2. Very little evidence of impact:

	Dependent variable: NDVI		
	(End of Main Season)	(All Dry Season)	(Jan-April)
Treatment	0.001	0.003***	0.012***
	(0.001)	(0.001)	(0.001)
Constant	0.326***	0.223***	0.184***
	(0.001)	(0.0003)	(0.0003)
Adjusted R^2	0.0001	0.002	0.056
Resid. SE	0.036	0.023	0.021
F (df = 1; 6178)	1.470	13.942***	368.897***

*p<0.1; **p<0.05; ***p<0.01
Includes unit fixed effects



Conclusions

We find very little evidence for long-run impact of a land titling program on investments in orchards or perennial crops, as measured by dry-season vegetation indices. We are unable to detect investments in main growing season crops due to cloud cover. We nevertheless show a method of ex-post evaluation of agricultural interventions using remotely-sensed vegetation indices and geospatial synthetic controls that could be applied to a broad array of development projects.

References

- Jushan Bai 'Panel Data Models with Interactive Fixed Effects' *Econometrica* (2009).
- Yiqing Xu 'Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models' *Political Analysis* (2017).
- Markus Goldstein et al 'Formalizing Rural Land Rights in West Africa' *Development Research* (2015).

Group Website: www.jaburney.net
Emails: lcsanford@ucsd.edu, rgoldblatt@ucsd.edu, jburney@ucsd.edu