

A top-down approach to estimating spatially heterogeneous impacts of development aid on vegetative carbon sequestration

Daniel Runfola^{*1}, Ariel BenYishay^{†1}, Jeff Tanner^{‡2}, Graeme Buchanan^{§3}, Jyothy Nagol^{¶4}, Matthias Leu^{||5}, Seth Goodman^{**1}, Rachel Trichler^{††1} and Rob Marty^{‡‡1}

¹Institute for the Theory and Practice of International Relations,
The College of William and Mary

²Independent Evaluation Group, World Bank

³Center for Conservation Science, Royal Society of Birds

⁴Global Land Cover Facility, University of Maryland

⁵Department of Biology, The College of William and Mary

Short Title: Estimating impacts of aid on carbon sequestration

Keywords: International Aid, Carbon Sequestration, Causal Identification, Heterogeneous Effects

Words in Abstract: 150

Words in Manuscript: 3000

Number of References: 30

Number of Figures and Tables: 10

Corresponding Author:

Dr. Daniel Runfola

Institute for the Theory and Practice of International Relations, *AidData*

Email: dsmillerrunfol@wm.edu Telephone: 508.316.9109 Fax: 757.221.4650

^{*}dsmillerrunfol@wm.edu

[†]abenyishay@wm.edu

[‡]jtanner@worldbank.org

[§]graeme.buchanan@rspb.org.uk

[¶]jnagol@umd.edu

^{||}mleu@wm.edu

^{**}smgoodman@wm.edu

^{††}rbtrichler@wm.edu

^{‡‡}ramarty@email.wm.edu

1 Abstract

Since 1945, over \$4.9 trillion dollars of international aid has been allocated (Tierney *et al.* 2011). To date there have been no estimates of the regional impact of this aid on the carbon cycle. We apply a geographically explicit matching method (Andam *et al.* 2008) to estimate the impact of World Bank projects implemented between 2000 and 2010 on sequestered carbon, using a novel and publically available dataset of 47,084 World Bank project locations. Considering only carbon sequestered due to fluctuations in vegetative biomass caused by World Bank projects, we estimate a global net positive increase in sequestration totaling 1,398,229 tonnes. However, we illustrate the apparent effectiveness of environmental safeguards (Laurance *et al.* 2015) implemented by the World Bank are variable across different geographic contexts. We argue that subnational data can be helpful in the identification of these heterogeneous impact effects, and highlight the considerable methodologic barriers which still exist.

Data Sources	
Data Name	Source
Gridded Population of the World	Center for International Earth Science Information Network ²
Nighttime Lights	Defense Meteorological Satellite Program ³
Precipitation and Temperature	University of Delaware (Willmott & Matsuura 2001) ⁴
Urban Travel Time	European Commission Joint Research Centre ⁵
Distance to Rivers	World Wildlife Fund ⁶
Vegetation	NASA LTDR ⁷
Carbon Storage	NASA JPL (Saatchi <i>et al.</i> 2011) ⁸
Ecofloristic Zone Carbon Fractions	Oak Ridge National Laboratory (Ruesch & Gibbs 2008)

Table 1: Data sources used in this analysis.

2 Introduction

International recognition of the dual challenges of international development and the mitigation of environmental change is resulting in a large redirection of resources towards developing countries. The United States alone has pledged nearly \$4 billion dollars of aid to mitigate climate vulnerability over the next 4 years, and the Paris convention urged donors to target \$100 billion annually by the year 2020 (Royal 2015). Coupled with increasing pressure from recipient nations, donors have and continue to introduce more stringent environmental safeguards, including requirements for compliance with national and international regulations, environmental management plans, reforestation goals, and others (Nielson & Tierney 2003, Gutner 2005).

Despite these shifts in the policies of donor agencies, there is a gap in empirical studies examining the impacts of these policies on environmental outcomes. This letter aims to illustrate an approach to overcoming this gap, as well as highlight many of the remaining methodological challenges. We specifically examine the impact of World Bank projects on vegetation and subsequent changes in carbon sequestration, leveraging a novel and publically available dataset of 47,084 World Bank project locations ¹ in conjunction with long-term satellite data quantifying vegetative biomass and a number of spatially-referenced control variables (see table 1),

¹<http://aiddata.org/level1/geocoded/worldbank>

We find that while the overall impact of World Bank projects on carbon sequestration appears to be positive (an increase of 1,398,229 sequestered tonnes), considerable temporal and spatial variation exists in these impacts. Further, we illustrate the advantages and limitations of a geographically explicit approach to estimating the causal effects of development aid projects, and outline a number of topics for further research. Finally, we introduce an enhanced, publically available dataset of the global, spatially-explicit distribution of World Bank activities encompassing projects initiated between 1995 and 2014.

3 Methods

Over the last five years, significant progress has been made on methods which integrate spatial data (i.e., satellite information on forest cover) to quantify the causal impact of interventions (i.e., projects aimed at the prevention of deforestation) (Buntaine *et al.* 2015; Nelson & Chomitz 2011). These methods largely rely on propensity score and other matching-based methods to select "control" cases where no or limited intervention occurred, and match these with similar "treatment" cases (Andam *et al.* 2008). We build on these approaches, implementing a geographically explicit two-stage Propensity Score Matching estimation strategy.

3.1 Geographically Explicit Impacts

First, we truncate the dataset to cover World Bank projects from 2000 to 2010 due to limitations of our ancillary information, resulting in a total of 47,084 World Bank project locations. Second, the area of influence within which we anticipate each World Bank project could plausibly have an impact on deforestation is calculated by examining the historic spatial distance at which forest cover is spatially correlated before any projects were implemented. To parameterize this distance, we calculate a Moran's I (**getis_the_1992**) score at increasing distances, identifying the distance at which spatial autocorrelation is no longer predominant for our outcome measure of forest cover, measured in 1999. We argue that this estimate is a highly conservative estimate of the possible area of influence a project could have (i.e., we will tend to over-estimate the buffer size), as it represents the totality of historic spillovers up to the year 1999.

²<http://sedac.ciesin.columbia.edu/data/collection/gpw-v3/sets/browse>

³Stable Lights retrieved from <http://ngdc.noaa.gov/eog/dmsp.html>

⁴Variables derived from these product included the average precipitation (P) and temperature (T) before a project was implemented (from 1992), the linear trend in P and T from 1992 to the project implementation, the average temperature from the date the project was implemented until the end of the temporal record (2012), and the post-project trend through 2012. Absolute measurements of each variable were also retained.

⁵<http://forobs.jrc.ec.europa.eu/products/gam/download.php>

⁶<http://hydrosheds.cr.usgs.gov/index.php>

⁷<http://ltdr.nascom.nasa.gov/cgi-bin/ltdr/ltdrPage.cgi>

⁸<http://click.jpl.nasa.gov/Archive/carbon/ftpdata/carbon/datasets/>

For each of 12 distance bins (between 0 and 2,200 kilometers, in increments of approximately 180km), Moran's I is calculated following:

$$I_h = \left(\frac{N}{\sum_i \sum_j w_{ij}} \right) * \left(\left(\sum_i \sum_j w_{ij} * (X_i - \bar{x}) * \frac{X_j - \bar{x}}{\sum_i (X_i - \bar{x})} \right)^2 \right) \quad (1)$$

where h represents each spatial bin, N the number of spatial units, i and j are indexes for each unit, X is the variable of interest, and W_{ij} represents the weights matrix. In this application, the weights matrix is specified according to the bin (h) being analyzed.

Once calculated, the distance at which Moran's I is equal to or less than .10 is identified, and used to parameterize a buffer around each project. The projects are then subdivided into six different project (sector) types - Health, Environment, Education, Industrial, Infrastructure, and Other. Within each of these sectors, projects are further subdivided into three equally-sized monetary bins of "low", "medium", and "high" dollars committed to the project.

Finally, for each sector-monetary bin combination (18 total groups), a geographically explicit propensity score model is fit. This is conducted following a three-stage process which is repeated for every unit of analysis, resulting in 47,084 total model estimations. In the first stage, a unit of observation is selected and all World Bank projects that fall within the distance estimated according to the Moran's I are selected as the relevant "sub population" for that point. All points within the sub-population are defined as treated or untreated pending the monetary value being analyzed - for example, if the unit of observation selected is a high dollar value project, all projects in the medium and low bins are assigned an untreated value (0), while all high dollar value projects are assigned as treated (1).

In the second stage, all points within the sub-population are matched according to a propensity score matching routine. Variables matched on can be seen in Table 1, and matches were limited to be within the same sector. The propensity scores are calculated following:

$$T = \beta_0 + \sum_{k=1}^k (\beta_k * x_k) \quad (2)$$

where T is the treatment binary, and β_k are the estimated coefficients for each covariate, x_k . The estimates from this equation are applied to each unit of observation within the sub-population, and the differences between propensity scores across different units of observation represent a univariate measure of similarity. For the set of treated units within the area of influence, the optimal set of matched untreated units was identified using a nearest-neighbor optimization Ho *et al.* 2011.

In the third stage, a regression relationship is calibrated between the outcome measure (The average LTDR NDVI value in the years after project implemen-

tation), the treatment binary, and all available covariates:

$$y_i = \beta_0 + \beta_1 * T + \sum_{k=1}^k (\beta_k * x_k) + D_p \quad (3)$$

where y_i represents the level of forest cover within each zone i , β_1 represents the estimated impact of the treatment, and D_p represent fixed effects for each paired observation. Every unit of observation n has a zone i , defined as all projects which fall within the distance calculated using Moran's I.

This three-step process is repeated for every unit of observaiton. In some cases, insufficient matches or eligible cases existed to approximate the impact of a region; these units were omitted from the analysis.

3.2 Estimating Carbon Sequestration

Because the outcome measure examined (NDVI) is only a proxy for tree cover, an additional step of modeling must be conducted to translate changes in NDVI into changes in estimated tonnes of carbon sequestered.

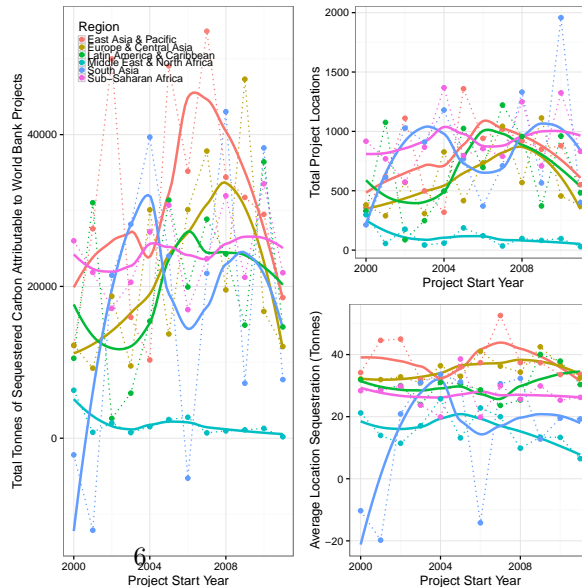
4 Results

The Results section should present the experiments that support the conclusions to be drawn later in the Discussion. The Results section should conform to a high standard of rigour. Extended lines of inference, arguments or speculations should not be placed in the Results.

5 Discussion

This approach has the benefit of comparing "apples to apples" - i.e., World Bank project locations are contrasted to others at which it is known a World Bank project exists. However, it also has the disadvantage of limiting the eligible set of matches to locations at which World Bank projects are.

The Discussion section should be separate from the Results section.



It allows authors to propose their interpretation of the results, and to suggest what they mean in a wider context. It should end with a clear statement of the main conclusions of the research, and a clear explanation of their importance and relevance to applied conservation and/or policy.

6 Acknowledgements

The authors would like to acknowledge the government of Sweden and the World Bank Independent Evaluation Group for partially funding this research. This work was performed in part using computational facilities at the College of William and Mary which were provided with the assistance of the National Science Foundation, Virginia Port Authority, Virginia's Commonwealth Technology Research Fund, and the Office of Naval Research. The authors would also like to thank Scott Stewart, Alex Kappel, Miranda Lv, Doug Nicholson, and Vinay Vijayan for their valuable contributions and insights.

7 References

References

1. Andam, K. S., Ferraro, P. J., Pfaff, A., Sanchez-Azofeifa, G. A. & Robalino, J. A. Measuring the effectiveness of protected area networks in reducing deforestation. en. *Proceedings of the National Academy of Sciences* **105**, 16089–16094. ISSN: 0027-8424, 1091-6490 (Oct. 2008).
2. Buntaine, M. T., Hamilton, S. E. & Millones, M. Titling community land to prevent deforestation: An evaluation of a best-case program in Morona-Santiago, Ecuador. *Global Environmental Change* **33**, 32–43. ISSN: 0959-3780 (July 2015).
3. Gutner, T. Explaining the Gaps between Mandate and Performance: Agency Theory and World Bank Environmental Reform. *Global Environmental Politics* **5**, 10–37. ISSN: 1526-3800 (May 2005).
4. Ho, D. E., Imai, K., King, G. & Stuart, E. A. MatchIt: Nonparametric Pre-processing for Parametric Causal Inference. *Journal of Statistical Software* **42**, 1–28 (2011).
5. Laurance, W. F. *et al.* Reducing the global environmental impacts of rapid infrastructure expansion. *Current Biology* **25**, R259–R262. ISSN: 0960-9822 (Mar. 2015).
6. Nelson, A. & Chomitz, K. M. Effectiveness of Strict vs. Multiple Use Protected Areas in Reducing Tropical Forest Fires: A Global Analysis Using Matching Methods. *PLOS ONE* **6**, e22722. ISSN: 1932-6203 (Aug. 2011).
7. Nielson, D. L. & Tierney, M. J. Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization* **57**, 241–276. ISSN: 0020-8183 (2003).
8. Royal, S. *United Nations Framework Convention on Climate Change ADOPTION OF THE PARIS AGREEMENT*. Dec. 2015. <<https://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf>>.
9. Ruesch, A. & Gibbs, H. *New IPCC Tier-1 Global Biomass Carbon Map For the Year 2000*. 2008. <<http://cdiac.ornl.gov/>>.
10. Saatchi, S. S. *et al.* Benchmark map of forest carbon stocks in tropical regions across three continents. en. *Proceedings of the National Academy of Sciences* **108**, 9899–9904. ISSN: 0027-8424, 1091-6490 (June 2011).
11. Tierney, M. J. *et al.* More Dollars than Sense: Refining Our Knowledge of Development Finance Using AidData. *World Development* **39**, 1891–1906. ISSN: 0305-750X (2011).
12. Willmott, C. & Matsuura, K. *Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950-1999)* 2001. <http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html>.