Title
Monitoring Forest Policy from Space

Author
Hammer, Daniel

Publication Date
2017

Peer reviewed|Thesis/dissertation

Monitoring Forest Policy from Space

by

Daniel Sussman Hammer

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Maximilian Auffhammer, Chair
Associate Professor Meredith Fowlie
Associate Professor Matthew Potts

Fall 2017

Monitoring Forest Policy from Space

Abstract

Monitoring Forest Policy from Space

by

Daniel Sussman Hammer

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Maximilian Auffhammer, Chair

Satellite imagery is an increasingly valuable data source in agricultural and resource economics. It offers credible, consistent, and globally comparable information on landscape-scale changes to the environment; and the supply of the information is accelerating with innovations in satellite technology. This dissertation presents the critical value of satellite imagery for environmental policy evaluation, focusing specifically on deforestation. Tropical deforestation may account for as much as one tenth of net global greenhouse gas emissions each year.[111, 62] Any viable effort to mitigate the impact of climate change must address deforestation. The effective design of forest policy is dependent on reliable impact analysis. Satellites offer a method of direct observation of forest cover loss, rather than incomplete or biased estimates. The dissertation first presents an algorithm to convert raw satellite imagery into a new and novel data source on tropical deforestation. Next, the data are used to explain the counter-intuitive outcomes from a 2012 conservation policy in Indonesia, relying on the spatial detail afforded by the new data. Finally, a set of empirical results describe the relationship of deforestation to interest rates, commodity prices, and other economic variables, using the temporal detail afforded by the new data. Together, these case studies demonstrate the effective use of a new data source in direct and credible policy analysis for natural resource economics.

To Emily

Contents

Contents ii

List of Figures iv

List of Tables v

1 Introduction 1

2 Alerts of forest disturbance from MODIS imagery 5
- 2.1 Introduction . . . 5
- 2.2 Data . . . 7
- 2.3 Methods . . . 8
- 2.4 Results . . . 14
- 2.5 Discussion . . . 17
- 2.6 Conclusion . . . 18

3 The spatial mechanism for Indonesia's failed forest conservation policy 20
- 3.1 Introduction . . . 20
- 3.2 Data . . . 21
- 3.3 Estimation strategy . . . 22
- 3.4 Results . . . 23
- 3.5 Conclusion . . . 24
- 3.6 Supplementary material . . . 24

4 Economic dynamics and forest clearing: A spatial econometric analysis 32
- 4.1 Introduction . . . 32
- 4.2 Prior research . . . 34
- 4.3 Expanding the scope of work . . . 35
- 4.4 Trends in Indonesian forest clearing (2005-2010) . . . 38
- 4.5 Model specification . . . 42
- 4.6 Data . . . 43
- 4.7 Econometric specification . . . 44
- 4.8 Results . . . 47

4.9 Conclusion 52
4.10 Supplement: estimation of spatial models 54

Bibliography **57**

List of Figures

1.1 Number of non-military, Earth imaging satellites in orbit through September 2017. 2

1.2 Three years of Landsat images of fields in the Central Valley, California, with each row depicting all scenes a given year. 3

2.1 Illustration of an abrupt shift in the 16-day NDVI time series (in blue). Sample NDVI series taken from the `R` package `BFAST`. 10

2.2 Overlaying NDVI trend t-statistic and PRODES deforestation data (black polygons) without and with spatial smoothing at 1km resolution. A large negative t-statistic appears in brown. This 100km×100km area is in Pará, Brazil (Figure 2.4a). 11

2.3 Algorithm processing chain 12

2.4 Spatial fit analysis covers all of Pará. Spatiotemporal analysis covers the red box in Figure 2.4a. 15

2.5 Comparison of 500 meter alerts to PRODES. PRODES deforestation from 2000 to 2006 appears as black outlines, while deforestation prior to 2000 is shown as gray polygons. False positive alerts appear as red pixels, while true positives appear in green. 17

3.1 Illustration of clusters 22

3.2 The impact of the moratorium on new forest concessions and commodity prices on the spatial distribution of forest clearing activity. 25

3.3 Illustration of an isoquant where the inputs exhibit a high degree of substitution in production and a low marginal rate of technical substitution. 29

3.4 The results of DTW matching. 30

4.1 Large-scale forest clearing in Indonesia (2005-2010) 39

4.2 Change in forest clearing index value: 2006 vs. 2010 39

4.3 Sumatra and Kalimantan: forest clearing rank of kabupatens. Among 1372 secondary administrative units in Southeast Asia. 40

4.4 Annualized forest clearing index values: top five Indonesian Provinces in January 2007. 41

4.5 Annualized forest clearing index values: top five kabupatens in Riau. 41

List of Tables

2.1 Accuracy assessment, PRODES versus FCLH and output, 2000-2006 and 2007-2010 in Pará, Brazil 16

3.1 Empirical results. 26

4.1 Definitions of estimation variables 46
4.2 Economic dynamics and forest clearing. All variables in logs except month and real interest rate. Lags (months) in brackets. Dependent variable: log(`Clear`). . 50
4.3 Introduction of cross-sectional variables. Variables logs except interest rate, area %s. Lags (months) in brackets. Dependent variable: log(`Clear`) 56

Acknowledgments

Thanks to Max Auffhammer, Jeff Hammer, David Wheeler, George Judge, and Jediphi Cabal. All mistakes are theirs alone.

Chapter 1

Introduction

Satellite imagery is increasingly common in economic research. It is used to measure crop yield, economic development, and deforestation, among many other applications. Data that is derived from satellite imagery is credible, consistent, and globally comparable. The cost of imagery is falling dramatically, alongside the accelerated pace of innovation in satellite technology. The spatial and temporal resolution of commercial satellite imagery is improving exponentially.[17] Researchers are able to monitor changes on the Earths surface at 30-centimeter resolution, with new imagery collected at least once each month. Medium resolution imagery at 1-10 meter resolution is now available for every point on Earth at least once each day.[1][6]

The pace of supply shows no signs of slowing. Figure 1.1 charts the growth in the number of active, non-military, Earth-observing satellites in orbit between January 1993 through September 2017.[2] Of the estimated 1,738 active satellites, 605 are used for applications in Earth Science and Earth Observation — as classified by the Union of Concerned Scientists.[32] The exponential growth since 2014 is attributed to the proliferation of small satellites (”smallsats” or ”cubesats”) from privately held companies like Planet Labs. A single launch may place over 80 satellites in orbit for a single company.

In parallel, the computational power and machine learning techniques to process satellite imagery into economically relevant information have become broadly accessible. For example, with training data that is freely available, it is now possible to extract, measure, and analyze building footprints from satellite imagery. Researchers are able to rely on direct observation of new home construction, rather than imperfect proxies from surveys or sampling. The fidelity and accessibility of direct economic measures has immediate benefits for

[1] Spatial resolution is measured by the size of a square section on the Earths surface. Thus, 30-centimeter resolution indicates that each pixel in an image represents a 30cm $\times$ 30cm square on Earth. High-resolution indicates a spatial resolution higher than 1m, while medium resolution is between 1-meter to 10-meter resolution.

[2] The United Nations maintains a Register of Objects Launched into Outer Space, which represents an estimated 92% of the total objects in orbit. The data for Figure 1.1 have been cleaned and collated by the Union of Concerned Scientists.

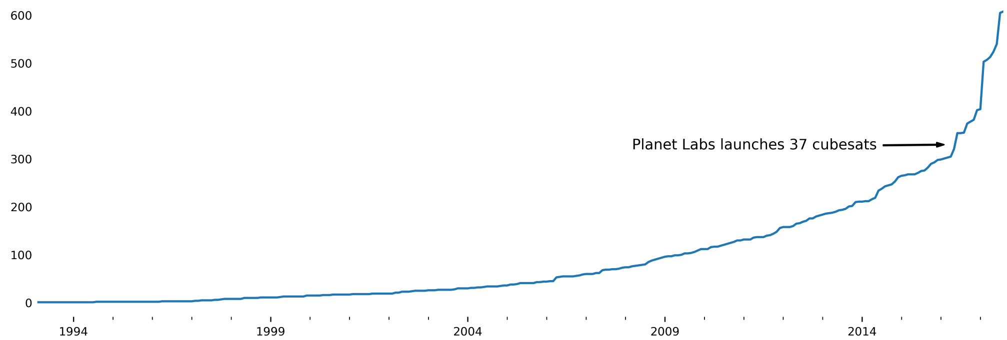


Figure 1.1: Number of non-military, Earth imaging satellites in orbit through September 2017.

economic research.

Direct observation of landscape-level change is especially promising for empirical questions in natural resource economics. Consider, for example, the foundational 1976 paper *The Optimal Development of Resource Pools* by Martin Weitzman.[117] Weitzman characterizes the optimal policy for extracting natural resources from distinct "pools" under conditions of variable costs — across space and through time. For decades, many of the model's implications remained untested and, frankly, untestable. Satellite imagery offers detailed, comprehensive, and time-consistent information on the extraction of resources that can be seen from space — gold mines, forests, agricultural areas, among others. Indeed, Chapter 3 relies heavily on Weitzman (1976) to frame the empirical study of deforestation. Figure 1.2 illustrates the spatial and temporal detail afforded by satellite imagery. The figure depicts NASA imagery of crop rotation for three years in California.[3] The imagery to directly observe and quantify crop rotation is freely accessible for the entire planet. The value of this information has not been fully realized by economic researchers; but satellite-derived information is becoming more commonplace in economic literature, slowly but surely.

Despite the broad value of satellite imagery, there are clear limitations of its use. Satellite imagery is not appropriate to measure gender roles or educational outcomes. It is appropriate only for measuring landscape-level changes that occur over the course of a week or a month — the frequency of image collection. Deforestation is particularly well-suited to be measured from satellite imagery. It is clearly identifiable from satellite imagery, collected over time. Furthermore, research into forest policy is critical for pressing environmental issues. Tropical deforestation may account for as much as one tenth of net global greenhouse gas emissions each year.[111, 62] Any viable effort to mitigate the impact of climate change must address tropical deforestation; and designing policies to reduce the rate of deforestation require

[3]Specifically, the imagery is from is a joint program of NASA and the United States Geological Survey, Landsat, providing 30-meter resolution imagery every eight days. Landsat data are free and open for public use.

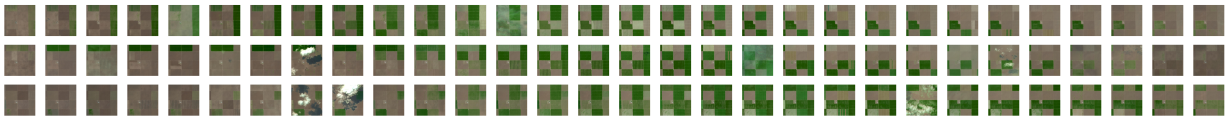


Figure 1.2: Three years of Landsat images of fields in the Central Valley, California, with each row depicting all scenes a given year.

careful study grounded in real data. The study of deforestation from satellite imagery is both opportunistic and critically important.

The three subsequent chapters address the economic drivers of deforestation using satellite imagery. First, an algorithm to measure deforestation from freely available satellite imagery is described and evaluated. The resultant data set, called FORMA for *Forest Monitoring for Action*, reports forest clearing activity at 500-meter resolution with 16-day updates for the entire humid tropical biome. The FORMA data is utilized the final two chapters, which use the unique features of the data set for policy and economic analysis. The unprecedented spatial and temporal resolution of FORMA afford the opportunity to link the spatiotemporal growth of deforestation to policy levers. It is notable that FORMA was the foundational data set for Global Forest Watch, an online platform to monitor deforestation from satellite imagery. Global Forest Watch has become the most popular environmental data platform ever released, indicating the operational value of the data in addition to the research value.

The next chapter (Chapter 3) describes the unintended consequences of forest policy in Indonesia. In an effort to reduce deforestation, the government of Indonesia enacted a moratorium on logging and agricultural concessions in May 2011.[9] Previous research has shown that this policy did not reduce the deforestation rate as intended; and, in fact, deforestation increased after the moratorium was enacted.[78] This empirical study demonstrates that the moratorium may have *caused* an increase in deforestation. The spatial distribution of deforestation was significantly impacted by the conservation policy, displacing clearing activity from one margin of production to another, latent margin of production. The net effect was an increase in the aggregate rate of clearing. The empirical result can be rationalized with basic production theory, examining the intensive and extensive margins of production. The detailed spatial information, derived from satellite imagery, allows for the separation of deforestation that occurs on the periphery of existing deforestation clusters (defined as intensive deforestation) or that seeds a new cluster (defined as extensive deforestation).

Finally, Chapter 4 utilizes the temporal resolution to examine the economic drivers of deforestation, including interest rates and commodity prices. An adequate study of the impact of interest rates on deforestation requires a measure of deforestation that is updated frequently – to match the time-scale of economic decision-making. Research into the dynamics of deforestation is possible with *at least* monthly updates. At best, deforestation data was updated annually, whereas decisions to invest in forest clearing activity were based on weekly or even daily information. The chapter further examines the dynamics of deforestation by employing methods in spatial econometrics to account for spatiotemporal autocorrelation.

While the focus of this dissertation is deforestation, the source data and methods are valuable to any study that relies on landscape-scale change. The new data, empirical methods, and computational capacity have, together, expanded the scope of questions that can be answered in empirical economics – and especially in agricultural and natural resource economics.

Chapter 2

Alerts of forest disturbance from MODIS imagery

2.1 Introduction

Deforestation and forest degradation contribute 12 percent of global anthropogenic greenhouse gas emissions each year.[60, 109] In addition, habitat loss and fragmentation of forest landscapes threaten ecosystem resilience and biodiversity.[45, 79] Environmental externalities do not necessarily enter the economic calculus for private decision makers.[90] Efforts to align private and social incentives depend crucially on measuring and evaluating the impacts of activities affecting forests. Thus, reliable, timely, and transparent information on forest disturbance is urgently needed, especially in the humid tropics, which accounts for the majority of global deforestation.[44]

Existing methodologies: Data

Current techniques to monitor forest disturbance must balance spatial and temporal resolution. Sampling techniques using high-resolution imagery have been shown to reliably detect small-scale disturbance. But infrequent updates, relatively high data and processing costs, and limited spatial coverage all constrain these techniques.[49, 8]

Moderate- and coarse-resolution imagery from Landsat and MODIS, respectively, is acquired at higher frequency and for broader geographic areas. Such data sets sacrifice spatial detail in favor of more timely information on forest cover disturbance.[38, 58, 99, 103, 102, 92, 3, 94]

A version of this chapter first appeared in the *International Journal of Applied Earth Observation and Geoinformation* (2014) Vol. **33**, with David Wheeler and Robin Kraft.

Despite the high frequency of coarse-resolution imagery products, monitoring is still often constrained by persistent cloud cover in the humid tropics. The DETER monitoring system [99], for example, only reports data in the legal Amazon in Brazil, where cloud cover is less problematic than in other tropical regions.[95, 52] Other systems [92] use resampled or aggregated data products to reduce the impact of cloud cover.

Existing methodologies: Algorithms

There are two broad approaches to change detection: cross-sectional differencing and time series analysis. Cross-sectional differencing is well-suited to relatively high resolution imagery that is acquired infrequently. However, persistent cloud cover often precludes the ability to identify transient change.[33] In Indonesia, for example, the conversion from primary forest to secondary forest or oil palm plantations can be completed between compositing periods, or even within the interval required to create cloud-free image composite.[15]

Time series analysis examines the spectral history of each pixel, with different techniques depending on the temporal resolution of the imagery. Other algorithms from search for specific temporal signatures of forest disturbance in Landsat pixel histories, using one observation out of a potential 22 observations each year.[74, 63, 15] Imaging twice a day by MODIS sensors aboard NASA's Terra and Aqua satellites yield far more potentially usable observations, reducing the impact of a cloudy scene on the overall time series. Choosing imagery with higher temporal resolution for sub-annual forest assessment requires discerning true change from a noisier time series with strong seasonal components, even in the tropics.[114]

The dense MODIS time series has inspired a varied set of techniques for identifying landuse/landcover (LULC) change more broadly. Mildrexler *et al.* (2009) detect disturbances in North America by generating and analyzing annual composites of maximum land surface temperature (LST) in conjunction with composited and 16-day enhanced vegetation index (EVI) values.[80] Campos and DiBella (2012) also identify several types of LULC around the world, but analyze full MODIS NDVI time series using wavelet transforms.[21] In the Dry Chaco ecoregion of South America, Clark *et al.* (2010) used TIMESAT to analyze multispectral MODIS time series and generate annual land use change statistics.[29] Finally, Kleynhans *et al.* (2010) show that time series analysis of NDVI using an extended Kalman filter and 3x3 pixel neighborhoods is more effective at change detection than image differencing.[75]

Proposed enhancements

This paper proposes a methodology for an alerting system for forest disturbance using imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) and statistical techniques borrowed from time series econometrics. The methodology has been implemented as an open source, automated forest monitoring system that covers the humid tropics. By leveraging the high-frequency MODIS imagery, the system produces forest disturbance alerts at 16-day, 500 meter resolution. It is intended to complement high-resolution forest change data sets that are updated less frequency but capture more variation in small-scale forest disturbance.

The algorithm relies on break detection algorithms to examine the spectral time series for each pixel, searching for structural changes in temporal patterns that are historically associated with forest disturbance. The structural changes are independent of both seasonal and random variation, helping to reduce the impact of clouds within the dense time series. In addition, the analysis incorporates rainfall in order to broadly control for transient climatic conditions during the sample period. A semi-supervised learning algorithm then characterizes each pixel's time series and compares it to historical forest cover loss data for 2000-2005, as identified in previous studies.[58] The classification rule derived from this process is used to classify each pixel for each 16-day interval after December 2005, the end of the training interval. Each pixel is assigned a normalized index representing the probability of disturbance.

The paper is organized as follows: Sections 2.1 and 2.1 present an overview of existing monitoring systems and algorithms to measure and quantify forest disturbance. Section 2.1 positions the proposed algorithm within previous research. Section 2.2 describes the raw imagery and other data used to develop the final data product. Section 2.3 provides a detailed, step-by-step explanation of the algorithm. Section 2.4 reports the correspondence of the proposed method with other remotely sensed forest disturbance data in Pará, Brazil. Section 2.5 describes the computational tradeoffs within the final feature selection, and Section 2.6 concludes.

2.2 Data

Time series data

The proposed algorithm uses the following time series data sets: (1) Normalized Difference Vegetation Index (NDVI) from the MODIS Terra 500 meter vegetation indices product (MOD13A1); (2) rainfall from the Precipitation Reconstruction over Land (PREC/L) data set; and (3) active fire detections identified by the Fire Information for Resource Management System (FIRMS).

The 500 meter NDVI layer in the MOD13A1 16-day composite [86] is a measure of pixel-level vegetation intensity used to flag extra-seasonal and persistent loss of vegetation. NDVI is particularly well-suited for use with forest disturbance alerts because it becomes saturated for dense vegetation such as tropical forest canopy.[48] Thus, deviations from a heavily forested norm are particularly striking.

The PREC/L precipitation data set [25] is used to control for broad climatological trends that may impact NDVI values across a wide geographic area [42, 100], independent of actual forest disturbance. For example, temporary fluctuations in NDVI during a period of abnormally low rainfall, such as during the El Niño Southern Oscillation, should not systematically lead to a forest disturbance alert. Rather, the algorithm calculates a time trend of NDVI conditional on precipitation. By including precipitation data, we account for the impact of precipitation on vegetation, which is well-documented in previous studies.[98, 50, 116, 14, 80, 42]

Finally, the Fire Information for Resource Management System (FIRMS) is based on 1km data from the MODIS MOD14/MYD14 Fire and Thermal Anomalies Product from Terra and Aqua.[36, 68] The use of fire for forest conversion varies with land management practices across the tropics and over time, but can be a reliable indicator of forest disturbance.[82, 110, 91]

Static data

Each MODIS pixel is also characterized using data sets that are static over time or rarely updated. These data sets are used to define the study area and group pixels to improve classification and ease computation. The Forest Cover Loss Hotspots (FCLH) data set [59] defines the study area and serves to train the classifier described in Section 2.3. This 500 meter data set is derived from MODIS imagery and only considers pixels with 25 percent or greater tree canopy cover, as defined in the Vegetation Continuous Fields (VCF) data set.[56] Moreover, it is the only publicly available pantropical forest cover loss data set at native MODIS resolution, ensuring a spatially consistent comparator. The study area in Hansen *et al.* (2008) is further restricted to the humid tropical biome, as defined in the World Wildlife Fund (WWF) ecoregions data set.[88, 59] The ecoregions data set is also used to group pixels that fall into the same ecoregion and may be considered broadly similar in terms of biology and climate. This allows for local tuning of the classifier (see Section 2.3), while also facilitating parallelization of data processing. Finally, the Global Administrative Areas (GADM) data set [61] is used to define coastlines, which in turn are used to control for the decreased reliability and quality of coastal pixels, as defined in MODIS quality flags.

2.3 Methods

The classification process analyzes the temporal characteristics of each pixel during the training period. A separate classification rule is generated for each terrestrial ecoregion in order to account for local vegetation, land use, or forest clearing patterns. The algorithm is then applied across the forested humid tropics.

Data cleaning

The algorithm is constructed to utilize as much separable variation in the NDVI as possible. An earlier version of the algorithm incorporated quality flags, but this eliminated most useful variation. Within the algorithm, heavy reliance on the quality flags precluded the ability to separate pixels subject to forest disturbance from those that were not. Rather, temporal smoothing using moving averages and ordinary least square (OLS) regression analysis of trends in the time series [93, 113] are used to limit the effect of contaminated observations. In addition, we incorporate (Section 2.3) a parameter instability test that is robust to seasonal cloud cover and random measurement error. [55] The test statistics are included as features that define the classification rule. The impact of clouds in a region that is prone to persistent cloud cover is therefore a muted output map, rather than a prevalence of falsely alerted pixels.

Other smoothing techniques and trend- and break- detection algorithms were examined [113, 77, 16, 30], and components of these algorithms were implemented and tested. But their computational requirements were deemed too costly to be applied tropics-wide. Implementation of such algorithms will undoubtedly be possible in the near future as computing costs continue to drop. These more sophisticated methods to improve data quality or mitigate the effect of low-quality input data may improve the quality of the output.

Time series feature extraction

Four features of the NDVI series are extracted for use in subsequent classification:

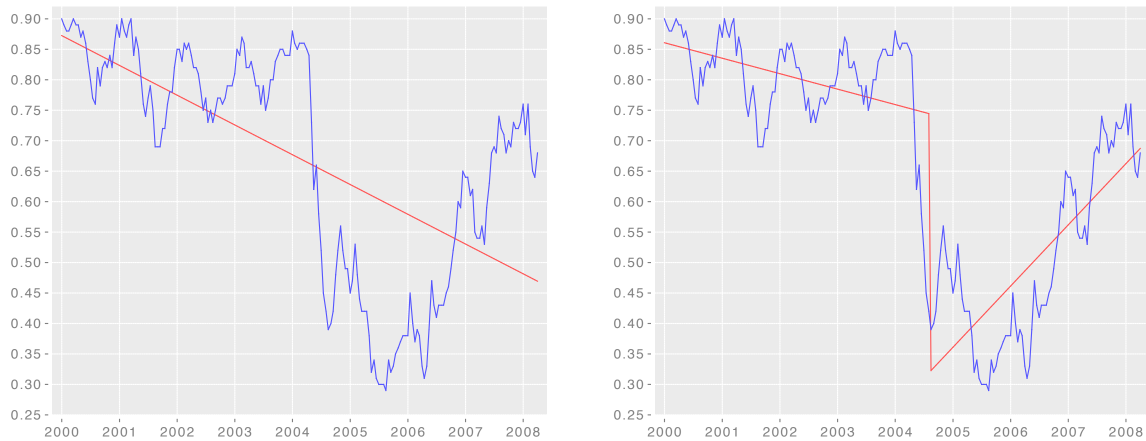
1. Long-term trend: The linear trend over the full NDVI time series, conditional on precipitation and monthly indicator variables to account for seasonality in the NDVI. A decreasing trend is correlated with a decrease in vegetation over time, which may indicate forest disturbance (Figure 2.1a).

2. Ratio of long-term linear trend to standard error: The linear trend divided by the standard error, or the Student t-statistic, of the linear NDVI trend indicates its statistical significance.

3. Minimum short-term trend: The NDVI series is split into moving blocks of 1.5 years. The linear trend is calculated for each block, and the most negative trend is retained.

4. Parameter instability test statistic: The test statistic described in Hansen *et al.* (2002) tracks error away from modeled time series stability.[55] The test statistic is extracted to indicate an abrupt shift away from paradigmatic behavior. Normal behavior can include seasonal variation induced by, for example, persistent cloud cover or abnormally low precipitation (Figure 2.1b).

For each pixel series, these characteristics are collected alongside the characteristics of thermal anomalies from the FIRMS data set. A running count of all thermal anomalies is extracted as a time-dependent feature. The anomalies are further categorized based on the measurement confidence (greater than or equal to 50 percent) and brightness (greater than or equal to 330 Kelvin), as specified in Morton *et al.* (2008).[82]



(a) Linear trend

(b) Parameter instability test

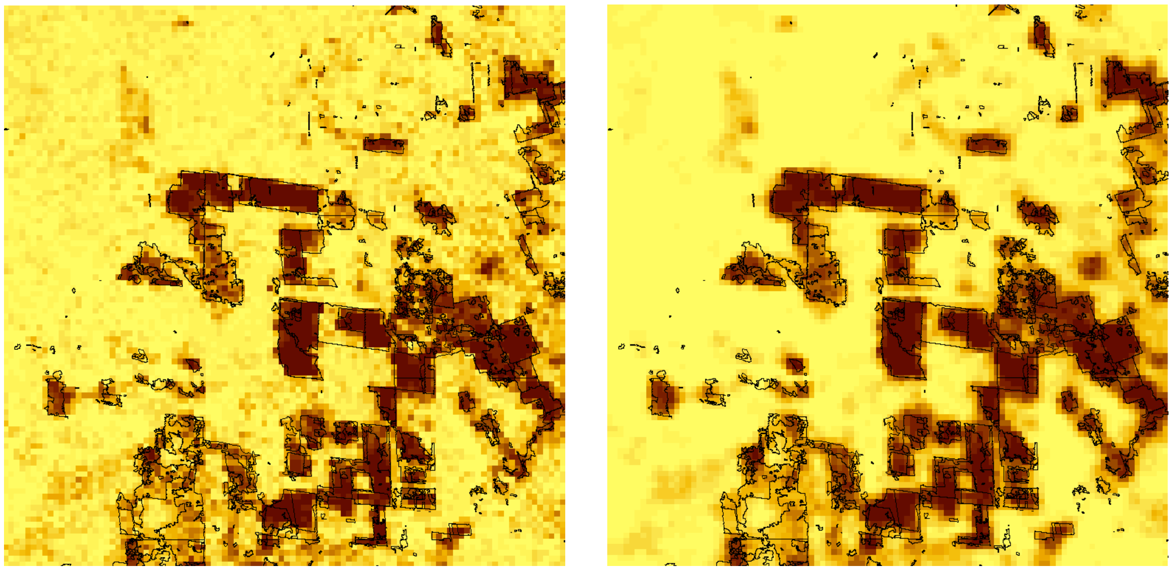
Figure 2.1: Illustration of an abrupt shift in the 16-day NDVI time series (in blue). Sample NDVI series taken from the R package BFAST.

Finally, information about the features described in Section 2.3 for each pixel's immediate neighbors is included in the algorithm to limit the effect of extreme values caused by random measurement error. Figure 2.2a maps the t-statistic of the NDVI long-term trend for an area of southeast Pará, corresponding to the signal-to-noise ratio of the trend, along with boundaries of PRODES deforestation.[20] It is notable that the t-statistic alone captures a significant proportion of variation in the PRODES data set. However, the spatially smoothed t-statistic illustrated in Figure 2.2b is a better predictor of PRODES deforestation, which demonstrates why we use neighborhood values as an input for the overall predictions (Section 2.3).

Classification rule

The interpretation of the array of features described in the previous section is based on a comparison of the features to historical forest cover loss data.[59] More precisely, a classification rule is developed for each ecoregion in the sample area, according to a modified logistic classifier.

Let A_j be the set of all pixels within a terrestrial ecoregion indexed by j. The classification rule for A_j is determined by the feature vectors $\mathbf{v}_i$ for all pixels $i \in A_j$. Each $\mathbf{v}_i$ has length 20, corresponding to 20 features derived from the NDVI and fires time series analysis (Section2.3). Paying due deference to uncertainty in empirical estimates, we calculate the probability of forest disturbance from the feature vector, where each feature is extracted during the training period (February 2000 through December 2005). Let $z_i = 1$ be the event that pixel $i \in A_j$ was subject to forest disturbance during the training period, and $z_i = 0$



(a) **Without** neighborhood features (b) **With** neighborhood features

Figure 2.2: Overlaying NDVI trend t-statistic and PRODES deforestation data (black polygons) without and with spatial smoothing at 1km resolution. A large negative t-statistic appears in brown. This 100km×100km area is in Pará, Brazil (Figure 2.4a).

otherwise. We model the probability based on the logistic function:

$$\mathbb{P}(z_i = 1|\mathbf{v}_i; \beta_j) = \frac{e^{\beta_j \mathbf{v}_i}}{e^{\beta_j \mathbf{v}_i} + 1}, \tag{2.1}$$

where β_j is the weighting or parameter vector for ecoregion j. We use the 500 meter FCLH data as the comparison data set, such that $z_i = 1$ if and only if the pixel i was classified as a forest cover loss hotspot between 2000 and 2005.

We estimate the parameter vector β_j using maximum likelihood, employing a variant of the Newton-Raphson method to optimize the likelihood function.[28] We selected the Newton-Raphson method because it could be implemented in a distributed computing environment. The parameter vector estimate $\hat{\beta}_j$ is calculated through an iterative process defined by

$$\beta_j^{(k+1)} = \beta_j^{(k)} + (\mathbf{V}'\mathbf{W}\mathbf{V})^{-1}\mathbf{V}'(\mathbf{z} - \mathbf{p}), \tag{2.2}$$

where $\beta^{(k)}$ indicates the value of the parameter vector at the k^{th} iteration; the matrix V is composed of the stacked feature vectors v_i for all $i \in A_j$; the vector $\mathbf{z}$ is the binary vector where z_i is defined as above; the vector $\mathbf{p}$ is the probability of forest disturbance based on the value of the parameter vector at the k^{th} iteration, or $p_i = \mathbb{P}(z_i = 1|\mathbf{v}_i; \beta_j^{(k)})$; and the matrix W is a diagonal matrix with $W_{ii} = p_i(1 - p_i)$ and zeros everywhere else.

The features in $\mathbf{v}_i$ are highly collinear. Persistent fires in forests, for example, are directly related to abrupt declines in vegetation. As a result, the matrix $\mathbf{V}'\mathbf{W}\mathbf{V}$ may be singular

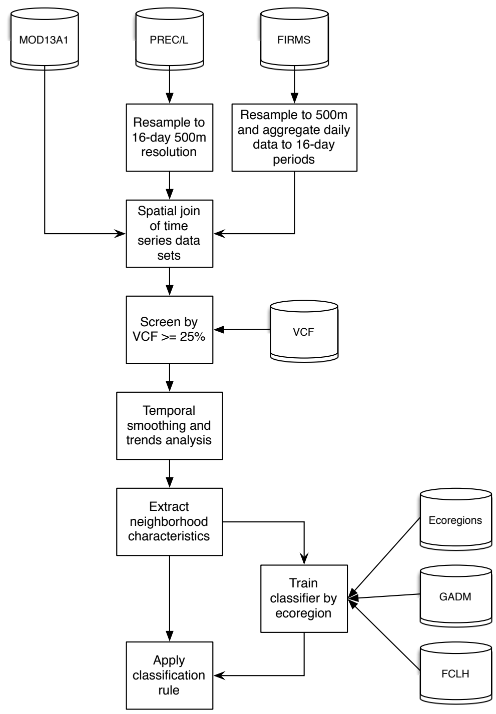


Figure 2.3: Algorithm processing chain

at machine precision at some iteration, which will break the estimation process because the matrix cannot be inverted. In order to force inversion, a very small ridge $c\mathbf{I}_n$ with $c \leq 10^{-8}$ is added element-wise to $\mathbf{W}$ when the determinant becomes dangerously close to zero, much like a ridge regression.[24] The effect on $\beta_j^{(k+1)}$ is then estimated in order to correct for the bias induced by this numerical adjustment.

The exact procedure is detailed in the source code, but amounts to the addition of the ridge and then a correction from the fallout. The final estimate, after convergence, $\hat{\beta}_j$ is used to classify each pixel in the corresponding ecoregion for a given interval.

Algorithm processing chain

The algorithm and accompanying data pipeline (Figure 2.3) run within a map-reduce framework to dramatically reduce computation time. The algorithm has been implemented in Clojure and Java to run on a Hadoop distributed computing cluster. All source code is fully documented and available in an online source code repository.[65]

Step 1: Build time series for all forested pixels

A time series of 16-day NDVI composites, monthly precipitation, and daily fires is created for each 500 meter MODIS pixel. Monthly precipitation data is resampled to match the 16-

day temporal resolution of the NDVI data. Daily fires are aggregated into counts for each 16-day period. The three datasets are then joined spatially into the native 500 meter MODIS grid. Non-forest pixels are removed using the VCF data set with a threshold of 25 percent tree canopy cover, following the methodology of Hansen *et al.* (2008) in the generation of the 500 meter FCLH data set.[59]

Step 2: Extract time series trends

The basis observation is each telescoping interval from February 18, 2000 to each 16-day period after 2005 for each 500 meter pixel. At this stage, the time series characteristics described in Section 2.3 are extracted and associated with each basis observation. More precisely, the feature vector that represents each basis observation is extended to include the four NDVI time series characteristics, as well as four fire characteristics based on (1) total fire count; (2) fire count with brightness greater than or equal to 330 Kelvin; (3) fire count with confidence greater than or equal to 50 percent; (4) and fires meeting the brightness and confidence thresholds.

Step 3: Collect neighborhood characteristics

The feature vector for each basis observation is further extended to include characteristics from adjacent pixels within the specified time period. This step extends the feature vector for each basis observation the following characteristics, based on at most 8 adjacent forested pixels: (1) number of detected fires; (2) average short-term drop statistic based on the NDVI; (3) most negative short-term drop statistic based on the NDVI; (4) average long-term trend in the NDVI; (5) most negative long-term trend in the NDVI; (6) average t-statistic in the long-term trend of the NDVI; (7) most negative t-statistic in the long-term trend of the NDVI; (8) average break statistic calculated from the NDVI; (9) maximum break statistic calculated from the NDVI; (10) number of fires with brightness greater than 330 Kelvin; (11) number of fires with confidence greater than 50 percent; and (12) number of fires matching both brightness and confidence thresholds.

Step 4: Train classifier

The time series characteristics for each pixel described in steps 2 and 3 are interpreted for the period 2000-5 in terms of forest disturbance given in the 500 meter FCLH data set. A logistic regression is used to model the probability of forest disturbance by weighting each time series characteristic based on all forested pixels in a given ecoregion (Section 2.3).

Step 5: Apply classification rule

For each interval from 2000 through each 16-day period after the training period, the classification rule from step 4 is applied to each pixel. This process interprets each pixel's time series characteristics over the given interval in terms of the spatiotemporal signature of historical forest disturbance in the pixel's ecoregion. The result is a probability series that represents the strength of the forest disturbance signal.

Step 6: Post-processing

The raw probability series for each pixel is smoothed using a 3-period moving average in order to limit false positives. This inevitably delays detection and increases false negatives, but allowing for evidence to accumulate improves the reliability of the system as an alarm. The smoothed probabilities are interpreted as the level of confidence that forest disturbance occurred in a given pixel.

2.4 Results

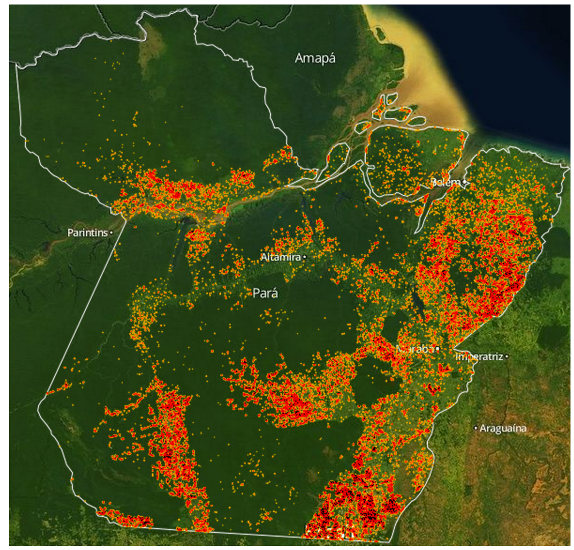
The geographic scope and high temporal frequency of the full data output is unique and there is currently no pantropical dataset with which to directly compare the output data. Furthermore, terms like forest clearing, loss, degradation and disturbance, used throughout the literature, are applied inconsistently, making direct comparisons difficult. Thus, we reiterate here that the output dataset is designed to generate alerts of forest disturbance, or forest cover change, not to estimate change area or classify the type of change. In this paper, we use data from Brazil as a comparator data set to evaluate spatial and spatiotemporal accuracy.

Spatial fit in Brazil

In Pará, Brazil, the Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite (PRODES) system generates the deforestation data set of record for the legal Amazon. Developed by the Instituto Nacional de Pesquisas Espaciais (INPE), PRODES deforestation data sets are derived each year from manually and automatically classified Landsat and CBERS imagery.[20] Polygons representing clear cutting from 2000 to 2010 were downloaded from the PRODES website [66], merged to generate deforestation masks, converted back into 60 meter resolution pixels, and aggregated to the 500 meter MODIS grid for use as the reference data set. The proposed algorithm and FCLH agreement with PRODES were assessed for the entire state of Pará, as shown in Figure 2.4a. Disturbance alerts for the entire period are shown in Figure 2.4b.

The Kappa statistic between FCLH and output alerts at the 50 percent threshold over the training period is 61.9 percent. Yet both perform similarly against PRODES, as shown in Table 2.1, which reports the producer, user, and overall accuracy of the FCLH and the output data sets against the PRODES reference data for 2000-2006 and 2007-2010 (accounting for PRODES years starting in August). The accuracy rates for the output data are reported for three different probability thresholds. The FCLH data set has already imposed an implicit confidence threshold, and the data set is binary.[59]

(a) Study areas



(b) Intensity map of alerts, 2000-2010.

Figure 2.4: Spatial fit analysis covers all of Pará. Spatiotemporal analysis covers the red box in Figure 2.4a.

For output data between 2000 and 2006 at all thresholds, user accuracy for detections is greater than 90 percent. User accuracy increases to 97.9 percent at the 80 percent threshold, reflecting more precise detections at the expense of producer accuracy, or detection sensitivity. The accuracy rates are not as high for the second time series (starting January 2006). However, visual inspection of the data indicates that false positives tend to be grouped near true positives, as will be shown in Section 2.4.

Producer accuracy is low for both FCLH and the alerts from proposed algorithm. This is not unexpected from the proposed alerting algorithm, which has an explicitly limited scope of detection and an aversion to false positives. As with previously published algorithms [75, 80], the low producer accuracy of the proposed algorithm is partly explained by the difference in resolutions (500 meter vs. 60 meter), timing of image acquisition, differing definitions of forest and forest disturbance, and the small size of many areas of disturbance.

Table 2.1 also reports the producer accuracy, highlighting the tradeoff between sensitivity and precision. Specifically, the producer accuracy ranges from 7.65 percent to 13.53 percent in a direct comparison to PRODES as the confidence threshold for the alerts ranges from 80 percent to 30 percent. There is a clear tradeoff with user accuracy, which declines from 93.38 percent to 78.72 over the same range of confidence thresholds. The implication is that as the threshold increases, the alerting system becomes more precise but less sensitive.

The accuracy assessment is a function of the threshold choice, suggesting that relevance or value of the system is dependent on the final use. A resource-constrained ranger may care more about precision than sensitivity, and place more weight on the higher threshold alerts. On the other hand, conservation groups interested in extremely early warning may find more value in hotspots of emerging, lower confidence alerts.

The accuracy assessment also depends on the definition of deforestation in the comparison data set. Table 2.1 reports accuracy for a direct pixel-by-pixel comparison with PRODES, aggregated to MODIS pixels. Any Landsat scale forest disturbance is counted as a disturbance event, even if only one 60 meter (0.36 ha) pixel is cleared within the 500 meter (25 ha) MODIS pixel. If, however, the comparison data set on forest disturbance is restricted to 500 meter pixels with PRODES deforestation covering more than half the area of the pixel, then the producer accuracy at the 80 percent confidence threshold more than doubles to 16.9 percent in the 2007-2010 period. The shifting definitions of accuracy suggest that the full range of alert probabilities should be used to support different use cases.

Table 2.1: Accuracy assessment, PRODES versus FCLH and output, 2000-2006 and 2007-2010 in Pará, Brazil

	FCLH	Output (30)	Output (50)	Output (80)
2000 to 2006				
Overall accuracy	89.59	89.55	88.89	88.09
Change producer	21.89	22.64	16.28	9.57
Change user	95.33	91.30	95.41	97.91
No change producer	99.84	99.67	99.88	99.97
No change user	89.41	89.49	88.74	87.96
2007 to 2010				
Overall accuracy	-	85.78	85.69	85.35
Change producer	-	13.53	10.84	7.65
Change user	-	78.72	87.34	93.38
No change producer	-	99.32	99.71	99.90
No change user	-	85.98	85.66	85.24

Spatiotemporal fit in Brazil

To evaluate temporal accuracy of the algorithm on an annual basis, a subsample of Pará covering approximately 100,000 hectares in the southeast corner near Vila Mandi was selected (Figure 2.4a). This area has experienced heavy deforestation, but also contains large swaths of intact forest, including portions of an indigenous area. As reported by PRODES, significant deforestation occurred from 2000 to 2008, but decreased markedly in 2009 and 2010.

To characterize temporal accuracy, pixels with probabilities 50 percent or greater were converted to points representing the pixel centroid. The date attribute for each detection point was snapped to the corresponding PRODES year. PRODES polygons were buffered by 100 meters to account in part for the different spatial resolution of the two data sets.

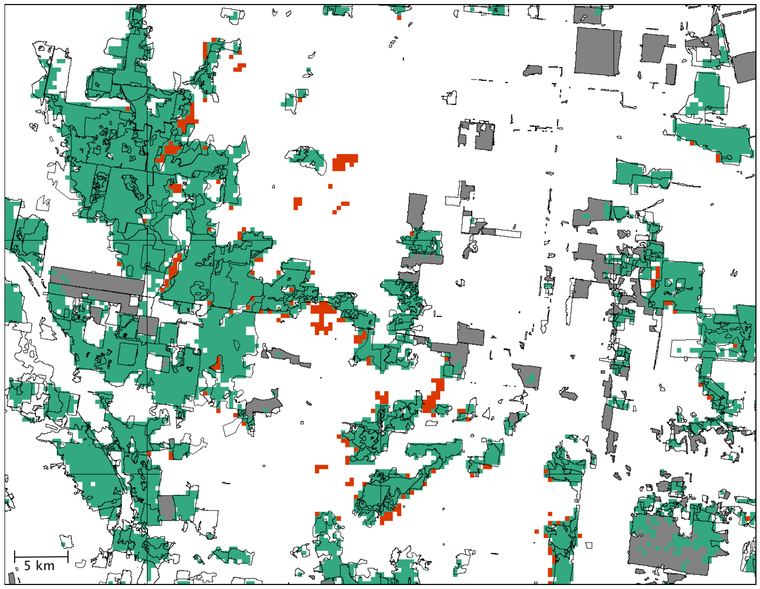


Figure 2.5: Comparison of 500 meter alerts to PRODES. PRODES deforestation from 2000 to 2006 appears as black outlines, while deforestation prior to 2000 is shown as gray polygons. False positive alerts appear as red pixels, while true positives appear in green.

Of 3,124 alerts in this area from 2007 to 2010 (Figure 2.5), 59 percent overlapped PRODES and were detected in the same calendar year, and 16 percent of detections overlapped PRODES but were detected later. Eighteen percent had no overlap with the PRODES data while percent of the alerts that overlapped with PRODES data were detected in years prior to a PRODES detection of change. Further investigation of the 562 alerts that did not initially show spatial overlap with PRODES revealed that 44 percent fell within 250 meters – half a pixel – of PRODES change polygons.

A visual inspection using Google Earth of the remaining alerts that did not intersect PRODES data showed that those alerts were typically immediately adjacent to areas of disturbance or areas of forest degradation and roads. This suggests that the output data would be useful as early warning of forest disturbance, prior to clear-cutting required for inclusion in the PRODES data set.

2.5 Discussion

Producer Accuracy

Producer accuracy of the proposed alerting system is currently low in the comparison with PRODES. The primary goal of the alerting system, however, to produce reliable alerts, i.e., high user accuracy. The results suggest that the proposed detection algorithm yields the desired features of an alerting system. A future iteration of the algorithm currently under development will move to 250 meter MODIS data – supporting detection of smaller patches of disturbance – and improve the classification data with recently released 30 meter resolution global tree cover loss data from Hansen *et al.* (2013).[57] We expect that accuracy will improve significantly as a result.

Computation

The computational demands of a timely and geographically extensive forest disturbance alerting system are non-trivial, and computational cost was a guiding factor in the development of the algorithm. However, much of the computation is at the pixel level, allowing for an "embarrassingly parallel" algorithm [76] that can be split across an arbitrarily large server cluster. The derived features of multiple pixels are only consolidated on a single (virtual) machine toward the end of the algorithm. The effect is that within a map-reduce computational framework, the data have been reduced sufficiently during the map phase such that only one reducer for each ecoregion is needed. This computational structure drastically reduces the computational cost, such that running this algorithm across the humid tropics for a single period, let alone 13 years, is eminently feasible at low cost.

Feature selection

The final set of 20 features used for classification are limited to those that are associated with an appreciable amount of independent explanatory power in the variation of historical deforestation. Through a series of trials, a large list of potential time series detection algorithms were distilled into the final set used in the algorithm.[113, 114] All experimentation is available for review in the open source code repository.

2.6 Conclusion

Timely and geographically consistent information on forest disturbance is a necessary – albeit insufficient – condition to effectively manage the use and protection of forest resources. The proposed alerting system reports timely information on large-scale forest disturbance, and while the classification rule is locally tuned the algorithm overall is globally consistent.

This has the advantage of generating consistent data that can be compared across tropical countries or ecoregions. However, more research is needed to ensure that signals from distinctive disturbance patterns at local scales can be detected, and to evaluate usefulness in detecting scattered, small-scale disturbance as opposed to industrial-scale plantation agriculture.

While this study focused on spatial agreement with detailed PRODES data, future research will examine in detail the temporal agreement of the alerts with the DETER monitoring system. Systematic accuracy assessment outside Brazil will also be pursued to characterize the algorithm's strengths and weaknesses in other regions.

The proposed alerting system is intended to complement high-resolution assessments of forest disturbance that are updated on an annual or less-frequent basis. With a relatively simple algorithm that can be implemented in parallel on a server cluster, the results are strikingly useful as a first-pass system for detecting forest disturbance.

Chapter 3

The spatial mechanism for Indonesia's failed forest conservation policy

3.1 Introduction

The 2015 United Nations Climate Change Conference prompted a multinational commitment of US$5 billion to reduce carbon emissions from deforestation.[67] The commitment doubles down on previous efforts to mitigate climate change with forest conservation. Previous research has shown, however, that historical commitments have not reduced deforestation as intended. Most notably, in 2010, Norway pledged US$1 billion in aid to Indonesia, conditional on a significant reduction of its deforestation rate.[84] The Indonesian government responded with a moratorium on new permits granted for development in primary forest.[9] Despite the resource commitment, Indonesia "hasn't seen actual progress in reducing deforestation," according to the Norwegian Minister of the Environment in 2016.[107] Indeed, Margano *et al.* (2014) show that the aggregate rate of deforestation *increased* after the moratorium was instituted; and specifically that within the first year of the moratorium, the rates of both lowland and wetland primary forest cover loss exceeded those within the previous decade.[78] This result is corroborated by both reports in the popular press and other scientific studies.[18, 107] While these inquiries report high rates of deforestation, they do not posit an explanation. Rather, they recommend that "questions concerning the moratorium as a driver of increased deforestation are worthy of investigation."[78] We offer an explanation as to why the moratorium failed, based on shifting spatial patterns of deforestation. We proceed to argue that the observed trends in deforestation patterns can be rationalized with basic production theory, and that economic incentives should be a first-order consideration when designing conservation policy.

The Indonesian moratorium was put into effect in May 2011, prohibiting the award of new licenses to clear natural forest. The moratorium was initially set to expire after two years in May 2013, but was repeatedly extended and continues to exist in 2016.[9] The moratorium originally covered 90% of primary forest in Indonesia, although the spatial extent has been revised several times.[84] Notably, licenses that were granted prior to the moratorium were not rescinded; development of natural forest was allowed to continue within existing concessions. Of the existing concessions, approximately 79% of allocated area was undeveloped in 2010, leaving a significant amount of natural forest designated for legal development.[22] In effect, the moratorium constrained development on one margin with alternatives available and abundant. The net effect of the moratorium is dependent on the rate of substitution between the two margins. We show that the substitution toward the latent, undeveloped margin exceeded the substitution away from business as usual development. We use the island of Borneo as a natural experiment to assess the causal impact of the moratorium. Specifically, we leverage the national boundary that bisects Borneo into Indonesia and Malaysia for identifying variation. Indonesia was subject to the policy, whereas Malaysia was not. The differential policy on the same island constitutes a natural experiment, with deforestation patterns in Malaysia serving as a baseline reference. The outcomes of the two groups (Indonesia and Malaysia) are compared across two time periods (before and after the moratorium) to yield a difference-in-differences estimator of the policy impact. The estimator measures the shift in spatial patterns of deforestation when the moratorium was enacted, conditional on economic variables that may confound or mask the effect. We employ a standard parametric regression, followed by a semi-parametric approach as a robustness check.

3.2 Data

We utilize data on forest clearing activity from Global Forest Watch, an online platform that tracks deforestation from satellite imagery. The Forest Monitoring for Action (FORMA) reports forest cover loss at 500-meter resolution and at 16-day intervals.[54] These data are openly accessible through a web service, so that the cost of fully replicating this study is low – arguably much lower than most empirical studies. The FORMA data set was selected from the available, remotely sensed data on deforestation because it is updated at a time scale commensurate with economic decision making. There is a trade-off between the spatial and temporal resolution of the data: higher spatial resolution information on deforestation is updated with lower frequency. With FORMA, we have data for each month between January 2008 and May 2012, yielding 218 observations for each of the forested pixels in Borneo. Forested pixels are defined by a Vegetative Continuous Field (VCF) value of 30% or greater in 2000 – a common standard for forest cover loss data based on NASA's Moderate Resolution Image Spectrometer (MODIS) sensor.[40, 59] The top panel in Figure 3.2 presents the aggregate number of pixels subject to forest clearing activity in Kalimantan, Indonesia and Sarawak, Malaysia – the states that partition Borneo. The vertical lines indicate the timing of the moratorium, when it was announced and when it was subsequently enacted. Note that the rates of both Indonesian and Malaysian clearing increased after the policy treatment; but also that the difference between the two series increased.

3.3 Estimation strategy

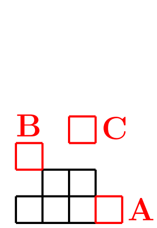


Figure 3.1: Illustration of clusters

Let ω_{it} indicate the aggregate rate of clearing activity in country $i \in \{\mathtt{idn}, \mathtt{mys}\}$ and time period $t \in \{1, \ldots, 109\}$. The aggregate rate can be partitioned into two types of clearing activity, remote (r_{it}) and peripheral (p_{it}) clearing, such that $\omega_{it} = r_{it} + p_{it}$. Peripheral clearing is defined as clearing activity that occurs on the periphery of previously existing clusters of cleared land. Remote clearing is forest cover loss is sufficiently far away from previously cleared land, such that it constitutes *new* clusters of cleared land. Figure 3.1 illustrates the two types of clearing activity. The black pixels indicate clearing that occurred in an previous period, considered an existing cluster of deforestation. The pixels labeled **A**, **B**, and **C** are considered peripheral clearing activity, whereas pixels **D** and **E** are remote clearing activity. The Let $\tilde{r}_{it} = r_{it}/\omega_{it}$ be the proportion of clearing activity in remote clusters for each country and time period. The estimating equations are given by

$$\omega_{it} = \gamma_0 + \gamma_1 m_t + \gamma_2 c_i + \tau_\omega (m_t \cdot c_i) + \epsilon_{it} \tag{3.1}$$

$$\tilde{r}_{it} = \alpha_0 + \alpha_1 m_t + \alpha_2 c_i + \tau_r (m_t \cdot c_i) + \epsilon_{it} \tag{3.2}$$

where $c_i \in \{0, 1\}$ is a binary indicator with $c_i = 1$ when $i = \mathtt{idn}$, and where $m_t \in \{0, 1\}$ is a binary indicator with $m_t = 1$ when t is after the policy treatment. We allow for two definitions of policy treatment because the moratorium was not immediately enacted when it was announced. The investment decisions of land developers may be impacted by the announcement alone, so we estimate the policy impact separately for the announcement and enactment, or January 2011 and May 2011, respectively. Then $(m_t \cdot c_i) = 1$ when the observation was in Indonesia, post-policy treatment. The coefficient estimates $\hat{\tau}_\omega$ and $\hat{\tau}_r$ indicate the causal effect of the policy on the target variable. A consistent estimator requires that $\epsilon_{it} \sim \mathscr{N}(0, 1)$, or that the unobserved variables impacting the target variables are uncorrelated with explanatory variables. One particularly concerning assumption that is implied by this normality condition is the parallel trends assumption. In context, the assumption for a consistent estimator is that the forest clearing response to economic conditions is the same for both Indonesia and Malaysia. The timing and length of response to, say, an abrupt change in global commodity prices may be different across countries. We relax the assumption with nonparametric matching of local trends, both before and after the moratorium. See the **Supplementary Material, Estimation** for further details.

3.4 Results

The bottom panel of Figure 3.2 displays the price of palm oil, which is an economically important commodity in Indonesia.[106] The price of the agricultural products determine the demand, and in turn the derived demand for cleared land – a critical input to agricultural.[47] Commodity prices across the board increased dramatically during this period.[10] Overlaid on the palm oil price, we report the *difference* in remote clearing as a proportion of total clearing between Indonesia and Malaysia for each time period ($\tilde{r}_{t,\texttt{idn}} - \tilde{r}_{t,\texttt{mys}}$ for each t). By inspection, this difference was lower after the moratorium than before the moratorium, relative to commodity prices. This visualization suggests that investment in new clusters was lower than expected in Indonesia after the moratorium. The quantified results are presented in Table 3.1. Columns (1) and (2) report the parametric $\hat{\tau}_\omega$ estimates for the two definitions of *post-policy* – when the moratorium was announced and when it was enacted. The estimates indicate that $\hat{\tau}_\omega > 0$, suggesting that the policy precipitated an increase in the aggregate rate of forest clearing activity. This result corroborates previous studies. Columns (3) and (4) report the comparable, parametric estimates for $\hat{\tau}_r$ for when the moratorium was enacted and announced. Columns (5) and (6) reestimate the equation with "snapped" time series, relaxing the assumption of strictly parallel trends near the treatment switch – which serves as a robustness check. Indeed, the estimated treatment effect is reasonably stable with $\hat{\tau}_r < 0$. The proportion of clearing activity in *remote* clusters significantly declined. The results are robust to placebo simulations.[13] In sum, the moratorium increased the aggregate rate of forest clearing activity and consolidated the spatial pattern of clearing activity.

The shift away from remote clearing was more than offset by the shift toward peripheral clearing, precipitating in a higher aggregate rate of clearing activity. The estimated parameters translate into a significant increase in aggregate deforestation, and a significant decrease in the proportion of clearing in peripheral clusters. In response to the moratorium, forest clearing activity in Indonesia became *less* fragmented but *more* intense, conditional on economic variables and relative to Malaysian forest clearing activity.

These results, when taken together, are consistent with basic production theory. A full treatment of the economic theory that underlies agricultural production in Indonesia is beyond the scope of this article, and even beyond the scope of the more extensive structure presented in the **Supplementary Material, Theory**. However, it is worth noting a minimally sufficient set of features that explain the substitution between the two margins of deforestation. Peripheral clearing activity can be conceived as development on the intensive margin, or intensifying clearing on existing clusters. Likewise, remote clearing activity is development on the extensive margin, or the creation of new centers of production. The rate of substitution between the two margins is a function of the relative cost and productivity of the cleared land – a critical input to agricultural or timber production. By definition, the moratorium raised the relative, expected cost of the extensive margin. Development within existing concessions was not *further* restricted by the moratorium. It is natural, then, that producers subject to the new constraint would shift away from extensive clearing at the margin. The net increase in aggregate clearing is consistent with peripheral land as a less productive input than remote land. While this assertion should be examined more extensively (at least to assess the magnitudes) the relation of spatial pattern to revealed prod backed by location theory in agricultural production.[cite] The basic insight is that a plot is seeded by the most productive land, and developed outward until the marginal cost exceeds the marginal benefit.[cite] Together, the differential in productivity and cost increases across margins would yield a net increase in deforestation, given that firms are profit maximizing. It is clear that limiting analysis to one island in the archipelago is not comprehensive; but the profit maximizing, economic structure that is imparted from the theory suggests that the results may be extensible to other areas.

3.5 Conclusion

The primary contribution of this paper is to offer empirical evidence that the spatial distribution of deforestation was an important factor in the 2011 moratorium's failure to control deforestation. The results are rationalized by economic theory, which provides a framework to extend the results to future policy design. The primary insight illustrated by the empirical results is that profit-maximizing agents will substitute away from more costly inputs to production. Policy that differentially impacts the costs to inputs will induce a shift toward the relatively cheaper inputs. The profit-maximizing substitution will temper the intended effect of conservation policy. It is important, then, to understand the margins of production when designing policy that is intended to rearrange production. Displacement of deforestation is often a primary policy concern across national borders (often called "leakage") but it is equally important to consider other margins across space *and* time.

3.6 Supplementary material

Theory

The empirical results are consistent with a basic economic model of production. Weitzman (1976) constructed a model of optimal extraction for a depletable resource from multiple sites. Here, cleared land is treated as a capital input to agriculture, and is stored in the form of primary forest. When primary forest is cleared, the land is activated for agricultural production. Once cleared, the land cannot return to primary forest. Agricultural investment in land can be partitioned into remote and peripheral clearing, depending on its location relative to previously cleared land. The two margins of production have different characteristics, which determine the rate of substitution between the two inputs. The aggregate effect on the input mix when the price of one input changes is a function of the marginal rate of substitution between the inputs. Weitzman's model explicates the conditions that yield higher aggregate deforestation when the expected, marginal cost of one input increases.

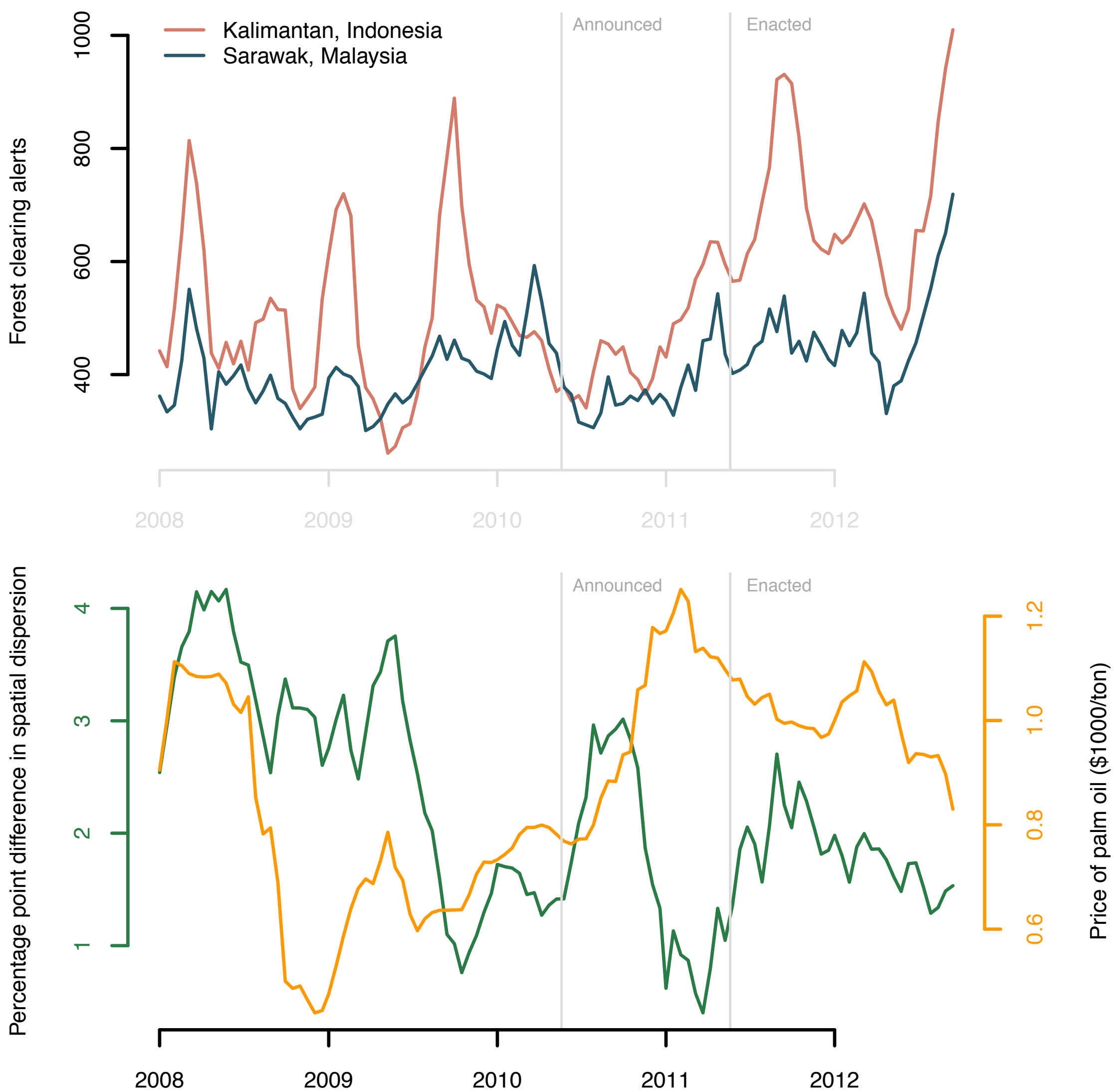


Figure 3.2: The impact of the moratorium on new forest concessions and commodity prices on the spatial distribution of forest clearing activity.

	total clearing (ω_{it})		proportion in peripheral clearing (p_{it})			
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	152.18	374.79***	3.70***	4.15***	4.27***	4.79***
	(140.81)	(125.37)	(0.62)	(0.75)	(0.67)	(0.84)
Indonesia	98.20***	94.32***	2.63***	2.37***	2.62***	2.31***
	(22.96)	(16.79)	(0.10)	(0.10)	(0.11)	(0.11)
post-policy	16.26	75.73***	−0.37***	−0.084	−0.50***	−0.11
	(26.32)	(23.50)	(0.12)	(0.14)	(0.13)	(0.16)
palm oil price	57.58*	3.29	4.41***	3.98**	2.47	2.00
	(34.13)	(30.51)	(1.51)	(1.82)	(1.62)	(2.05)
(palm oil price)^2	−3.09	−0.13	−2.09**	−2.37**	−0.62	−1.02
	(1.98)	(1.77)	(0.88)	(1.06)	(0.94)	(1.19)
policy impact	65.43***	127.62***	−0.84*	−0.54***	−1.04***	−0.72***
	(32.62)	(31.48)	(0.14)	(0.19)	(0.16)	(0.21)
R^2	0.290	0.462	0.834	0.769	0.808	0.708
Adj. R^2	0.274	0.449	0.830	0.763	0.804	0.701
Num. obs.	218	218	218	218	218	218

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.1: Empirical results.

Following Weitzman's (2003) notation, define $G(\mathbf{K}, \mathbf{I})$ as the net current cash flow of agriculture, where $\mathbf{K}$ is a vector of capital inputs and $\mathbf{I}$ is a vector of the associated fixed-cost investments. The profit-maximizing developer will choose investment to maximize the present value of $G(\cdot)$ over time. The detailed time-path of development is incidental in determining the aggregate effect of a change in the investment vector. For this type of application, Weitzman suggests an "old economist's trick" to collapse the dynamic problem to its stationary equivalent. Consider the prototypical optimal control problem:

$$\max \int_0^\infty e^{-\rho t} G(\mathbf{K}(t), \mathbf{I}(t))\, dt$$

$$\begin{aligned} \text{subject to} \quad & \dot{\mathbf{K}}(t) = \mathbf{I}(t) \\ \text{and} \quad & \mathbf{K}(t) \geq 0 \end{aligned}$$

where $\mathbf{K}(t)$ is the cumulative stock of capital inputs in time t, $\dot{\mathbf{K}}(t)$ is the change in the capital stock in time t, and $\mathbf{I}(t)$ is the instantaneous investment in the corresponding capital inputs. The parameter ρ indicates the competitive interest rate. Define $R(\hat{\mathbf{K}})$ to be the stationary rate of capital return. The stationary solution requires that there exists a time T such that for any $\epsilon_i > 0$ and $t > T$, the optimal solution maintains $\mathbf{I}(t) < \epsilon$. The vector $\hat{\mathbf{K}}$ is the capital input mix that satisfies the conditions for a stationary solution. The stationary rate of capital return is thus defined as

$$R(\hat{\mathbf{K}}) = \frac{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{K}}{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{I}} \tag{3.3}$$

The fundamental theorem of capital theory sets the stationary rate of return equal to the competitive interest rate, resulting in the the system of equations

$$R(\hat{\mathbf{K}}) = \rho \tag{3.4}$$

Equation (3.4) suggests that the investment mix is subject to an external valve, such that the decision to invest in each capital input will be weighed against the going interest rate. Note that the stationary solution may never be reached, but investment decisions will push the capital mix toward the stationary solution over time. An implication of Equation (3.4) is that, for any two inputs i and j,

$$\frac{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{K}_i}{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{I}_i} = \frac{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{K}_j}{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{I}_j} \qquad \Rightarrow \qquad \frac{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{K}_i}{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{K}_j} = \frac{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{I}_i}{\partial G(\mathbf{K}, \mathbf{0})/\partial \mathbf{I}_j} \tag{3.5}$$

The system described in Equation (3.5) demonstrates that at the stationary solution, the bang-per-buck for each input is equal. Otherwise, the profit could be increased by investing more in a different input mix, implying that the solution is not stationary. The implications are not so different from the static, two-factor production model. The present value of the marginal rate of technical substitution should equal the present value of the relative investment costs at the optimum. Two elements of the vector $\hat{\mathbf{K}}$ are cleared land on the periphery of existing clusters and cleared land that would constitute a new, remote cluster. Let $\hat{\mathbf{K}}_1$ be the stationary capital usage for peripheral land, and let $\hat{\mathbf{K}}_2$ be the stationary usage of remote land. These two inputs can be combined to produce a certain level of agricultural product at a competitive market price. The associated revenue, or value to the land developer, is the gross gain Weitzman's $G(\cdot)$ function. At this point, the dynamic problem has been sufficiently collapsed to use the standard insight from a static two-factor production model. The derivation from the dynamic problem ensures that the subsequent insight is robust up to the dynamic considerations faced by the land developer.

The implication in Equation (3.5) suggests that the optimal solution will remain on a level set of $G(\cdot)$, even after a change in relative input prices. We refer to this level set symbolically as $\bar{G}$, and is depicted in Figure (3.3). The two inputs, peripheral and remote cleared land, are highly substitutable in agricultural production, such that the level set is almost linear. The rate of substitution, or the slope of the level set, is determined by the marginal productivity of each land type. If peripheral land is, on average, less productive than remote land, then the slope of $\bar{G}$ will be shallow. If peripheral land is relatively more productive, then the slope of $\bar{G}$ will be steep. Note that the slope is not determined by the input price; but rather just the relative productivity of the inputs. Equation (3.5) provides the link between input price and productivity at the optimum. Specifically, after an input price change, the cost-minimizing input mix will remain on the level set, albeit at a different point.

Suppose that $\hat{\mathbf{K}}_1^{\prime}$ and $\hat{\mathbf{K}}_2^{\prime}$ satisfy Equation (3.5) under an initial cost regime. If the required investment for a unit remote land increases, then the optimum input mix will move along the isoquant to $(\hat{\mathbf{K}}_1^{\prime\prime}, \hat{\mathbf{K}}_2^{\prime\prime})$. This situation corresponds to the moratorium. The moratorium differentially impacted the cost of investment in remote clusters by increasing the uncertainty surrounding the maintenance of the capital input. Given that the moratorium map is uncertain and changes every six-months, the likelihood that a concession granted after May 2011 may be revoked is non-trivial. At best, the moratorium increases the uncertainty of a stranded capital asset (cleared land), and at worst, the moratorium provides leverage to local administrators to extort money from land developers. The rate of corruption surrounding land tenure and development in Indonesia has skyrocketed since the moratorium, according to various local news reports. Either way, the requisite investment for remote clusters increased relative to peripheral clusters as a direct result of the moratorium. Figure (3.3) indicates that the relative intensity of remote land decreases in response to the price increase.

The effect of the moratorium on the aggregate use of cleared land depends on the average slope of the present value isoquant, which is in turn determined by the relative productivity of the two land types. The dominant use for land cleared at large-scale in Borneo is palm oil. The palm oil production process requires that the raw kernels be processed by a central facility within 24 hours of harvesting. The kernels spoil quickly, and the proportion of spoiled kernels increases in time. The time required to transport the harvested kernels to the processing facility is substantial, given a network of poor, dirt roads. Cleared land that is close to the processing facility therefore has a higher per-acre yield of processed oil than cleared land that is further away. Land on the periphery of existing clusters is, by definition, further away from the seed of the deforestation cluster than the seed itself. New clusters in remote forest landscapes therefore have a higher productivity over the course of the plantation development. Peripheral deforestation indicates that the plantation is further along in its development than remote deforestation, which indicates initial clearing activity. The argument is, in effect, a geometric argument, and reflects the diminishing productivity of a unit of land as the plantation grows.

The characteristics of the two land types support this argument. Note that the tangency of the isocost line would imply that the cost of investment tends to be higher $\hat{\mathbf{K}}_2$ than for $\hat{\mathbf{K}}_1$. It is more difficult to prepare cleared land for agriculture at higher elevations and at higher slope, all else equal. We use the elevation data from the SRTM digital elevation model to examine the characteristics of the two land types. For both Indonesia and Malaysia, the slope and elevation are significantly higher for remote deforestation than for peripheral deforestation (with p-values less than 0.001). This result is consistent with the slope of the isoquant in Figure (3.3).

Note that, assuming the shallow isoquant in Figure (3.3), an increase in the cost of

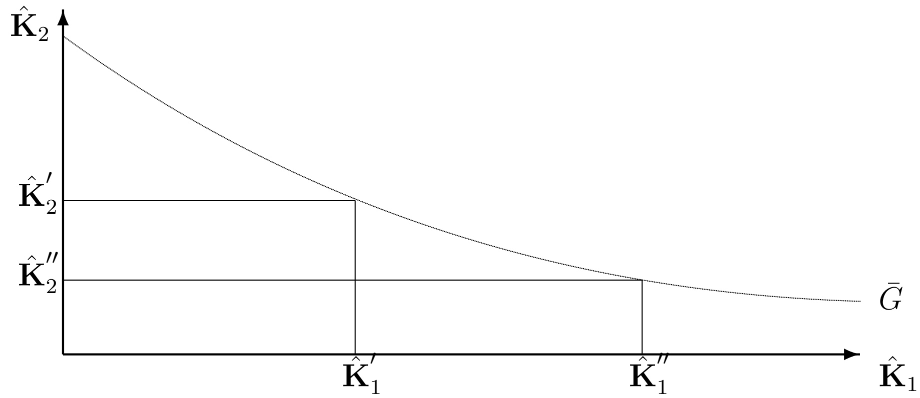


Figure 3.3: Illustration of an isoquant where the inputs exhibit a high degree of substitution in production and a low marginal rate of technical substitution.

investment in $\hat{\mathbf{K}}_2$ will yield an increase in the aggregate level of cleared land at the optimum, i.e.,

$$\hat{\mathbf{K}}_1' + \hat{\mathbf{K}}_2' < \hat{\mathbf{K}}_1'' + \hat{\mathbf{K}}_2'' \tag{3.6}$$

The decrease in $\hat{\mathbf{K}}_2$ is more than offset by the increase in $\hat{\mathbf{K}}_1$ as land developers shift agriculture to the periphery of existing clusters, despite the lower marginal production. After the dynamic investment decisions are collapsed to their stationary equivalents, there is nothing particularly deep about this structure. The empirics indicate that, indeed, more land was cleared in the aggregate after the moratorium, even with a decrease in remote clusters.

Estimation

The estimation strategy is based on the difference-in-difference (DiD) approach, a quasi-experimental method to measure the impact of an abrupt policy change. DiD estimation is common in economics because of its simplicity, but it comes with limitations.[13] An especially concerning limitation in this context is the assumption of parallel paths. The assumption insists that, conditional on the covariates, the potential path in the treatment group is the same as that for the control group. Here, the parallel path assumption requires that (among other things) the patterns in deforestation in the treatment and control groups respond in lock step to exogenous cofactors. It is more likely, however, that producers in Malaysia may respond differently than producers in Indonesia to interest rates, commodity prices, or other economic factors that drive the demand for inputs to agricultural production. We relax the parallel path assumption by employing methods in dynamic time warping (DTW), which are often used in speech processing and gesture recognition.

The basic empirical model is specified in Equations (3.1) and (3.2). The target variables, ω_{it} and $\tilde{r}_{it}$, are directly observed from the raw data. There is no processing or smoothing of the measures prior to estimation. The parallel paths assumption applies to these two variables. Consider, however, the top panel of Figure 3.2. The short-term variation in the

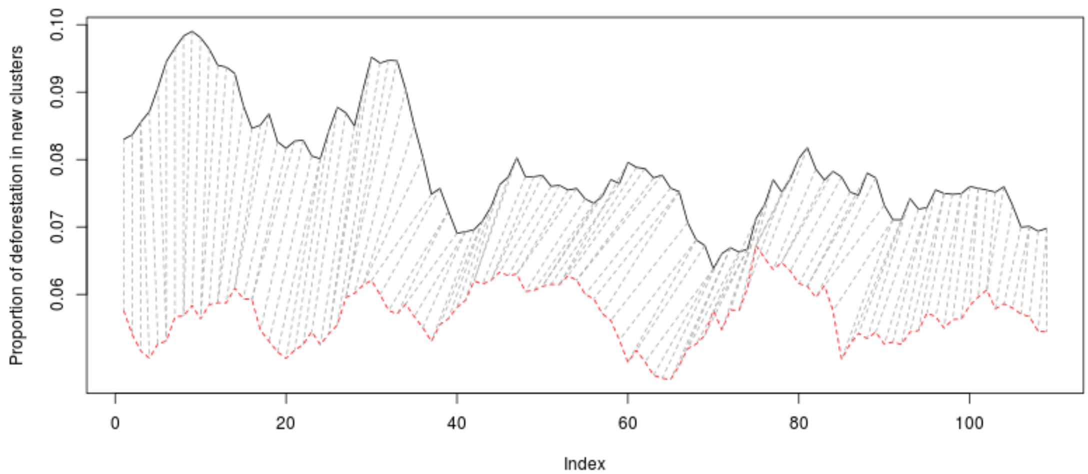


Figure 3.4: The results of DTW matching.

aggregate rates of deforestation tend to move together across the two countries. A small bump in deforestation in Indonesia is often accompanied by a small bump in deforestation in Malaysia, albeit shifted or elongated through time (the x-axis). The "small bumps" are clearly induced by an external, possibly unobserved factor. If the external force influences the target variables (a) around the same time as the policy change, and (b) differentially in the control and treatment groups, then the parallel path assumption may be violated. For example, if the response to an external force is more delayed in Indonesia than in Malaysia, then the effect may be falsely ascribed to the policy – just by virtue of the differential response time.

A possible remedy is to preprocess the target variables to remove transient differences in the series that are not caused by the policy. We are interested in "snapping" the time series for the treatment group to the time series of the control group, based on short-term variation. The objective is to remove any variation induced by heterogenous responses to exogenous factors that may bias the DiD estimator. To this end, we employ DTW methods. DTW effectively warps the time series, searching for the arrangement of the treatment series that minimizes the difference between the treatment and control series, based on short-term variation. The minimization problem is restricted to short-term variation by enforcing a moving window of a given length, although there are many ways to perform dynamic warping. We rely on the `dtw` package in `R`.[] All supporting code is posted online, with assumptions documented. The results of the warping are presented in Figure 3.4.

The first difference – between treatment and control – is determined by the match shown in Figure 3.4. The matched series becomes the basis for the second difference – before and after treatment. In some sense, then, the target variables for the treatment group (Indonesia) is preprocessed before the models in Equations (3.1) and (3.2) are estimated.

Note that DTW will not solve the problem of non-parallel paths that are persistent over time. Rather, it only addresses the edge effects – the potential time shifting close to the time of policy implementation. The DTW procedure should be viewed as a robustness check, alongside the placebo tests also used to check the stability of the results.

Chapter 4

Economic dynamics and forest clearing: A spatial econometric analysis

4.1 Introduction

Forest clearing is an enormous contributor to global warming, accounting for some 15% of annual greenhouse gas emissions.[31] Most forest clearing occurs in developing countries that have limited resources and regulatory capacity. Since these countries understandably focus their energy and resources on poverty alleviation, their support for forest conservation will be weak as long as forested land has a higher market value in other uses. Under these conditions, many actors will continue clearing their forested land unless they are given conservation payments that match or exceed the opportunity cost of the land. This economic insight has led to the establishment of REDD+ (Reducing Emissions from Deforestation and Forest Degradation in Developing Countries), an international mechanism for compensating proprietors for forest conservation.

While the conceptual foundations of REDD+ are straightforward, its actual success will depend on program designs tailored to the economic dynamics of forest clearing in tropical forest countries. Stern *et al.* (2006) and Enkvist *et al.* (2009) have asserted that carbon emissions abatement from forest conservation is generally lower-cost than abating emissions from fossil fuels.[41, 105] For example, the UNFCCC's estimate of CO2 emissions from forest clearing (5.8 Gt) implies an aver-age abatement cost of only $2.10/tonne.[1] While such general

A version of this chapter first appeared in the *Ecological Economics* (2013) Vol. **85**, with David Wheeler, Robin Kraft, Susmita Dasgupta, and Brian Blankespoor.

[1]This estimate is based on the opportunity cost of forested land.

estimates inform the policy dialog, they cannot guide specific conservation programs because the economic returns to forest clearing vary widely over space and time. For many agents, land clearing for production is a lumpy investment driven by expectations about future prices and demand conditions. These expectations and investor interest in land clearing will change with market conditions, creating problems for REDD+ incentive programs based on fixed payments, or traditional forest conservation programs that focus on protection of designated areas.

Extensive theoretical work has considered the role of economic dynamics in forest clearing. However, relevant empirical research has been severely hindered by the lack of spatially-disaggregated time series data. Until recently, the translation of satellite images into credible estimates of forest clearing has been so cumbersome that updates have taken years for many countries. As a result, empirical research has focused on multi-year clearing and its relationship to demographic and geographic factors. The forced exclusion of fluctuating economic conditions from most studies has left researchers and policymakers uncertain about the timing, magnitude and spatial incidence of their effects. At the same time, long lags in forest monitoring have left conservation managers blind to new threats in many areas, and program evaluators unable to provide timely assessments of conservation measures.

Faced with these limitations, donors and governments have traditionally focused on legally-protected areas. Mounting empirical evidence suggests that some forms of protection have significantly reduced forest clearing, but it has remained difficult to defend fixed conservation frontiers with limited monitoring information.[87] In addition, the advent of massive REDD+ payment programs will force donors to account for billions in expenditure for targeted reductions of carbon emissions from forest clearing. Such programs are not likely to survive taxpayer scrutiny unless they incorporate much more accurate and timely information.[2]

Fortunately, the state of the art in forest monitoring is now advancing rapidly. Work by Hansen *et al.* (2008), Souza (2006), Townshend *et al.* (2008), Hammer *et al.* (2009), Asner (2009), Jarvis (2009) and others is creating new, high-resolution forest information systems based on NASA's MODIS (Moderate Resolution Imaging Spectrometer) and Landsat programs, as well as airborne light detection and ranging (LiDAR).[59, 101, 108, 54, 8] Drawing on advances for the MODIS system, Hammer *et al.* (2009) have recently published FORMA (Forest Monitoring for Action), a monthly database for forest clearing in Indonesia at 1 km resolution since 2005.

Equipped with the vast new information resource available from FORMA, this paper focuses on Indonesia for an in-depth econometric study of economic dynamics and forest clearing at a high level of spatial and temporal disaggregation. Economic dynamics are clearly important for the Indonesian case, which is heavily driven by forest clearing for palm-oil and wood-processing exports to fast-changing Asian markets.

The remainder of the paper is organized as follows. Section 4.2 reviews the extensive

[2]In addition, land which is most appropriate for REDD+ programs may not be in currently-protected areas.

prior research on the economics of forest clearing. In Section 4.3 we discuss past limitations imposed by data scarcity, and the implications of recent technical advances for expanded research in this domain. Section 4.4 introduces the FORMA database, a critical contributor to the expanded prospectus, and uses FORMA data to investigate patterns of national and local forest clearing in Indonesia since 2005. In Section 4.5, we develop a model of forest clearing based on expected profitability calculations by potential investors in commercial production on currently-forested land. Section 4.6 describes the available data and Section 4.7 specifies a model for econometric panel estimation. We present and discuss our econometric results in Section 4.8, while Section 4.9 summarizes and concludes the paper. Section 4.10 provides additional details on the econometric strategy, beyond the scope of the core chapter.

4.2 Prior research

Previous empirical research has assessed the relative importance of numerous factors that may influence the conversion value of forested land. These include local population scale and density, distance from markets, the quality of transport infrastructure, agricultural input prices, physical factors such as topography, precipitation and soil quality, and zoning into categories that include protected areas. The results are generally consistent with a model in which the conversion of forested land varies with potential profitability.[3] Among studies that control for protection zoning, Nelson and Chomitz (2009) provide the most rigorous and comprehensive assessment.[87] Their finding for the tropical forest biome — that protected areas have less land clearing, *ceteris paribus* — supports the specific results of Gaveau *et al.* (2009) for Sumatra, Indonesia.[46]

The existing research provides many useful insights about long-run drivers of forest clearing. Nelson and Chomitz (2009) and Rudel *et al.* (2009) have studied land-use change across countries over multi-year intervals.[87, 96] Within counties, numerous econometric studies have estimated the impact of deforestation drivers across local areas during multi-year intervals. Some studies have used aggregate data for states, provinces or sub-provinces (e.g., studies for Brazilian municipios by Pfaff (1997) and Igliori (2006), and Mexican states by Barbier and Burgess (1996)).[89, 64, 11] Many studies have also used GIS-based techniques to obtain multi-year estimates at a higher level of spatial disaggregation (e.g., Cropper *et al.* (1999, 2001) for Thailand; Agarwal *et al.* (2005) for Madagascar; Deininger and Minten (2002); Chowdhury (2006) and Vance and Geoghegan (2002) for Mexico; Kaimowitz *et al.* (2002) for Bolivia; and De Pinto and Nelson (2009) for Panama).[34, 35, 2, 39, 27, 112, 69, 37] In rarer cases, studies have used annual national or regional aggregate time series over extended periods (e.g. Zikri (2009) for Indonesia; Ewers *et al.* (2008) for Brazil). These

[3] For detailed summaries, see particularly Chomitz (2006); also Igliori (2006) and Wunder and Verbist (2003).[26, 64, 118]

studies are hindered by limited degrees of freedom, since they must control for many factors and available observations are annual at best.[119, 43]

While econometric work on long-run forest clearing drivers is well-advanced, data problems have limited treatments of short-run economic dynamics to theoretical work and simulation modeling. Arcanda *et al.* (2008) and others have studied the theoretical relationships between economic drivers and forest clearing.[7] Notable simulation exercises include Cattaneo (2001) for Brazil and San *et al.* (2000) for Indonesia.[23, 97] The latter study investigates economic drivers of forest clearing in Sumatra using a multisectoral, multiregional computable general equilibrium model. Since short-period data were not available to the authors, they use changes in deforestation-related sectors (e.g. plantation agriculture and wood products) as proxies. While the results are interesting and suggestive, they depend entirely on the researchers' specification of CGE parameters, and are unable to provide any estimates for areas smaller than provinces.

While more temporally- and spatially-disaggregated studies have been awaiting the advent of better data, econometric theorists have been laying the groundwork for efficient estimation of more highly- disaggregated models. Notable contributions to the literature on computable approaches to spatial econometric analysis have been made by Agarwal *et al.* (2002), Anselin (2001, 2002), Barrios *et al.* (2012), Kapoor *et al.* (2007), Kelejian and Prucha (1998, 2010), and Kelejian *et al.* (2004).[1, 4, 5, 12, 70, 71, 72, 73]

4.3 Expanding the scope of work

Past contributions

Many estimates of forest clearing are based on remotely-sensed data that have been available in various forms for decades. Perhaps the most impressive contribution has been made by Brazil's PRODES (2009), which has provided yearly maps of Amazonian forest clearing since 1988. Since 2004, these have been augmented by twice-monthly estimates from Brazil's DETER system.[4] Another noteworthy Brazilian contribution is Imazon's Forest Transparency Initiative, which has utilized MODIS data to produce and rapidly disseminate information about forest clearing in Mato Grosso State (Souza *et al.*, 2009).[102]

Several global-scale studies of forest clearing have been reported in scientific journals. Although they have laid the groundwork for global monitoring, these studies have not replicated the Brazilian contribution by providing updated, online reporting. Nor are they accessible to non-specialists who do not have a deep understanding of Geographic Information Systems (GIS) and remote sensing techniques. As Grainger (2008) has noted, tracking the long-term trend in tropical forest clearing has been problematic.[51] Hansen *et al.* (2008) identify global forest clearing in humid tropical forests using MODIS and Landsat images for the period

[4] Detailed descriptions of PRODES and DETER are available from Brazil's National Institute for Space Research (INPE).

2000-2005.[59] Mulligan (2008) uses remotely sensed data for assessment of land-use changes in and around protected areas from 2000 to 2005.[83] Carroll *et al.* (2006) identify changes in vegetation cover from 2001 to 2005.[56]

Several institutions provide detailed information on forest clearing with varying quality, but they have not attempted continuous global monitoring at high resolution. The FAO provides a detailed Global Forest Resources Assessment at the country level, updated at 5-year intervals. The World Resources Institute (WRI) has published detailed maps of forest clearing hotspots in Latin America, Asia and Africa for the period 20002006. The website maintained by Global Forest Watch has provided global information, but with non-standardized spatial and temporal coverage of different datasets by country, infrequent updates, and a map interface that does not permit integrated global views. In summary, outside of Brazil, policy researchers have not been able to access panel databases sufficient for in-depth investigations of country-specific dynamics.

Recent advances

Recently, a group affiliated with the Center for Global Development, the World Resources Institute and the University of Maryland has laid the groundwork for a global database that will permit much more rigorous empirical work on the economic dynamics of forest clearing. Called FORMA (Forest Monitoring for Action), the system utilizes data recorded daily by the Moderate Resolution Imaging Spectrometer (MODIS), which operates on NASA's Terra and Aqua (EOS PM) satellite platforms. Although its signal-processing algorithms are relatively complex, FORMA is based on a common-sense observation; tropical forest clearing involves the burning of biomass and a pronounced temporary or long-term change in vegetation color, as the original forest is cleared and replaced by pastures, croplands or plantations. Accordingly, FORMA constructs indicators from MODIS-derived data on the incidence of fires and changes in vegetation color as identified by the Normalized Difference Vegetation Index (NDVI). It then calibrates to local forest clearing by fitting a statistical model that relates the MODIS-based indicator values to the best available information on actual clearing in each area.

FORMA incorporates biological, economic and social diversity by dividing the monitored territory into blocks and separately fitting the model to data for the parcels in each block. The dependent variable for each pixel is coded 1 if it has actually experienced forest clearing within the relevant time period, and 0 otherwise. The MODIS-based indicator values are the independent variables. For all tropical countries except Brazil, the best identification of recent forest clearing has been published in Proceedings of the National Academy of Sciences by Hansen *et al.* (2008), who provide estimates for 500m parcels in the humid tropics.[59] FORMA is calibrated using the map of forest cover loss hotspots (henceforth referred to as the FCLH dataset) published by Hansen *et al.* for the period 2000-2005.

Using the FCLH pan-tropical dataset for 2000-2005, FORMA fits the calibration model to observations on forest clearing for 1 km^2 cells in each country and ecoregion. As Hammer *et al.* (2009) document, the model's predicted spatial probability distribution provides a very close match to the spatial incidence of FCLH forest clearing.[53] FORMA then applies the fitted model to monthly MODIS indicator data for the period after December 2005. The output for each month is a predicted forest clearing probability for each 1 km^2 parcel outside of previously-deforested areas, as identified in the FCLH map.

FORMA selects parcels whose probabilities exceed 50%.[5] It calculates the total number of selected parcels within a geographic area to produce an index of forest clearing activity in that area. Even small geographic areas can include thousands of 1 km cells, so error-averaging ensures robust index values.[6] FORMA's outputs consistently aggregate to forest clearing indicators for subnational, national and regional entities.

While FORMA represents a significant advance in monitoring forest clearing dynamics, its limitations must also be understood. FORMA's identification of parcels where significant forest clearing has occurred does not imply that these parcels have been completely deforested. As we have noted above, FORMA's probability estimates are generated by a model fitted to data from Hansen *et al.* (2008), who assign values of 1 to pixels where forest clearing at agro-industrial scale (FCAS) is estimated to have occurred during 2000-2005, and 0 otherwise. By implication, each pixel with FCAS has experienced significant, but not necessarily total, clearing. Of necessity, FORMA's estimates reflect the same logic. After aggregation to the kabupaten level and time-differencing, each FORMA observation in our Indonesian panel database should be interpreted as the monthly change in 1 km parcels that have experienced FCAS with a probability greater than 0.50.

This new dataset permits panel estimation of spatially-disaggregated forest clearing models that incorporate short- and medium-term economic dynamics, as well as previously-studied demographic and geographic determinants of forest clearing. It also permits explicit consideration of differences in clearing dynamics across land-use categories, including protected areas and areas zoned for commercial exploitation. The results can provide important new insights into the behavior of forest clearing agents who constantly adjust expectations as market conditions change.

Such econometric analysis can provide two major benefits for conservation policymakers and project planners. First, its incorporation of important economic variables will provide

[5]FORMA explicitly relates forest clearing to changes in vegetation color, the incidence of fires and variations in rainfall. Changes in its predicted forest-clearing probabilities are therefore best interpreted as changes in the strength of the composite signal provided by these variables. Translation of the continuous signal to an optimal single indicator of significant forest-clearing requires a judgment about the relative importance of two types of error: False positives, which incorrectly identify clearing in areas where it has not occurred; and false negatives, which fail to identify clearing in areas where it has occurred. In recent research, Hammer *et al.* (2012) have used high-resolution information from PRODES, Brazil's annual deforestation report, to investigate changes in the incidence of false positives and negatives as FORMA's identification threshold for forest clearing is increased or decreased from the current threshold probability of 0.50. The results indicate a relatively low incidence of false positives at 0.50, which decrease at a modest rate as the threshold probability is increased. At the same time, increasing the threshold probability toward 1.00 increases the incidence of false negatives at a rapid rate. On balance, the evidence to date does not provide compelling support for alteration of the 0.50 threshold probability.

[6]For example, a square area 50 km on a side contains 2500 1 km cells.

measures of their relative significance as drivers of forest clearing. By providing a better understanding of economic incentives in this context, the results can inform the design and implementation of incentive payment systems for REDD+ programs and similar arrangements. Second, the estimation of dynamic econometric models will provide a quantitative foundation for tracking area-specific risks of forest clearing as economic and other conditions change.

4.4 Trends in Indonesian forest clearing (2005-2010)

The advent of monthly forest clearing data permits a much more timely, detailed view of forest clearing than has previously been possible. In this section, we use the FORMA Indonesia database to develop a detailed view of forest clearing patterns since 2005.

National trends

Figure (4.1) displays FORMA-estimated indices for monthly forest clearing in Indonesia from December 2005 to December 2010. The graph indexes monthly changes on the left axis and annualized changes on the right axis. The monthly series displays marked seasonality; annualizing the data with a 12-month moving sum removes the seasonal component, revealing a broadly-declining trend during the past five years.[7]

Regional trends

Changes in the index of national forest clearing summarize complex patterns of change within Indonesia. In Figure (4.2), we investigate relative changes at the kabupaten level by dividing the total of monthly index values in 2010 by the total in 2006. We color the map dark for ratios greater than one (larger index values in 2010 than in 2006); lighter for ratios equal to one (no change); and lightest for ratios less than one (smaller index values in 2010).

Clear interregional patterns are evident in Figure (4.2): Forest clearing activity has increased in northern Sumatra and decreased in the southern and central parts of the island. Kalimantan exhibits increased activity in the west and north, and either constant or decreased activity in the south-central and eastern areas. Increased activity also appears in central Sulawesi, and parts of western and southern Irian Jaya.

Figure (4.2) displays patterns of change without providing any information about the scale of activity. Figure (4.3) provides an alternative view by identifying Indonesian kabupatens whose index values are top-ranked among 1372 secondary administrative units in Southeast Asia. On this map we color units dark if they are in the top 20, lighter if they are in the next 30, and lightest otherwise. In 2006, Indonesia's highest regional index values

[7] The moving sum is equivalent to a 12-month moving monthly average, multiplied by 12. For each month, we compute the moving series using that month and the previous 11 months.

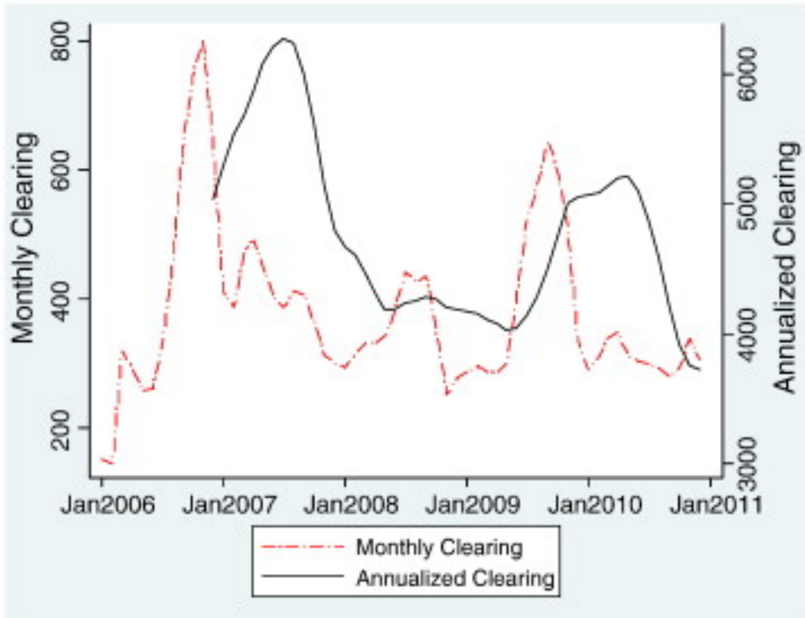


Figure 4.1: Large-scale forest clearing in Indonesia (2005-2010)

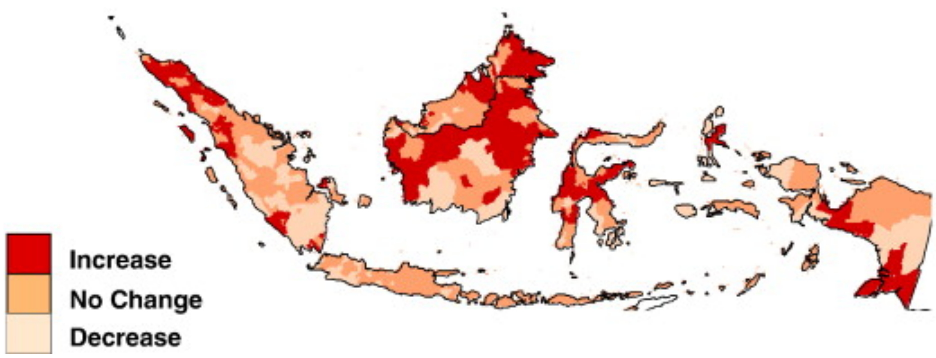


Figure 4.2: Change in forest clearing index value: 2006 vs. 2010

were concentrated in east-central Sumatra, southern Sumatra and south-central and extreme northwest-central Kalimantan.

Substantial changes are evident by 2008, with a reduction of top-ranked areas in southern Sumatra, some new areas in northern Sumatra, and a shift westward in southern Kalimantan. These pat- terns become more pronounced in 2010, with continued shrinking of the clusters in south Sumatra and southern Kalimantan, and in- creased cluster size in the northern frontier area of Kalimantan.

While parts of the pattern have significantly changed since 2006, some parts have also remained stable, with large top-ranking clusters persisting in east-central and southern Sumatra.

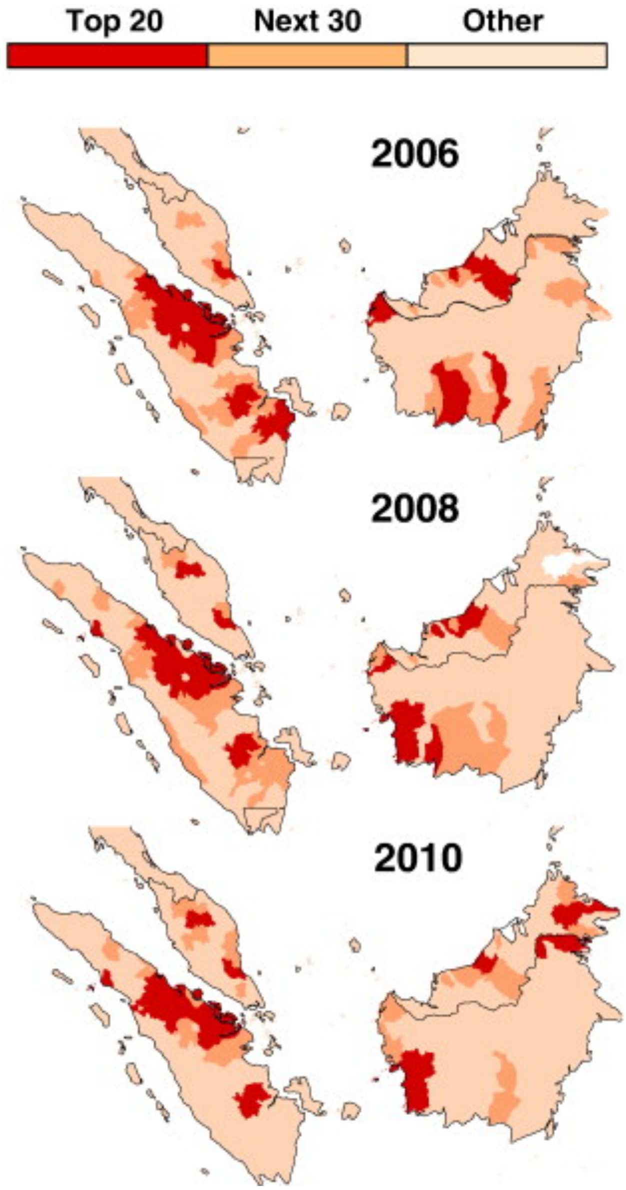


Figure 4.3: Sumatra and Kalimantan: forest clearing rank of kabupatens. Among 1372 secondary administrative units in Southeast Asia.

Provincial trends

Figure (4.4) illustrates trends in annualized forest clearing indices for the five Indonesian provinces with the largest index values in January, 2007. In the first year, the series are dominated by the clearing indices of Riau (east-central Sumatra) and Central Kalimantan. The more recent period has witnessed strong convergence, with steep declines in both Riau and Central Kalimantan, along with increases in West Kalimantan and North Sumatra and

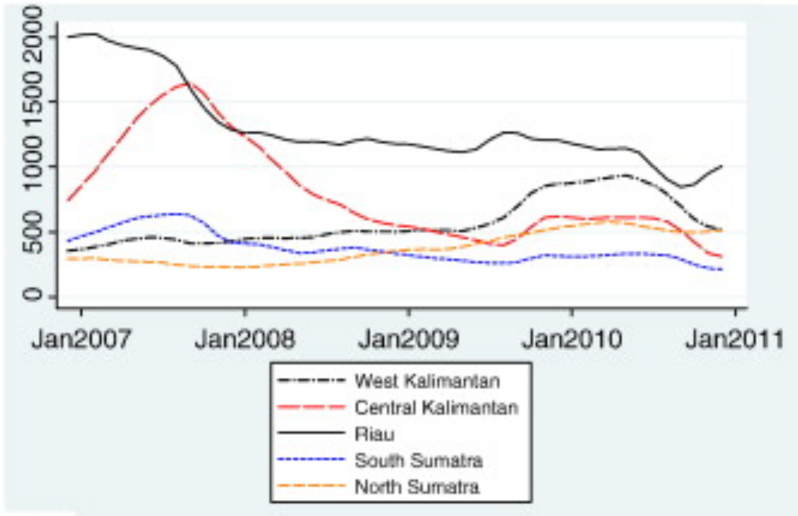


Figure 4.4: Annualized forest clearing index values: top five Indonesian Provinces in January 2007.

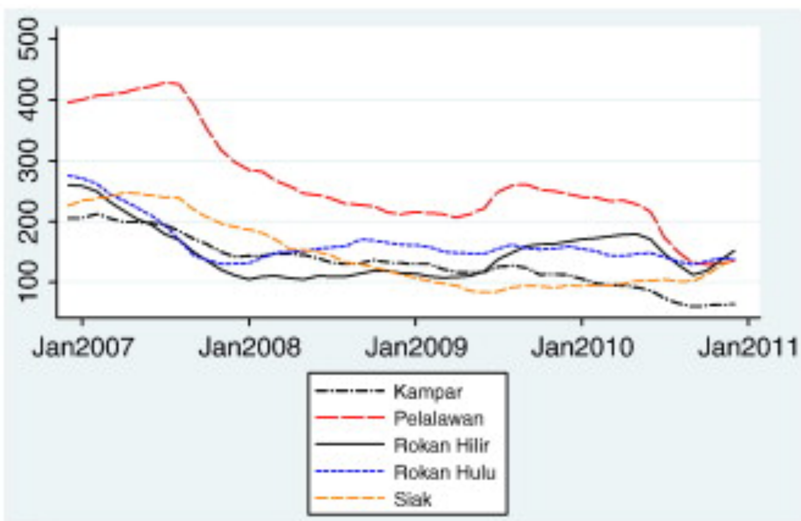


Figure 4.5: Annualized forest clearing index values: top five kabupatens in Riau.

a more modest decline (in ab- solute terms) in South Sumatra.

These changes are reflected in the overall patterns displayed in Figure (4.2). Riau has experienced the greatest decline, although it retains the highest index value in Indonesia. Within Riau, Figure (4.5) shows that convergence of kabupatens has also occurred. In January 2007, Pelalawan dominated the other top-ranking units in Riau. All five units have experienced a decline since then, but it has been most pronounced in Pelalawan. Slower declines in three of the other units (Rokan Hilir, Rokan Hulu and Siak) have brought them to approximate parity with Pelalawan.

4.5 Model specification

To explore the determinants of the patterns revealed by Figures (4.1 - 4.5), we mobilize a FORMA panel dataset for Indonesia that includes monthly observations on forest clearing from January 2006 to December 2010 for over 950,000 1-km parcels.[8] This dataset permits construction of a large panel database at the kabupaten level.

We posit an intertemporal model in which the representative proprietor or occupant of a forested area considers the relative profitability of maintaining or clearing the area. In each period, the agent compares the present-value profitability of sustainably-harvested forest products with the clear-cut value of forest products, plus the cleared land's present-value profitability in its best use (e.g., plantation (palm oil, wood products), pasture, smallholder agriculture, and settlement). Forest clearing dynamics are likely to be quite different in cases where commercial exploitation rights are well- or poorly-defined.

The decision to hold or clear a parcel depends on many factors, including expected revenues, input costs and the exchange rate. Expected revenues are a function of expected international prices and demand, particularly for wood products and palm oil in the Indonesian case. These factors and the exchange rate are constant across areas but vary over time, while other factors vary over both areas and time.

The relative significance of forest clearing determinants may well depend on the nature of particular forested areas, because they may be occupied by different types of agents with different incentives. A recently-produced GIS database enables us to separate Indonesia's forested land into areas zoned for activity in five categories: protected natural forest, palm oil plantations, timber plantations, logging concessions and unzoned areas.

In our specification, the relative profitability of forest clearing for agriculture or settlement, for the representative proprietor or occupant in area i, time t, is given by:

$$\pi_{it}^e = \pi_{it}^e(p_t^e, q_t^e, n_{it}, t_{it}, c_{it}, i_t^e, x_t^e, g_t^e, r_t^e, u_t^e, h_t^e, y_t^e, w_t^e, s_i) \tag{4.1}$$

[8] Indonesia's natural forest area in 2000 was 951,160 km^2.[31]

π_{it}^e	expected relative profitability of forest clearing
p_t^e	vector of expected prices for relevant products
q_t^e	vector of expected demands for relevant products
n_{it}	rupiah-denominated input cost per unit of output
t_{it}	transport cost per unit of output
c_{it}	communications cost per unit of output
i_t^e	expected interest rate
x_t^e	expected exchange rate (rupiah/dollar)
g_t^e	quality of governance from investors' perspective
r_t^e	regulatory quality
u_t^e	officially-designated use (among the five categories specified above)
h_t^e	population density
y_t^e	unskilled wage rate
w_t^e	precipitation (forest-burning is more difficult when rainfall is heavier)
s_i	slope of the terrain

In this specification, the expected profitability of forest clearing relative to forest conservation increases with expected revenues for outputs produced on cleared land, which in turn depends on the expected prices and levels of demand for those outputs. Expectations adjust to changes in prices and quantities with product-specific lags. The expected profitability of clearing declines with increases in the unit costs of low-skill labor, capital, transport and communications. Forest-sector outputs are traded internationally; dollar-denominated input costs decrease (and profitability increases) when the exchange rate increases. Governance has two anticipated effects in this context. Local government efficiency and integrity should increase the expected profitability of forest-sector production, which will in turn promote forest clearing. On the other hand, greater regulatory effectiveness may discourage forest clearing in protected areas, if local governments are actually concerned about clearing.

We posit effects for local structural factors as well. Higher population density should increase the demand for cleared land. Production will be more costly on more steeply-sloped land, and clearing will be more costly in areas (and months) with heavier precipitation.

4.6 Data

We have drawn the data for our estimation exercise from a variety of sources. All forest clearing information for the period December 2005 - December 2010 comes from FORMA, which, as we have previously noted, provides indicators of large-scale forest clearing at 1 km resolution for all forested areas in Indonesia. To index determinants of expected revenue, we use international market prices and world demand for hardwood sawlogs (our proxy for tropical wood products) and palm oil. We draw the price series from IMF data[9] and adjust to constant-dollar prices using the US GDP deflator.[10] Data on world palm oil production have been provided by the US Department of Agriculture,[11] while world production statistics for sawlogs have been obtained from the FAO.[12]

Among local input price variables, the only available time series is a proxy for communications cost. Our index is an estimate of mobile phone coverage that we construct from high-resolution data provided by GSM World, Inc.[13] In addition, we include three cross-sectional proxies: (1) An index of the economic opportunity cost of forested land developed by Resources for the Future and Climate Advisors; (2) Estimated travel time to the nearest city of 50,000 or more people in the year 2000, from Nelson (2008). The travel time data are available at a high level of spatial resolution. For this exercise, we estimate kabupaten means, as well as standard deviations to control for within-kabupaten variation. (3) The average poverty rate in 2000, a proxy for the prevalence of low-skill, low-cost labor, obtained from the World Bank.

Our interest rate series is the one-month rate on notes issued by Bank Indonesia,[14] adjusted for inflation using annual estimates from the World Bank.[15] We have drawn exchange rate data from OANDA's historical database.[16] Our land-use data are from a high-resolution digital map of Indonesia. All indices of governance quality have been drawn from the KPPOD survey database for Indonesia. Our precipitation data come from the PREC/L (PRECipitation REConstruction over Land) database as described by Chen *et al.* (2002).[25] The terrain slope data are kabupaten averages from Verdin *et al.* (2007).[115] Since the underlying slope data are at higher resolution, we also calculate standard deviations to control for within-kabupaten variations.

4.7 Econometric specification

The model developed in Section 4.5 relates the expected profitability of forest clearing to the determinants incorporated in Equation (4.1). We translate this general function to an estimating equation in two steps. First, we assume that forest clearing varies directly with its expected probability. Second, we replace the general determinants in Equation (4.1)

[9] The relevant IMF data series are prices for Hard Logs, Best Quality Malaysian Meranti, import price Japan, US$ per cubic meter; and Palm Oil, Malaysia Palm Oil Futures (first contract forward) 45% FFA, US$ per metric tonne. Source: IMF Primary Commodity Prices.

[10] Source: Bureau of Economic Analysis, Table 1.1.9. Implicit Price Deflator for Gross Domestic Product.

[11] US Department of Agriculture, Foreign Agricultural Service.

[12] Source: FAO, World production of sawlogs and veneer logs.

[13] Mobile phone coverage is reported at frequent intervals; we interpolate to produce a full monthly dataset.

[14] Source: Division of Economic & Monetary Data & Information Processing, Bank Indonesia, Table 1.25, Sertifikat Bank Indonesia, 1 Bulan.

[15] Source: World Development Indicators.

[16] Source: OANDA, Historical Exchange Rates.

with the specific variables that are relevant and available for this exercise. We specify the estimating Equation (4.2) as follows:

$$
\begin{aligned}
\log(\texttt{Clear})_{it} = {} & \beta_0 + \beta_1 \log(\texttt{PalmPrice})_{i,t-i} + \beta_2 \log(\texttt{LogPrice})_{i,t-j} \\
& + \beta_3 \log(\texttt{PalmQuant})_{i,t-k} + \beta_4 \log(\texttt{LogQuant})_{i,t-1} + \beta_5 \log(\texttt{OppCost})_i \\
& + \beta_6 \log(\texttt{PovRate})_i + \beta_7 \log(\texttt{MobileCov})_i + \beta_8 \log(\texttt{MeanTravT})_i \\
& + \beta_9 \log(\texttt{sdTravT})_i + \beta_{10} \texttt{IntRate}_{t-m} + \beta_{11} \log(\texttt{XRate})_{t-n} \\
& + \beta_{12} \log(\texttt{InvGovQual})_i + \beta_{13} \log(\texttt{RegQual})_i + \beta_{14} \texttt{LogPct}_i \\
& + \beta_{15} \texttt{TimbPct}_i + \beta_{16} \texttt{PalmPct}_i + \beta_{17} \texttt{ProtPct}_i \\
& + \beta_{18} \log(\texttt{PopDens})_{it} + \beta_{19} \log(\texttt{Precip})_{i,t-w} + \beta_{20} \log(\texttt{MeanSlope})_i \\
& + \beta_{21} \log(\texttt{sdSlope})_i + \beta_{22} \log(\texttt{Forest2000}) + \epsilon_{it} \qquad (4.2)
\end{aligned}
$$

Data specifications are summarized in Table 4.1. The model includes six short-term market variables: prices and quantities for palm oil and sawlogs, the interest rate and the exchange rate. We expect the standard investment calculus to produce a negative effect for the real interest rate. Palm oil and sawlogs are traded internationally; their expected profitability and associated forest clearing should be positively associated with the rupiah/dollar exchange rate, because increases in that rate will lower the dollar cost of local inputs while leaving the dollar-denominated prices of exports unchanged. The expected prices and global demands for palm oil and sawlogs are positively associated with expected profitability, *ceteris paribus*, so they should be positively associated with forest clearing as well. We have no basis for *a priori* specification of appropriate lags for expectations-formation; they are quite likely to differ by variable. During the estimation process, we retain the single lagged value of each variable that provides the best fit. We would expect the palm oil price variable to have the shortest lag because our measure is the futures price.

While all six market variables, rainfall and mobile phone coverage are observed in time series, we have only single cross-sectional observations for the four other proxies for local input prices. We expect the agricultural opportunity cost of forested land and the poverty rate to be positively associated with forest clearing: the former because it provides a measure of conversion profitability, and the latter because it proxies the local availability of low-cost, low-skill labor for forest clearing. We acknowledge some ambiguity in the latter expectation, since some kinds of agricultural production on cleared land will re- quire labor of the same type. We expect mobile phone coverage to be positively associated with forest clearing, because greater coverage lowers investor costs. Unit transport cost should be negatively associated with clearing, since palm oil and sawlogs are bulk commodities.[17] In the same vein, we would expect travel time to the nearest port to be negatively associated with forest

[17] A possible caveat is introduced by the impact of transport cost on local forest regulation, which is less effective in remote areas. If this factor dominates in Indonesia, the composite sign on transport cost could actually be positive.

Table 4.1: Definitions of estimation variables

Variable	Definition
`Clear`	FORMA forest clearing index
`PalmPrice`	Constant-dollar palm oil futures price
`LogPrice`	Constant-dollar sawlog price
`PalmQuant`	World palm oil production
`LogQuant`	World sawlog production
`OppCost`	Agricultural opportunity cost
`PovRate`	Poverty rate
`MobileCov`	Coverage by mobile phone networks
`MeanTravT`	Mean travel time to the nearest city of 50,000+
`sdTravT`	St. dev. travel time to the nearest city of 50,000+
`IntRate`	Real interest rate
`XRate`	Rupiah/dollar exchange rate
`InvGovQual`	KPPOD index of governance quality for investors
`RegQual`	KPPOD index of regulatory quality
`LogPct`	% of area zoned for logging concessions as of 2005
`TimbPct`	% of area zoned for timber plantation concessions as of 2005
`PalmPct`	% of area zoned for palm oil plantation concessions as of 2005
`ProtPct`	% of area zoned for protection of natural forest
`PopDens`	Population density
`Precip`	Precipitation
`MeanSlope`	Mean slope
`sdSlope`	St. dev. slope
`Forest2000`	Uncleared natural forest area in 2000

clearing. Although our indicator is the best available, it measures travel time to the nearest city of significant size, rather than time to the nearest port. In light of this difference, we remain agnostic about the potential size and significance of the measured effect.

Equation 4.2 includes a measure of governance quality for investors. Our database includes three relevant variables from the KPPOD survey: Quality of Assistance with Land Access; Capacity and Integrity of the Government; and Security and Conflict Resolution. A higher score on each variable should indicate a better environment for investment in forest-sector production. We would therefore expect a positive association between each variable and forest clearing. The other governance measure in Equation 4.2, Regulatory Quality, may be negatively associated with forest clearing in local protected areas. This will occur if local governments treat forest protection as a regulatory issue on par with other forms of local regulation.

We incorporate five types of land use, measured as percents of total area in each kabupaten. We include four types in the regression, excluding areas that are not explicitly zoned for protection or commercial exploitation.[18] *A priori*, we would expect areas zoned for commercial production to have more forest clearing than protected areas.[19]

Our model includes population density, which is related to local settlement and demand for forest-sector products. We would expect this variable to be positively associated with forest clearing. Finally, our specification includes two local physical factors: monthly precipitation and terrain slope. Forest clearing is more costly when precipitation hinders burning, so we would expect a strong negative association between the two variables. We expect a very short lag, if any, in the impact of precipitation on forest clearing. Suitability for plantation production declines with terrain slope, so we would expect a negative association with this variable as well. Our kabupaten-level measure of average terrain slope is calculated from highly-disaggregated spatial data, and kabupatens with the same average slope may have very different patterns of variation around the average. We capture this variation with the standard deviation, which we expect to moderate the measured effect of slope. If the marginal effect of mean slope is negative, as we expect, then we would expect the marginal effect of the standard deviation to be positive. In a similar vein, we have included the standard deviation of travel time along with our measure of kabupaten mean access time.

4.8 Results

In this section, we present our estimation results for Indonesian islands that are significant forest clearing sites. We exclude Java and Bali because they are heavily populated and largely cleared, and the islands of Nusa Tenggara that are not in the tropical forest zone.

Core model estimates: major geographic divisions

Table 4.2 reports results for a core model that includes the variables in Equation 4.2 that can be used for panel estimation by fixed effects. The first two columns of Table 4.2 present fixed and random effects estimates for all kabupatens in the tropical forest areas of Indonesia. Random effects estimation is preferable because it is more efficient, but its use depends on failure of the appropriate Hausman test to reject the null hypothesis of equal parameters in random and fixed effects estimation. Failure occurs in this case ($\chi^2 = 2.74$, $p = 0.9494$), so we adopt the random effects estimator. We retain this estimator for the remainder of the work reported in the paper.

As the strong Hausman results indicate, columns (1) and (2) have effectively-identical parameter estimates. In both equations, all estimated parameters have the expected signs and high levels of significance. Rainfall affects forest clearing with a short lag (one month

[18] We exclude one land use type to prevent perfect collinearity with the regression constant.

[19] Indonesian areas identified as "protected" may vary substantially in the actual degree of protection, lending some uncertainty to the assessed effectiveness of protection. We have no information on the relative allocation of monitoring and enforcement resources to different protected areas.

provides the best fit). Our results indicate that both expected prices and demands have strong effects on forest clearing, with substantially higher elasticities for the quantity effects. Our final estimates use the lags that provide the best fit to the data. As expected, we find no lag for the palm oil futures price, since it already incorporates expectations. Our best result for the sawlog price suggests that a lag of about nine months characterizes the process of price expectation revision and translation of revised expectations into forest clearing. The results for palm oil and sawlog demands suggest lags of 15 and 12 months, respectively, before market changes induce changes in forest clearing. Changes in the real interest rate take significantly longer to induce changes in forest clearing: 23 months in our best estimate. The response to changes in the real exchange rate is faster. We find approximately equal effects for lags of 9, 10 and 11 months, so we use the three-month average. We lag mobile phone coverage by one year to guard against simultaneity (a contemporaneous effect could easily reflect two-way causation). As Table 4.2 shows, the lagged variable is highly significant.

Columns (3)-(7) report random effects estimates for five islands and island groups: Sumatra, Kalimantan, Sulawesi, Maluku and Irian Jaya. Sample sizes vary from 3969 observations for Sumatra to 294 for Maluku. In light of this variation, it would not be surprising to see substantial variation in estimation results. However, the quality of the results is quite similar. The estimated impact of rainfall has the expected sign in all cases and statistical significance in three cases. Product price elasticities have the expected signs in eight of ten cases and are statistically significant in six cases. Quantity effects have the expected signs in all cases, and statistical significance in seven of ten cases. The real interest rate has the expected sign in all five cases, although it has statistical significance in only two (Kalimantan and Sulawesi). The exchange rate is statistically significant in three of five cases and has the expected sign in all cases.

Estimation of the fully specified model

Consistent, efficient panel estimation techniques permit inclusion of cross-sectional variables that may influence average levels of deforestation in Indonesian kabupatens. We have included several of these variables in the full specification in Equation 4.2. Table 4.3 reports results obtained from the random effects estimator programmed in `Stata` (column (1)), as well as three alternative estimators programmed in `R` (columns (2)-(4)). We include the latter because appropriate adjustments for spatial dependence were not available in `Stata`. Column (2) reports estimation results equivalent to those in (1), obtained using the method of Swamy and Arora (1972) for a fully-balanced panel.[104][20] Column (3) reports estimates from the method of Kapoor *et al.* (2007), which adjusts for spatial autocorrelation across kabupatens.[70] Column (4) reports estimates from the method of Millo and Piras (2009),

[20] Implementation in `R` requires a fully-balanced panel, which we produce by spatial interpolation of available monthly rainfall observations to replace missing observations in some kabupatens. The result is an enlarged panel (193 kabupatens) relative to the dataset used for our `Stata` estimate in column (1) (142 kabupatens).

which adjust for both spatial autocorrelation and spatial lags across kabupatens.[70, 81] We provide a brief introduction to the three estimators in the supplement Section 4.10.

Table 4.2: Economic dynamics and forest clearing. All variables in logs except month and real interest rate. Lags (months) in brackets. Dependent variable: log(`Clear`).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	All	All	Sumatra	Kalimantan	Sulawesi	Maluku	Irian Jaya
Estimator	FE	RE	RE	RE	RE	RE	RE
Rainfall [-1]	0.381	0.379	0.557	0.308	0.085	0.407	0.152
	(10.19)**	(10.14)**	(9.42)**	(3.92)**	(1.13)	(2.62)**	(0.92)
Palm oil futures price	0.816	0.815	0.385	1.716	0.898	0.270	0.688
	(10.71)**	(10.70)**	(3.40)**	(11.22)**	(4.53)**	(0.59)	(3.51)**
Sawlog price [-9]	1.134	1.129	1.562	2.438	1.316	2.809	1.209
	(2.97)**	(2.96)**	(2.76)**	(3.12)**	(1.32)	(1.23)	(1.24)
Global palm oil prod. [-15]	5.225	5.231	6.346	5.808	3.959	13.842	0.740
	(6.53)**	(6.53)**	(5.35)**	(3.60)**	(1.90)	(2.89)**	(0.36)
Global sawlog prod. [-12]	14.651	14.594	12.376	22.205	10.448	28.808	7.541
	(9.27)**	(9.24)**	(5.12)**	(7.04)**	(2.56)*	(3.11)**	(1.89)
Real interest rate [-23]	0.044	0.044	0.018	0.105	0.052	0.046	0.031
	(7.03)**	(7.07)**	(1.96)	(8.37)**	(3.19)**	(1.24)	(1.93)
Exchange rate [-9,10,11]	2.827	2.823	1.246	3.881	4.961	9.025	2.813
	(4.94)**	(4.93)**	(1.47)	(3.33)**	(3.30)**	(2.63)**	(1.93)
Mobile phone coverage [-12]	0.117	0.112	0.095	0.115	0.511	4.106	0.008
	(7.09)**	(6.98)**	(4.77)**	(0.96)	(2.81)**	(3.15)**	(0.06)
Constant	396.0	394.9	344.6	579.7	307.8	854.4	189.8
	(9.15)**	(9.13)**	(5.26)**	(6.69)**	(2.74)**	(3.34)**	(1.73)
Observations	8673	8673	3969	1960	1225	294	1225
Kabupatens	177	177	81	40	25	6	25

Hausman Test for Fixed Effects (1) vs. Random Effects (2): $\chi^2 = 2.74$ ($p = 0.9494$).
Absolute value of t statistics in parentheses; Significance: **0.01; *0.05.

Inspection of Table 4.3 reveals remarkable similarity in the results obtained by all of our estimators: Signs and general significance levels are identical for 81 estimates and different for only three. We assign particular importance to column (4), which reports high significance for both spatial autocorrelation (ρ) and spatial lags (λ), and adjusts for both error components. All variables in the KPPOD governance survey are insignificant in every case, so we have excluded them from the table.

In Table 4.3, our results for core-model time series variables are very similar to those reported in Table 4.2. Signs and elevated significance levels match throughout, as do estimated parameter sizes (except for the rainfall results in the R-based estimates). Table 4.3 includes the cross-section variables in Equation (4.2): uncleared forest in 2000; terrain slope (mean and SD); land opportunity cost, population density; the poverty rate; and access time to the nearest city (mean and SD). We also include information on zoning status: the percent of area in each kabupaten zoned for forest protection, as well as commercial zoning in three categories: palm oil plantations, timber plantations and logging concessions. Among the cross-sectional variables, only three are highly significant and have the expected signs in all four estimates: uncleared forest in 2000; mean slope (modified by its standard deviation[21]); and percent area designated for palm oil plantations. Zoning for logging concessions is also consistently signed, with a high significance level in two of four cases. The large, positive, significant results for palm oil plantations and the insignificance of protected-area status are notable in all four estimates.

For the comparable variables, our results differ somewhat from those obtained by a recent cross-section estimation exercise for Indonesia. Using pixel-level forest clearing estimates at 500m resolution from Hansen *et al.* (2008), Busch *et al.* (2012) find appropriately-signed, significant impacts on deforestation for a measure of land opportunity cost based on estimated net present potential gross agricultural revenue (Naidoo and Iwamura, 2007), as well as variables related to land use, topography and transport costs.[58, 19, 85] We obtain similar results for one topographical variable (slope) and one land-use variable (logging concession percent). However, as we have noted above, we do not find significance for our measures of direct land opportunity cost, protection, and transport cost.

There are several plausible reasons for the differences in results. First, the two studies cover different time periods and time increments: the entire period from 2000 to 2005 for Busch *et al.* (2012), versus monthly data from December, 2005 to August, 2011 in our own case.[19] Second, the Busch study is purely cross-sectional, while ours is a panel estimation exercise that incorporates time-varying national-level variables, time-varying local variables (rainfall and cell phone coverage), and time-invariant local variables. Direct comparison is possible only for the time-invariant local variables. Third, the Busch exercise is at the pixel level, while ours aggregates to kabupaten averages. Their pixel-level database provides huge nominal degrees of freedom ($N = 166,297$), which should permit estimation of underlying parameter values with considerably higher precision than our panel exercise at the kabu-

[21]The result for the standard deviation (SD) indicates that slope effects are more pronounced in areas with relatively small variations in slope.

paten level. However, a major caveat is introduced by their failure to correct for spatial autocorrelation, which is likely to be very high at the pixel level of spatial disaggregation.[22] Absent the needed correction, we cannot draw strong inferences from the estimated parameters, signs and significance levels in their exercise. Fourth, our introduction of variables measured at the kabupaten level (e.g., the local poverty index) forces us to use area averages for the geographic variables in Table 4.2 that are observed at much higher resolution. This spatial averaging seems likely to attenuate the significance of geographic variables that are used at much higher resolution in the pixel-level study of Busch *et al.* (2012).[19] Fifth, our two significant time-varying local variablesrainfall and cell phone coverageare not incorporated in the Busch study. Either could be correlated with variables included in that study, and significant correlation seems particularly likely for the relationship between cell phone coverage in our case and proximity to transport infrastructure in the Busch study. If so, then our results for cross-sectional factors may be closer to the Busch results than they appear to be. Sixth, we use a more recently-derived land opportunity cost index, measured at a higher level of spatial resolution that incorporates more determinants than the Naidoo-Iwamura measure employed by Busch.

For all these reasons, we remain uncertain about the actual degree of difference between our cross-section results and those of the Busch study. In this context, it is worth noting that incorporation of the needed spatial autocorrelation adjustment in such pixel-level exercises can easily overwhelm existing estimation packages. It may therefore be some time before more definitive results become available from pixel-level research.

4.9 Conclusion

In this paper, we have employed a large panel database to investigate the determinants of forest clearing in Indonesian kabupatens since 2005. Using monthly forest clearing data from FORMA (Forest Monitoring for Action), the paper provides the first Indonesian impact assessment for short-run economic variables, as well as impact estimators for indicators of area zoning, forest protection, the opportunity cost of forested land, the availability of communications infrastructure, and the quality of local governance. In addition, we test the effects of variables that have been included in more traditional analyses of forest clearing: rainfall, terrain characteristics, the poverty rate, population density and transport cost.

Our results strikingly demonstrate the importance of economic factors in the dynamics of forest clearing. In our full estimation model, significant roles are played by short-run changes in several economic variables, as well as communications infrastructure, zoning for palm oil plantations, and three physical factorsuncleared forest in 2000, rainfall and terrain slope. In counterpoint, many cross-section variables prove to be insignificant: local governance quality, a direct estimate of land opportunity cost, travel time, population density, the poverty rate,

[22]Indeed, it remains high even at the much greater aggregation level we use for our sub-provincial panel exercise.

protected-area status, and zoning for timber plantations. In the case of access time (our proxy for transport cost), we recognize that the insignificance of the measured effect could be attributable to collinearity with our proxy for communications cost, and/or the offsetting impacts of transport cost on the direct profitability of forest clearing (negative) and the effectiveness of local forest regulation (positive).

We believe that the most distinctive feature of our approach is its inclusion of short-run economic variables, which was simply not possible before the advent of FORMA. As we have noted in the paper, economic theory has long posited critical roles for expected forest product prices, quantity demands, interest rates and exchange rates in the investor calculations that lead to large-scale clearing for commercial production. The econometric analysis reported in this paper introduces all of these variables and explores the time lags that characterize their impact on forest clearing. We find highly-variable lags: less than a year for product prices; around one year for product demands and the exchange rate; and closer to two years for the real interest rate. All variables are highly significant in our panel analysis, and their fluctuations, along with variations in rainfall, explain a major portion of the changes in Indonesian forest clearing that are strikingly visible in Figure (4.1).

From a policy perspective, our results highlight the importance of incorporating economic dynamics into arrangements that offer financial compensation for forest conservation. Our findings are strongly consistent with a model of forest clearing as an investment that is highly sensitive to expectations about future forest product prices and demands, as well as changes in the cost of capital (indexed by the real interest rate), the relative cost of local inputs (indexed by the exchange rate), and the cost of land clearing (indexed by local precipitation). By implication, the opportunity cost of forested land fluctuates widely as changes occur in international markets, local weather conditions, and decisions by Indonesia's financial authorities about the exchange and interest rates.

Among the land-use variables available for our analysis, we view the protected area indicator with particular caution. In practice, protected areas are strongly differentiated by sponsor identity (local government, national government, local or national NGO), local community engagement, enforcement priority, administrative arrangements and other factors. It would certainly be misleading to claim that the insignificant result for our general protection variable implies that none of the variants identified above can have a significant effect. To cite one recent counterexample, Nelson and Chomitz (2011) find significant effects for multiple use and indigenous areas in a global study that controls for a set of cross-sectional variables similar to those employed in this study and Busch *et al.* (2012).[87, 19] With the advent of FORMA and other high-resolution panel datasets, future research will be able to explore the relationship between protected-area characteristics and forest clearing in much more detail.

Our results may also provide new insights for compensation-based approaches, since they suggest that the perceived opportunity cost of forested land varies widely over time, and in response to numerous dynamic factors.[23] By implication, compensation schemes for forest

[23] Here it is important to distinguish our use of the general term opportunity cost from the single measure of

conservation may have to incorporate arrangements for adjusting compensation as economic conditions change. One possible approach, for example, could employ a weighted index of the economic factors that affect clearing decisions.[24] We propose another approach in Hammer *et al.* (2011): A "cash-on-delivery" system that ties annual payments to changes in national forest conservation performance over time. Our proposed system focuses solely on incentive payments to national governments, leaving them free to make flexible arrangements with local forest proprietors.[54]

In summary, our empirical results and other recent research suggest that forest protection programs will be more likely to succeed if they recognize the importance of several local, national and international factors. Locally, appropriate compensation for forest conservation will need to recognize the effects of topography, variable rainfall and infrastructure improvements on the potential profit- ability of converting forested land to other uses. At the national level, effective compensation will need to recognize the importance of key macro-policy variables (e.g., interest and exchange rates set by national authorities) in investor calculations. And international market variables will undoubtedly be important as well, as changes in expected prices and demands for forest products affect the expected profitability of forest clearing.

4.10 Supplement: estimation of spatial models

Columns (2)(4) of Table 4.3 report results from a succession of panel estimators implemented in `R`. Estimation requires a complete, balanced panel, so we have used spatial interpolation to replace some missing monthly rainfall observations in our kabupaten-level dataset. The estimation panel includes five years of monthly data (Jan. 2006 - Dec. 2010) for 193 kabupatens, yielding 11,580 observations. We construct a weighting matrix from the GIS shapefile for the 193 kabupatens, defining neighbors to be kabupatens with borders that are within 0.5 arc degrees (roughly 50 km) of each other. The weighting matrix $\mathbf{W}$ is row-standardized, so that each row sums to one.

Our objective is to efficiently and consistently estimate the following model:

$$\mathbf{y} = \alpha + \mathbf{X}\beta + \mathbf{u} \tag{4.3}$$

where $\mathbf{y}$ is a panel of logs of forest clearing activity (by kabupaten and month), $\mathbf{X}$ is a panel of explanatory variables and $\mathbf{u}$ is an error term. OLS estimates of model parameters are

---

agricultural opportunity cost used for the econometric analysis. The latter provides one temporal snapshot of the value of forested land in alternative uses. The former includes all the intertemporal factors (e.g. expected prices and demands, the exchange rate, the interest rate, rainfall, and cell phone coverage) that have time-variable effects on the profitability of clearing, and therefore on the opportunity cost of keeping forested land uncleared.

[24] We are indebted to an anonymous referee for this suggestion.

not efficient or consistent if $\mathbf{u}$ is subject to serial or spatial autocorrelation. Standard panel estimation employs the model:

$$\mathbf{y} = \alpha + \mathbf{X}\beta + \mu + \mathbf{E} \tag{4.4}$$

where the error component μ is specific to panel groups and $\mathbf{E}$ is assumed to be uncorrelated with μ and the regressors in $\mathbf{X}$. Estimation of Equation (4.4 via GLS will produce efficient results for non-autocorrelated data. We report estimates obtained by the method of Swamy and Arora (1972) in column (2) of Table 4.2.[104]

For forest clearing analysis, standard GLS estimation may be insufficient because of spatial dependence. As Figures (4.2) and (4.3) indicate, forest clearing clusters frequently cross kabupaten boundaries. Kapoor *et al.* (2007) propose the following model for the case where error terms are correlated across spatial units, where the error term $\mathbf{u}$ in Equation (4.3) follows a first-order spatial autoreggresive process:

$$\mathbf{u} = \rho\left(\mathbf{I}_T \otimes \mathbf{W}\right)\mathbf{u} + \mathbf{E} \tag{4.5}$$

and observations are stacked by time period rather than panel group.[70] This model can be interpreted as a time series of kabupaten cross sections, and ρ is the coefficient of spatial spillover of the errors. To allow for temporal autocorrelation, $\mathbf{E}$ is specified as:

$$\mathbf{E} = \left(\mathbf{e}_T \otimes \mathbf{I}_N\right)\mu + \nu \tag{4.6}$$

where μ is a vector of kabupaten-specific error components, $\mathbf{e}_T$ is an appropriately-dimensioned unit vector, and ν contains idiosyncratic error components that vary over time and space. We report the results of estimation by this model in column (3) of Table 4.2.

Dynamic economic factors propel growing forest clearing clusters across kabupaten boundaries. It is therefore likely that our panel data are also characterized by spatial lags, in which clearing in one kabupaten is related to clearing in neighboring kabupatens. Following Millo and Piras (2009), we specify and estimate a model with general spatial autocorrelation:

$$\mathbf{y} = \lambda\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{u} = \rho\mathbf{W}\mathbf{u} + \eta \tag{4.7}$$

where $\eta \sim N(0, \Omega)$, $\Omega \neq \sigma^2\mathbf{I}$.[81] This specification incorporates both a pure spatial error model when $\lambda = 0$ and a pure spatial lag model when $\rho = 0$. We report our results in column (4) of Table 4.2.

Table 4.3: Introduction of cross-sectional variables. Variables logs except interest rate, area %s. Lags (months) in brackets. Dependent variable: log(`Clear`)

Regression variables	**(1)**	**(2)**	**(3)**	**(4)**
Rainfall [-1]	0.366	0.004	0.003	0.002
	(9.08)	(6.842)	(5.242)	(4.993)
Palm oil futures price	0.904	0.784	0.782	0.356
	(10.67)	(11.347)	(8.520)	(7.669)
Sawlog price [-9]	1.342	1.354	1.343	0.593
	(3.15)	(3.984)	(2.981)	(2.595)
Global palm oil production [-15]	5.811	4.636	4.458	2.000
	(6.53)	(6.398)	(4.636)	(4.109)
Real interest rate [-23]	0.054	0.048	0.046	0.020
	(7.75)	(8.407)	(6.088)	(5.201)
Exchange rate [-9,10,11]	2.544	2.675	2.605	1.190
	(3.99)	(5.225)	(3.836)	(3.458)
Mobile phone coverage [-12]	0.073	0.062	0.083	0.045
	(4.37)	(4.174)	(5.092)	(3.772)
Mean slope	2.783	2.459	2.513	1.674
	(5.98)	(6.497)	(6.340)	(5.124)
St. dev. slope	1.535	0.965	1.090	0.929
	(3.46)	(2.879)	(3.251)	(3.028)
Land opportunity cost	0.133	0.061	0.058	0.097
	(0.65)	(0.417)	(0.376)	(0.786)
Protected area %	0.117	0.398	0.390	0.147
	(0.10)	(0.380)	(0.378)	(0.153)
Timber plantation area %	0.696	0.033	0.253	2.266
	(0.36)	(0.018)	(0.139)	(1.275)
Logging concession area %	4.292	1.920	1.288	3.156
	(2.71)	(1.402)	(0.951)	(2.483)
Palm oil plantation %	16.361	16.438	15.056	13.787
	(4.95)	(5.234)	(4.734)	(4.871)
Population density	0.128	0.067	0.158	0.071
	(0.46)	(0.299)	(0.706)	(0.342)
Poverty rate (2000)	1.761	1.183	1.137	0.302
	(0.80)	(0.652)	(0.591)	(0.199)
Access time to nearest city (50,000+)	1.543	0.008	0.102	0.046
	(2.46)	(0.020)	(0.235)	(0.118)
St. dev. access time	1.081	0.016	0.086	0.195
	(1.83)	(0.038)	(0.209)	(0.512)
Constant	417.412	337.874	331.540	152.641
	(8.61)	(8.882)	(6.590)	(5.944)
ρ			.257	.411
				(9.41)
λ				.538
				(20.933)
Adj. R^2	.556	.130	.827	.405
Kabupatens	142	193	193	193

Bibliography

[1] Deepak K Agarwal, Alan E Gelfand, and John A Silander. "Investigating tropical deforestation using two-stage spatially misaligned regression models". In: *Journal of agricultural, biological, and environmental statistics* 7.3 (2002), pp. 420–439.

[2] Deepak K Agarwal et al. "Tropical deforestation in Madagascar: analysis using hierarchical, spatially explicit, Bayesian regression models". In: *Ecological Modelling* 185.1 (2005), pp. 105–131.

[3] L. Anderson et al. "Assessment of Deforestation in Near Real Time Over the Brazilian Amazon Using Multitemporal Fraction Images Derived From Terra MODIS". In: *IEEE Geoscience and Remote Sensing Letters* 2.3 (2005), pp. 315–318.

[4] Luc Anselin. "Spatial effects in econometric practice in environmental and resource economics". In: *American Journal of Agricultural Economics* 83.3 (2001), pp. 705–710.

[5] Luc Anselin. "Under the hood issues in the specification and interpretation of spatial regression models". In: *Agricultural economics* 27.3 (2002), pp. 247–267.

[6] Eillie Anzilotti. "Every Day, This Satellite Company Takes A Snapshot Of The Entire Planet". In: *Fast Company* (Nov. 2017).

[7] Jean-Louis Arcand, Patrick Guillaumont, and Sylviane Guillaumont Jeanneney. "Deforestation and the real exchange rate". In: *Journal of Development Economics* 86.2 (June 2008), pp. 242–262.

[8] Gregory P Asner et al. "High-resolution forest carbon stocks and emissions in the Amazon". In: *Proceedings of the National Academy of Sciences* 107.38 (Sept. 2010), pp. 16738–16742.

[9] Kemen Austin, Fred Stolle, and Stuart Sheppard. "Indonesia's Moratorium on New Forest Concessions:Key Findings and Next Steps". In: *World Resources Institute Working Paper* (Feb. 2012).

[10] "Back with a vengeance: Rising commodity prices both reflect and threaten the world's economic recovery". In: *The Economist* (Jan. 2011).

[11] Edward B Barbier and Joanne C Burgess. "Economic analysis of deforestation in Mexico". In: *Environment and Development Economics* 1.2 (1996), pp. 203–239.

[12] Thomas Barrios et al. "Clustering, spatial correlations, and randomization inference". In: *Journal of the American Statistical Association* 107.498 (2012), pp. 578–591.

[13] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. "How Much Should We Trust Differences-In-Differences Estimates?" In: *The Quarterly Journal of Economics* 119.1 (2004), pp. 249–275.

[14] Roberto Bonifaz-Alfonzo. *Assessing Seasonal Features of Tropical Forests Using Remote Sensing*. Tech. rep. University of Nebraska-Lincoln, Apr. 2011.

[15] Mark Broich et al. "Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia". In: *International Journal of Applied Earth Observations and Geoinformation* 13.2 (Apr. 2011), pp. 277–291.

[16] Lori Mann Bruce, Abhinav Mathur, and John D Byrd Jr. "Denoising and Wavelet-Based Feature Extraction of MODIS Multi-Temporal Vegetation Signatures". In: *GIScience & Remote Sensing* 43.1 (Sept. 2007), pp. 67–77.

[17] Maria Burns. "Remote Sensing Services Market Worth 21.62 Billion USD by 2022". In: *The Daily Telescope* (Nov. 2017).

[18] Jonah Busch et al. "Reductions in emissions from deforestation from Indonesia's moratorium on new oil palm, timber, and logging concessions". In: *Proceedings of the National Academy of Sciences* 112.5 (2015), pp. 1328–1333.

[19] Jonah Busch et al. "Structuring economic incentives to reduce emissions from deforestation within Indonesia". In: *Proceedings of the National Academy of Sciences* 109.4 (2012), pp. 1062–1067.

[20] Gilberto Camara, Dalton de Morisson Valeriano, and Joao Vianei Soares. *Metodologia para o Calculo da Taxa Anual de Desmatamento na Amazonia Legal*. Jan. 2006. URL: `http://www.obt.inpe.br/prodes/metodologia.pdf`.

[21] Alfredo Nicolas Campos and Carlos Marcelo Di Bella. "Multi-Temporal Analysis of Remotely Sensed Information Using Wavelets". In: *Journal of Geographic Information System* 4.4 (Aug. 2012), pp. 383–391.

[22] Kimberly M Carlson et al. "Carbon emissions from forest conversion by Kalimantan oil palm plantations". In: *Nature Climate Change* 3.3 (2013), pp. 283–287.

[23] Andrea Cattaneo, Edward B Barbier, and Joanne C Burgess. "Deforestation in the Brazilian Amazon: comparing the impacts of macroeconomic shocks, land tenure, and technological change". In: *Land Economics* 77.2 (2001), pp. 219–240.

[24] S Le Cessie and J C Van Houwelingen. "Ridge Estimators in Logistic Regression". In: *Applied Statistics* 41.1 (1992), pp. 191–201.

[25] Mingyue Chen et al. "Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations". In: *Journal of Hydrometeorology* 3 (May 2002), pp. 249–266.

[26] Kenneth M Chomitz. “At Loggerheads?” In: (2007).

[27] Rinku Roy Chowdhury. “Driving forces of tropical deforestation: The role of remote sensing and spatial models”. In: *Singapore Journal of Tropical Geography* 27.1 (2006), pp. 82–101.

[28] Changbum Chun. “Construction of Newton-like iteration methods for solving nonlinear equations”. In: *Numerische Mathematik* 104.3 (Aug. 2006), pp. 297–315.

[29] Matthew L Clark et al. “A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America”. In: *Remote Sensing of Environment* 114.11 (Nov. 2010), pp. 2816–2832.

[30] William S Cleveland and Susan J Devlin. “Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting”. In: *Journal of the American Statistical Association* 83.403 (Sept. 1988), pp. 596–610.

[31] *Climate Analysis Indicators Tool (CAIT)*. Tech. rep. WRI, 2010. URL: `cait.wri.org`.

[32] Union of Concerned Scientists. *UCS Satellite Database*. Nov. 2017.

[33] P Coppin et al. “Digital change detection methods in ecosystem monitoring: a review”. In: *International Journal of Remote Sensing* 25.9 (May 2004), pp. 1565–1596.

[34] Maureen Cropper, Charles Griffiths, and Muthukumara Mani. “Roads, population pressures, and deforestation in Thailand, 1976-1989”. In: *Land Economics* (1999), pp. 58–73.

[35] Maureen Cropper et al. “Predicting the location of deforestation: The role of roads and protected areas in North Thailand”. In: *Land Economics* 77.2 (2001), pp. 172–186.

[36] D K Davies et al. “Fire Information for Resource Management System: Archiving and Distributing MODIS Active Fire Data”. In: *IEEE Transactions on Geoscience and Remote Sensing* 47.1 (Nov. 2008), pp. 72–79.

[37] Alessandro De Pinto and Gerald C Nelson. “Land use change with spatially explicit data: a dynamic approach”. In: *Environmental and Resource Economics* 43.2 (2009), pp. 209–229.

[38] R DeFries et al. *Reducing Greenhouse Gas Emissions from Deforestation in Developing Countries: Considerations for Monitoring and Measuring*. Tech. rep. 26. GOFC-GOLD, Aug. 2006.

[39] Klaus W Deininger and Bart Minten. “Poverty, policies, and deforestation: the case of Mexico”. In: *Economic Development and Cultural Change* 47.2 (1999), pp. 313–344.

[40] C.M. DiMiceli et al. “Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 - 2010”. In: *University of Maryland* (2011).

[41] P Enkvist, Tomas Naucler, and Jerker Rosander. “A cost curve for greenhouse gas reduction”. In: *McKinsey Quarterly* 1 (2007), p. 34.

[42] Stefan Erasmi et al. “Spatial Patterns of NDVI Variation over Indonesia and Their Relationship to ENSO Warm Events during the Period 1982–2006”. In: *Journal of Climate* 22.24 (Dec. 2009), pp. 6612–6623.

[43] Robert M Ewers, William F Laurance, and Carlos M Souza. “Temporal fluctuations in Amazonian deforestation rates”. In: *Environmental Conservation* 35.4 (2008), pp. 303–310.

[44] FAO. *Global forest land-use change 1990–2005*. Tech. rep. 169. Food and Agriculture Organization of the United Nations, Jan. 2012.

[45] Carl Folke et al. “Regime Shifts, Resilience, and Biodiversity in Ecosystem Management”. In: *Annual Review of Ecology, Evolution, and Systematics* 35.1 (Dec. 2004), pp. 557–581.

[46] David LA Gaveau et al. “Evaluating whether protected areas reduce tropical deforestation in Sumatra”. In: *Journal of biogeography* 36.11 (2009), pp. 2165–2175.

[47] H. K. Gibbs et al. “Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s”. In: *Proceedings of the National Academy of Sciences* 107.38 (2010), pp. 16732–16737.

[48] Anatoly A Gitelson, Yoram J Kaufman, and Mark N Merzlyak. “Use of a green channel in remote sensing of global vegetation from EOS-MODIS”. In: *Remote Sensing of Environment* 58.3 (Dec. 1996), pp. 289–298.

[49] Scott J Goetz et al. “Mapping and monitoring carbon stocks with satellite observations: a comparison of methods”. In: *Carbon Balance and Management* 4.1 (Mar. 2009).

[50] L Gomez-Mendoza et al. “Assessing onset and length of greening period in six vegetation types in Oaxaca, Mexico, using NDVI-precipitation relationships”. In: *International Journal of Biometeorology* 52.6 (Feb. 2008), pp. 511–520. URL: `http://dx.doi.org/10.1007/s00484-008-0147-6`.

[51] Alan Grainger. “Difficulties in tracking the long-term global trend in tropical forest area”. In: *Proceedings of the National Academy of Sciences* 105.2 (2008), pp. 818–823.

[52] Adam Gunderson and Mark Chodas. “An investigation of cloud cover probability for the HyspIRI mission using MODIS cloud mask data”. In: *2011 IEEE Aerospace Conference*. IEEE, Mar. 2011, pp. 1–14.

[53] D. Hammer, R. Kraft, and D. Wheeler. “FORMA: Forest Monitoring for Action — Rapid Identification of Pan-tropical Deforestation Using Moderate-Resolution Remotely Sensed Data”. In: *CGD Working Paper* 192 (2009).

[54] Dan Hammer, Robin Kraft, and David Wheeler. "Alerts of forest disturbance from {MODIS} imagery". In: *International Journal of Applied Earth Observation and Geoinformation* 33.0 (2014), pp. 1–9.

[55] Bruce E Hansen. "Tests for Parameter Instability in Regressions with I (1) Processes". In: *Journal of Business & Economic Statistics* 20.1 (2002), pp. 45–59.

[56] M C Hansen et al. "Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm". In: *Earth Interactions* 7 (Sept. 2003), pp. 1–15.

[57] M C Hansen et al. "High-Resolution Global Maps of 21st-Century Forest Cover Change". In: *Science* 342.6160 (Nov. 2013), pp. 850–853.

[58] M. Hansen et al. "A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin". In: *Remote Sensing of Environment* 112.5 (2008), pp. 2495–2513.

[59] M. Hansen et al. "Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data". In: *Proceedings of the National Academy of Sciences* 105 (2008), pp. 9439–9444.

[60] N L Harris et al. "Baseline Map of Carbon Emissions from Deforestation in Tropical Regions". In: *Science* 336.6088 (June 2012), pp. 1573–1576.

[61] Robert Hijmans. *Global Administrative Database, version 2*. 2013. URL: `http://www.gadm.org`.

[62] R. A. Houghton, Brett Byers, and Alexander A. Nassikas. "A role for tropical forests in stabilizing atmospheric CO2". In: *Nature Climate Change* 5 (Nov. 2015), pp. 1022–1023.

[63] Chengquan Huang et al. "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks". In: *Remote Sensing of Environment* 114.1 (Jan. 2010), pp. 183–198.

[64] Danilo Igliori. "Deforestation, Growth and Agglomeration Effects: Evidence from Agriculture in the Brazilian Amazon". In: (2006).

[65] World Resources Institute. *forma-clj*. `www.github.com/wri/forma-clj`. 2014.

[66] *PRODES data server*. URL: `http://www.dpi.inpe.br/prodesdigital/dadosn`.

[67] "Joint Statement by Germany, Norway and the United Kingdom of Great Britain and Northern Ireland: Unlocking the Potential of Forests and Land Use". In: (Nov. 2015).

[68] Christopher. Justice et al. "MODIS-Derived Global Fire Products". English. In: *Land Remote Sensing and Global Environmental Change*. Ed. by Bhaskar Ramachandran, Christopher O. Justice, and Michael J. Abrams. Vol. 11. Remote Sensing and Digital Image Processing. Springer New York, 2011, pp. 661–679. ISBN: 978-1-4419-6748-0.

[69] David Kaimowitz et al. "Spatial regression analysis of deforestation in Santa Cruz, Bolivia". In: (2002).

[70] Mudit Kapoor, Harry H Kelejian, and Ingmar R Prucha. "Panel data models with spatially correlated error components". In: *Journal of econometrics* 140.1 (2007), pp. 97–130.

[71] Harry H Kelejian and Ingmar R Prucha. "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances". In: *The Journal of Real Estate Finance and Economics* 17.1 (1998), pp. 99–121.

[72] Harry H Kelejian and Ingmar R Prucha. "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances". In: *Journal of Econometrics* 157.1 (2010), pp. 53–67.

[73] Harry H Kelejian, Ingmar R Prucha, and Yevgeny Yuzefovich. "Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results". In: *Spatial and spatiotemporal econometrics*. Emerald Group Publishing Limited, 2004, pp. 163–198.

[74] Robert E Kennedy, Warren B Cohen, and Todd A Schroeder. "Trajectory-based change detection for automated characterization of forest disturbance dynamics". In: *Remote Sensing of Environment* 110.3 (Oct. 2007), pp. 370–386.

[75] W Kleynhans et al. "Detecting Land Cover Change Using an Extended Kalman Filter on MODIS NDVI Time Series Data". In: *Remote Sensing of Environment* (Oct. 2010), pp. 1–5.

[76] Y Li, M Mascagni, and A Gorin. "A decentralized parallel implementation for parallel tempering algorithm". In: *Parallel Computing* 35.5 (2009).

[77] Ross S Lunetta et al. "Land-cover change detection using multi-temporal MODIS NDVI data". In: *Remote Sensing of Environment* 105.2 (Nov. 2006), pp. 142–154.

[78] Belinda Arunarwati Margono et al. "Primary forest cover loss in Indonesia over 2000-2012". In: *Nature Climate Change* 4 (8 2014), pp. 730–735.

[79] T R Meagher. "A fragmented future for the forest flora". In: *Heredity* 105.2 (May 2010), pp. 163–164.

[80] David J Mildrexler, Maosheng Zhao, and Steven W Running. "Testing a MODIS Global Disturbance Index across North America". In: *Remote Sensing of Environment* 113.10 (Oct. 2009), pp. 2103–2117.

[81] Giovanni Millo and Gianfranco Piras. "splm: econometric analysis of spatial panel data". In:

[82] D C Morton et al. "Agricultural intensification increases deforestation fire activity in Amazonia". In: *Global Change Biology* 14.10 (2008), pp. 2262–2275.

[83] Mike Mulligan. *RALUCIAPA : Rapid assessment of land use change in and around protected areas (2000-2005).* Tech. rep. 2008. URL: `unep-wcmc.org`.

[84] D. Murdiyarso et al. “Indonesia’s forest moratorium: A stepping stone to better forest governance?” In: *CIFOR Working Paper* 3561 (2011).

[85] Robin Naidoo and Takuya Iwamura. “Global-scale mapping of economic benefits from agricultural lands: Implications for conservation priorities”. In: *Biological Conservation* 140.1 (2007), pp. 40–49. ISSN: 0006-3207. DOI: `https://doi.org/10.1016/j.biocon.2007.07.025`. URL: `http://www.sciencedirect.com/science/article/pii/S0006320707003011`.

[86] NASA Land Processes Distributed Active Archive Center. *MODIS/Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid*. USGS/Earth Resources Observation and Science (EROS) Center. Sioux Falls, South Dakota, 2001.

[87] Andrew Nelson and Kenneth M Chomitz. “Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods”. In: *PloS one* 6.8 (2011), e22722.

[88] David M Olson et al. “Terrestrial Ecoregions of the World: A New Map of Life on Earth”. In: *BioScience* 51.11 (2001), p. 933.

[89] Alexander SP Pfaff. “What Drives Deforestation in the Brazilian Amazon?” In: *Journal of Environmental Economics and Management* 37 (1999), p. 2643.

[90] A. C. Pigou. *The economics of welfare*. Macmillan, London, 1920.

[91] C Portillo-Quintero and A Sanchez-Azofeifa. “Monitoring deforestation with MODIS Active Fires in Neotropical dry forests: An analysis of local-scale assessments in Mexico, Brazil and Bolivia”. In: *Journal of Arid Environments* 97 (2013), pp. 150–159.

[92] Christoper Potter et al. *CASA Quarterly Indicator of Cover Change (QUICC)*. 2012.

[93] B C Reed and J F Brown. “Trend analysis of time-series phenology derived from satellite data”. In: *International Workshop on the Analysis of Multi-Temporal Remote Sensing Images, 2005.* IEEE, May 2005, pp. 166–168.

[94] Louis Reymondin et al. “A methodology for near real-time monitoring of habitat change at continental scales using MODIS-0NDVI and TRMM.” In: *Submitted* (2012).

[95] D. P. Roy et al. “The Global Impact of Clouds on the Production of MODIS Bidirectional Reflectance Model-Based Composites for Terrestrial Monitoring”. In: *IEEE Geoscience and Remote Sensing Letters* 3.4 (Oct. 2006), pp. 452–456.

[96] Thomas K Rudel et al. “Agricultural intensification and changes in cultivated areas, 1970–2005”. In: *Proceedings of the National Academy of Sciences* 106.49 (2009), pp. 20675–20680.

[97] Nu Nu San, Hans Lofgren, and Sherman Robinson. “Structural Adjustment, Agriculture, and Deforestation in the Sumatera Regional Economy”. In: (2000).

[98] P A Schultz and M S Halpert. “Global correlation of temperature, NDVI and precipitation”. In: *Advances in Space Research* 13.5 (May 1993), pp. 277–280.

[99] Yosio E Shimabukuro et al. “Near real time detection of deforestation in the Brazilian Amazon using MODIS imagery”. In: *Ambiente e Agua* 1 (Jan. 2006), pp. 37–47.

[100] Masato Shinoda. “Seasonal phase lag between rainfall and vegetation activity in tropical Africa as revealed by NOAA satellite data”. In: *International Journal of Climatology* 15.6 (June 1995), pp. 639–656.

[101] Carlos M Souza. “Mapping land use of tropical regions from space”. In: *Proceedings of the National Academy of Sciences* 103.39 (2006), pp. 14261–14262.

[102] Carlos M de Souza, Sanae Hayashi, and Adalberto Verissimo. “Near real-time deforestation detection for enforcement of forest reserves in Mato Grosso”. In: *Land Governance in Support of the MDGs: Responding to New Challenges*. 2009.

[103] Carlos Souza et al. “Normalized Difference Fraction Index (NDFI): a new spectral index for enhanced detection of forest canopy damage caused by selective logging and forest fires”. In: *LBA-ECO 9th Science Team Meeting*. Sao Paulo, Nov. 2005.

[104] PAVB Swamy and Swarnjit S Arora. “The exact finite sample properties of the estimators of coefficients in the error components regression models”. In: *Econometrica: Journal of the Econometric Society* (1972), pp. 261–275.

[105] “The Stern Review on the Economic Effects of Climate Change”. In: *Population and Development Review* 32.4 (2006), pp. 793–798. ISSN: 1728-4457. DOI: 10.1111/j.1728-4457.2006.00153.x. URL: http://dx.doi.org/10.1111/j.1728-4457.2006.00153.x.

[106] The World Bank. “Indonesia Economic Quarterly”. In: (Mar. 2016). Ed. by World Bank Group.

[107] The Financial Times. “The great land rush Indonesia: Saving the earth”. In: (May 2016).

[108] John R Townshend et al. “Improving the Utilization of Remotely Sensed Data”. In: *Advances in Land Remote Sensing*. Springer, 2008, pp. 465–483.

[109] G R Van Der Werf et al. “CO2 emissions from forest loss”. In: *Nature Geoscience* 2.11 (Nov. 2009), pp. 737–738.

[110] G R Van Der Werf et al. “Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009)”. In: *Atmospheric Chemistry and Physics* 10.23 (2010), pp. 11707–11735.

[111] Guido R Van der Werf et al. “CO2 emissions from forest loss”. In: *Nature geoscience* 2.11 (2009), pp. 737–738.

[112] Colin Vance and Jacqueline Geoghegan. “Temporal and spatial modelling of tropical deforestation: a survival analysis linking satellite and household survey data”. In: *Agricultural Economics* 27.3 (2002), pp. 317–332.

[113] J. Verbesselt et al. "Detecting trend and seasonal changes in satellite image time series". In: *Remote Sensing of Environment* 114.1 (2010), pp. 106–115.

[114] J. Verbesselt et al. "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series". In: *Remote Sensing of Environment* 114.12 (2010), pp. 2970–2980.

[115] Kristine Verdin et al. *Development of a global slope database for estimation of landslide occurrence resulting from earthquakes*. Tech. rep. 2007.

[116] J Wang, P M Rich, and K P Price. "Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA". In: *International Journal of Remote Sensing* 24.11 (Jan. 2003), pp. 2345–2364. DOI: doi:10.1080/014311603764713427.

[117] Martin Weitzman. "The Optimal Development of Resource Pools". In: *Journal of Economic Theory* June (1976).

[118] S Wunder and B Verbist. "The impact of trade and macroeconomic policies on frontier deforestation". In: (2003).

[119] Muhammad Zikri et al. *An econometric model for deforestation in Indonesia*. Tech. rep. 2009.