Understanding and predicting socioeconomic determinants of deforestation in Vietnam's Central Annamites Landscape (CAL): Pilot study implementing a spatial econometric approach

KATIE P. BERNHARD (Corresponding author), United Nations Development Programme,
Kampala, Uganda¹

STEFANO ZENOBI, World Wide Fund for Nature - Greater Mekong, Thailand

AURÉLIE C. SHAPIRO, Geography Department, Humboldt Universität-zu-Berlin, Berlin,
Germany; Forestry Division, Food and Agricultural Organization, Rome, Italy

Working paper

June 10, 2019

Revised for pre-print March 14, 2021²

¹ Affiliation at time of pilot study: Department of Geography and Environment, London School of Economics and Political Science. Corresponding author email: katiepbernhard@gmail.com

² This pilot study was conducted and working paper drafted initially as a World Wide Fund for Nature (WWF) – Greater Mekong report in 2018-19. The authors are grateful to WWF Vietnam, and to Benjamin Rawson, Thibault Ledecq and Tam Le Viet for data collection support, review, and comments. This project received funding from WWF Greater Mekong and WWF Switzerland.

Abstract

The forests of the Greater Mekong Subregion, consisting of Myanmar, Thailand, Cambodia, Laos and Vietnam, are under high pressure from economic development and exploitation of natural resources, including but not limited to land concession, smallholder plantations and commercial agriculture, agroforestry development, mining, and road infrastructure development. While these threats are well-known, the magnitude and dynamics of their individual and interacting effects on forest cover are not fully understood. This pilot study aims to apply existing, publicly available macro, micro, and socioeconomic data in addition to remote sensing forest cover data to explore economic determinants of deforestation in the Greater Mekong, using the case of the Central Annamites Landscape (CAL) ecoregion of Vietnam. A longitudinal panel was constructed for 2000-2017, containing 1,658 observations for 144 variables across 95 Tier 2 (district) administrative units in CAL from 2000-2017 for modeling macroeconomic and socioeconomic conditions against annual tree cover loss aggregated to the Tier 2 administrative unit. The first phase of the study used tiered spatial regression analysis to correlatively identify which commodities, economic development activities, and social conditions have historically had the greatest effect on forest cover by magnitude in CAL. Based on first phase results, we selected a subset of these determinants for scenario modelling to predict possible deforestation outcomes given certain economic scenarios. The results, among others, indicate that, when spatially collocated, high poverty rates and smaller scale agricultural land conversion are key immediate determinants of deforestation. This therefore provides evidence to support programs targeting rubber, acacia harvesting, artisanal mining, and land conversion for cash crop plantations. Education is also key immediate socioeconomic factor, as poverty rate is consistently associated with tree cover loss and increased educational attainment is consistently associated with reduced tree cover loss, at approximately 12 ha per percentage increase in secondary school graduation. On the macro level, economic growth in China and Vietnam are correlatively associated with tree cover loss, as are rising trade in Myanmar and Laos. This study provides a methodological contribution to the current academic literature identifying socioeconomic dimensions of deforestation through spatial econometric analysis and scenario modelling at the landscape level. For practitioner work, this pilot provides a model or tool for strategic planning of conservation interventions in light of economic conditions and factors.

Keywords: Tree cover loss, deforestation, economic development, land conversion, Environmental Kuznets Curve, poverty-environment, spatial econometrics, Vietnam

1. Introduction

The Greater Mekong Subregion, which includes Myanmar, Thailand, Cambodia, Laos and Vietnam, hosts one of the world's most biodiverse forests while simultaneously supporting a large human population. The region's forests are under high pressure from economic development and exploitation of natural resources, including but not limited to economic land concession, smallholder plantations and commercial agriculture, agroforestry development, mining, and roads. While these threats are well-known, the magnitude and dynamics of their individual and interacting effects on forest cover are not fully understood.

Between 1990 and 2015, approximately 4.7 million hectares (ha) of forest are reported to have been lost in the Greater Mekong Subregion (GMS), amounting to 2.5% of the total land area (Costenbader et. al, 2015). Average annual decrease in forest cover across the GMS has been approximately 0.21% over the period. Biodiversity loss and species extinctions in the GMS have been attributed to this tree cover loss through habitat fragmentation, degradation, and destruction (WWF-GM Forests Report, 2018). In Vietnam, this tree cover loss has persisted despite reforestation progress in some areas. For example, mountainous regions of Vietnam experienced “forest transition” during the 1990s, with observed overall annual forest cover increases driven largely by reduced forest-to-agriculture conversion as a result of intensification of agriculture on existing land (Meyfroidt and Lambin, 2007; Hansen et al., 2013). However, economic and socioeconomic drivers of deforestation persist, and, as the Central Annamites Landscape (CAL) demonstrates, reforestation is neither spatially nor temporally consistent (Figures 1 and 3). While some subnational regions may undergo this “forest transition,” other regions may still experience deforestation. Indeed, Figure 1 shows average annual tree cover loss in CAL districts from 2001-2017. Reforestation areas may actually accelerate deforestation in other areas by displacement or leakage, when the drivers of deforestation persist (Meyfroidt and Lambin, 2009). Additionally, reforestation may not eliminate forest edge effects and habitat fragmentation, as reforestation does not necessarily ensure density or quality, in terms of biodiversity, of new forest. (Meyfroidt and Lambin, 2008b).

Figure 1. Average tree cover loss trends by CAL province, in hectares, for the landscape area of interest. Overall, tree cover loss increased from 2001-2017 for all provinces, with a slow in tree cover loss 2016-2017 (Data: GFW, 2018).

The objectives of this pilot study are twofold: First, we sought to systematically assess available macro and microeconomic data that can be used for modeling economic determinants of deforestation in the Central Annamites Landscape (CAL) ecoregion of Vietnam (Figures 2 and 3).

Figure 2. Area of interest in Vietnam. Left: Vietnamese provinces, the Tier 1 administrative level. Right: Vietnamese districts, the Tier 2 administrative level. In red are the Tier 2 districts of interest in the Central Annamites Landscape. These districts are within nine central provinces: Ha Tinh, Nghe An, Quang Binh, Quang Tri, Thua Thien Hue, Da Nang, Quang Nam, Quang Ngai, Kon Tum, and Binh Dinh (Data: GADM, 2016). Binh Dinh was ultimately removed from analysis it has insufficient complete forest cover.

Second, we use these data to conduct initial modeling of specific socioeconomic determinants of deforestation in the CAL, focusing on economic activities such as production of key commodities in addition to social conditions. Our main research questions were the following:

- Which commodities, economic development activities, and social conditions have historically had the greatest effect on forest cover by magnitude in the CAL?
- How do different commodities interact in leading to this deforestation?
- If the current trends of economic growth are maintained, to what extent will deforestation be affected? Spatially, which areas in the CAL are most “at risk”?

To answer these questions, the initial phase of this study identified principal determinants and their historical magnitude and interaction from 2000-2017. Based on the results of the first phase, we selected a subset of these determinants of interest for scenario modelling in order to predict possible deforestation outcomes given scenarios in these interest areas. This study provides a methodological contribution to the current academic literature identifying socioeconomic dimensions of deforestation through spatial econometric analysis and scenario modelling at the landscape level. From the practitioner perspective, we hope that this pilot study provides a useful tool for strategic planning of conservation interventions in light of economic conditions and factors, not only through the empirical analyses conducted but also through our systematic assessment of currently available data for understanding these issues. As a pilot, we hope to highlight areas of need for additional data collection and future research. Lastly, we hope to move beyond or provide empirical justification for the “deforesting commodity of the year” effect for

certain commodities based on visible investments and media attention, as this focus does not allow for long term programmatic planning.

Figure 3. Left: Forest cover in 2017 in the CAL area of interest. Green indicates forest cover. Data: WWF Forest Cover (Data: GFW, 2018). Forest cover and forest cover loss aggregated to district for analysis to account for economic and socioeconomic data. 1m cell. Right: Average annual forest cover loss in the CAL area of interest (sq m), using inverse distance weight (IDW) interpolation of district level forest cover loss averages (power = 0.2 to smooth across districts; interpolated to 5 surrounding points) (Data: GFW, 2018).

1.1 Conceptual Framework

There is a robust literature seeking to understand both the drivers and the consequences of deforestation. Similar to several studies, such as Khuc et al. (2018), we utilized the Angelsen and Kaimowitz (1999) theoretical model of socioeconomic determinants of deforestation, further detailed in Figure 4. We selected framework because even alternative models in the literature have in many cases been variations of the Angelsen and Kaimowitz approach, which is a foundational model in the study of socioeconomic determinants of deforestation (Chomitz, 2006). The Angelsen and Kaimowitz framework identifies three levels of determinants: underlying macroeconomic causes, immediate decision parameters, and agent-level decision-making.

Figure 4. Variables affecting deforestation. Adapted from Angelsen and Kaimowitz (1999).

As we apply this conceptual and model framework in this study, deforestation at location i and time t is thus a function of underlying and immediate determinants and agent decision parameters,

$$Deforestation_{it} = f(\text{underlying}_{it}, \text{immediate}_{it}, \text{agent}_{it})$$

and each of these determinants is, in turn, a nested function of key economic and socioeconomic conditions including commodity prices, GDP per capita, GDP growth, trade liberalization, income and poverty, population density, education levels, off-farm wage opportunities, agricultural suitability of land, and an agent's background, preferences, and resources.

$$\text{underlying}_{it} = f(\text{commodityprices}_t, \text{gdp}_t, \text{growth}_t, \text{tradelib}_t,)$$

$$\text{immediate}_{it} = f(\text{poverty}_{it}, \text{popdens}_{it}, \text{education}_{it}, \text{wageopps}_{it}, \text{agsuitability}_{it})$$

$$\text{agent}_{it} = f(\text{background}_i, \text{preferences}_i, \text{resources}_{it})$$

(Angelsen and Kaimowitz, 1999; Khuc et al., 2018)

As discussed in Scrieciu (2007), the inclusion of all three levels of determinants in a single model can result in endogeneity and lack of clarity in identifying causal links, as ultimately, variables determining agent decision-making are affected by immediate determinants, in turn effected by underlying determinants. We apply this in our application of the framework and have discussed the variables used in further detail in Table 1. We also review the literature describing our main variables below.

First, poverty-environment linkages have long been the subject of scholarly debate (Chomitz, 2006). Barbier (2017) finds that the relationship between income and poverty and deforestation is complex and often context-dependent. Off-farm employment opportunities are an important determinant for the effect that poverty may have on deforestation (Busch and Ferretti-Gallon, 2017). In Vietnam, the provincial level averages of poverty rates (Figure 5) demonstrate that the overall trend is improving (GSOV, 2018). This decline is the result of several economic and policy mechanisms. In this study we seek to understand both the effect of poverty on deforestation in CAL but also to describe the relationships between some of these economic mechanisms. In addition to provincial-level poverty data, we use the intensity of nighttime lights at the district level to proxy human development and industrial development, as this variable is collinear with road infrastructure. We also proxy access to off-farm wages with the proportion of each district that is covered by urban area.

Figure 5. Poverty rate trends by CAL province, according to GSOV data, for roughly 2006-2016 with dropout years. Importantly, GSOV states: “Monthly average income per capita of household which is updated by CPI as follows: 400 thousand dong for rural areas and 500 thousand dong for urban areas in 2010; 570 thousand dong and 710 thousand dong in 2013; 605 thousand dong and 750 thousand dong in 2014; 615 thousand dong and 760 thousand dong in 2015 and 630 thousand dong and 780 thousand dong in 2016, respectively.” (Data: GSOV)

Similar to poverty-environment linkages, the effect of economic growth and GDP on the environment is contested in the literature. According to Environmental Kuznets Curve (EKC) theories, economic growth degrades the environment until a certain point, after which the benefits of economic growth and development, such as enhanced technology coupled with stronger regulations and institutions, leads to improved environmental outcomes (Panayotou, 1994). However, evidence for this relationship exists only for certain types of environmental degradation, such as air pollution, while the effect on, other degradation types, such as deforestation, is unclear (Cropper and Griffiths, 1994). Bhattari and Hammig (2001) find evidence of a “deforestation EKC” mostly reliant on shifting energy use away from forest-based resources, and improved institutional structures with increased economic development. However, economic growth in Vietnam and its neighbors may have resulted in increasing pressure on forests to support both economic activities and growing populations (Figure 6). Identifying the links between deforestation and economic growth is thus a priority for conservation and the environmental sustainability of economic activities moving forward. In part, this study fits into the EKC literature by assessing the relationship between various indicators of economic growth and development at the landscape level and deforestation.

Figure 6. Economic growth from 2001-2017 for Vietnam and selected neighboring economies, Laos, Cambodia, and China. The causes and consequences of negative Cambodian growth rate in 2009 for environmental degradation could also be an interesting focus of future research (Data: WDI, 2018).

Next, increasing human population, particularly increasing population density in proximity to forests and protected areas can be an important determinant of deforestation. Supporting a large human population can lead to more extraction of forest resources, while population density in close proximity to forests, especially in high-poverty areas, can result in community forest dependence (Cropper and Griffiths, 1994; DeFries et al., 2010). Additionally, Pfaff (1999) examined satellite & socio-economic municipality-level data and found deforestation driven by land characteristics, such as soil quality and vegetation type, and factors affecting transport costs such as distance to markets and roads. According to Pfaff, the effects of population growth on deforestation, modeled through migration and credit availability, depended on the preexisting population. We use population density estimates at the district level to understand the dynamics of this population and deforestation in CAL.

Lastly, land use change for agricultural and industrial production has indeed proven to be a key driver of deforestation in CAL. A number of studies have previously used spatial and econometric methods to understand the effect of land use change or socioeconomic conditions on deforestation. These have primarily been at one of two scales: finer, sub-district level, and national-scale. First, several studies have focused on a fine spatial scale, within a single district or province, relying on small scale surveys and/or data collection at the household level (Müller and Zeller, 2002; Webb and Kiyoshi, 2007; Chi et al., 2013). These have not generally incorporated multiple temporal waves due to reliance on limited spatial land use data availability. Müller and Zeller (2002) used spatial econometric methods, supplemented by interviews, to identify drivers of deforestation in the Dak Lak Province specifically in 1975, 1992, and 2000. This study found that deforestation in this province has historically been driven by land-intensive agricultural production. However, as previously discussed, implementing protected areas and “policies discouraging shifting cultivation” in the later 1990s reduced forest pressure while promoting economic growth. (Müller and Zeller, 2002; Meyfroidt and Lambin, 2007). In this study, we hope to contribute insights from CAL from the 2000s.

Second, a number of studies have been conducted at the national scale, incorporating all regions and ecoregions of Vietnam. Khuc et al. (2018) found through spatial and regression-based methods that, at the national level, primary underlying indicators of deforestation and forest degradation include initial forest cover, per capita income, agricultural production, governance, population growth, food, and poverty. Further, the study identified north central, northeast, central highland, and northwest areas of Vietnam as particular risk areas. We use these findings as key support for our identification strategy for econometric analysis in this study.

So far no other studies have sought to quantify the relationship between socioeconomic conditions and deforestation in the CAL at the landscape level using spatially explicit methods. Because the spatially explicit previous studies have also relied on one, or in some cases two, maps of land use change, our study is among the first to use dense time series commodity price data for underlying conditions and as a possible signal of land use change leading to deforestation in Vietnam (Munroe and Müller, 2007). Cropper and Griffiths (1994) found hard log price to be correlated with deforestation at the continent level in Latin America, but neither Africa nor Asia. Scricciu (2007) found that global commodity price increases are linked to increasing deforestation, but this study was conducted on a broad scale, similar to that of Cropper and Griffiths, using a panel of fifty tropical countries over a period of eighteen years. Figure 7 shows trends in selected commodity prices relative to tree cover loss trends.

Figure 7. Combination plot, commodity price trends (left axis) relative to tree cover loss in CAL by m^2 (right axis) over time from 2000-2015 with 2016 and 2017 not included due to dropout (Data: IMF, 2018; GFW, 2018).

In summary, in this study we apply the Angelsen and Kaimowitz (1999) framework for underlying and immediate determinants of deforestation and base our identification and econometric analysis methods on the results of previous spatial econometric studies in Vietnam, such as Khuc et al. (2018). Based on this literature, income and poverty, economic growth, population density, and land use change are key variables in our analysis. Finally, as previously mentioned, Table 1 describes the rationale behind the use of data for specific variables in this study to capture the key socioeconomic conditions emphasized in the literature, and to capture parameters of interest in alignment with the conceptual framework.

2. Methods

2.1 Data

We constructed a panel of ecological, economic, and social variables for analysis. Prior to constructing this panel, as a key component of this study we systematically reviewed open-source secondary data available for forest cover and economic variables, the findings of which are available in Appendix II. Data availability and quality confined the scope of our study, but also identified important areas for future data collection and research.

The primary response variable used in this study was annual tree cover loss aggregated to the Tier 2 administrative unit. This was extracted from annual raster data acquired from WWF and Global Forest Watch (Hansen, 2013), then aggregated to coarse district-level means. Economic data was sourced from a variety of institutions including IMF, World Bank DataBank, World Database on Protected Areas, the AidData Geoquery platform and University of Delaware, and the General Statistics Office of Vietnam (GSOV). Biophysical factors were included in spatial regression, such as elevation, slope, temperature, and precipitation (aggregated to Tier 2 administrative unit by mean), but time-invariant biophysical factors were not included in site-demeaned fixed effects as those specifications account for time-invariant factors that vary between sites. Proportion of a Tier 2 unit in a protected area, in addition to previously existing forest cover extent as a control variable, were also included in spatial regression. Aggregation to administrative unit rather than grid cells was selected to maintain consistency with the socioeconomic variables from household surveys, which are weighting and representative according to administrative units.

Coarse-scale variables for underlying macroeconomic conditions driving deforestation included per capita GDP (PPP, international), trade as a percentage of GDP, and annual GDP growth for Vietnam and surrounding economies; global commodity prices for commodities relevant to Vietnamese and CAL economies (including but not limited to rubber, rice, coffee, hardwood price indices, softwood price indices, metals price indices, cereals price indices, etc.). We also included annual quantity of foreign direct investment flow delivered to specific districts and earmarked by World Bank. These variables generally exhibited temporal but not spatial variation. For the immediate determinants of deforestation, we include variables which had both spatial and temporal variation, such as population density, poverty rates, proportion of households with access to electricity, district-level contributions to national GDP, proportion of urban area per district, land area devoted to irrigated agriculture per district, nighttime lights, and accessibility by roads. Table 1 outlines the key variables used and expected relationships with deforestation based on the literature and the CAL context.

Table 1. Expected and observed outcomes for driver variables based on CAL context-specific and metanalysis evidence from the landscape change and deforestation literatures. See Summary Statistics (Appendix I) for variable detail and units.

		Variable	Expected Effect on Deforestation	Observed Effect on Deforestation (see Results)	Rationale for proxy (if proxy)	Literature Notes
Underlying causes: Macro	Biophysical	Elevation	-	-	N/A	These control variables are in addition to the existing forest cover in each district. Biophysical variables and preexisting forest extent will limit agricultural production and the amount of deforestation possible. This is clearly true for time-varying precipitation and temperature but also is an important factor for mountainous forest regions like CAL
		Slope	+	+		
		Precipitation	+	+		
		Temperature	+	+		
	Economic growth	Per capita GDP (PPP)	+/-	-	N/A	Literature is not conclusive on role of economic growth in deforestation. See Environmental Kuznets Curve debate. (Chomitz, 2006; Busch and Ferretti-Gallon, 2017)
		GDP growth annual				
	Population pressure	Population density	+	+	N/A	Growing population needs more land for extracting forest products, population growth can also lead to tech progress, institutional development. Overall lit finds + relationship (Busch and Ferretti, 2017; Angelsen and Kaimowitz, 1999)
	Income level	Consumption	-	+	Testing a variety of proxies, incl. rate, quintile, and consumption In this case increased consumption would be a proxy for poverty/reduction and increased economic choice Although consumption can also lead to increased deforestation through consumer demand...	Literature is not conclusive on role of poverty, income. This also returns to the Environmental Kuznets Curve debate. (Chomitz, 2006; Busch and Ferretti-Gallon, 2017)
	Tech progress	Ag technology	-	...	Could be proxied by “irrigated ag” as well if in reference to the	Literature establishes that ag tech increases yields without

					transition from low yield to industrial	requiring increasing deforestation (Chomitz, 2006)
	Global price indices	Timber price	+	+	N/A	Controversial in literature review: can increase defor by making logging more profitable, but also can maintain forest by preventing forested areas from being completely cleared for other ag or industrial dev. Depends on species, plantation requirements, etc. (Angelsen and Kaimowitz, 1999; Chomitz, 2006; Busch and Ferretti-Gallon, 2017) Also differentiating here between underlying global price fluctuation and local price fluctuation as a direct driver
		Coffee price	+	+	N/A	Increase in price for ag commodities has, in previous studies, encouraged transition to ag from forest, unless farmer is operating on a subsistence model in which case higher prices allow for less labor and thus less land clearing (Angelsen and Kaimowitz, 1999)
		Rice price	+	+	N/A	
		Metals price	+	+	N/A	
		Commodities indices	+	+	General commodity indices for fuel and nonfuel	
	Neighboring and regional economies (NRE): Laos Myanmar Cambodia China Bangladesh India Japan Korea Singapore Hong Kong Nepal Indonesia Malaysia	Per capita GDP (PPP) % GDP growth	+/-	-	N/A	Literature is not conclusive on role of economic growth in deforestation. However, we expect that neighboring and regional economies have not reached a point of passing beyond the level of growth at which environmental benefits can be expected according to EKC theory. (Chomitz, 2006; Busch and Ferretti-Gallon, 2017)
	Trade liberalization (Vietnam and NRE)	Trade as a percentage of GDP	+/-	-		Literature is not conclusive on role of trade liberalization in deforestation. See Environmental Kuznets Curve debate. (Chomitz, 2006; Busch and Ferretti-Gallon, 2017)

Immediate causes: Micro	Poverty rate	Poverty	+/-	...	Testing a variety of proxies, incl. rate, quintile, and consumption	Literature is not conclusive (Chomitz, 2006; Busch and Ferretti-Gallon, 2017)
		Poverty quintile	+/-	+	Testing a variety of proxies, incl. rate, quintile, and consumption	(Angelsen and Kaimowitz, 1999)
	Irrigated cropland vs. rainfed cropland	Area of irrigated cropland by Tier 2 unit per year	+	+	Based on ESA landcover and imperfect, but looking into Croplands and FORMIS as well Proxying for commercial scale, more industrialized agriculture vs. subsistence agriculture (but this also depends on the crop itself)	Farmers will not always be incentivized for land transition if ag on existing land becomes more productive (CGE=assumes that farmers are profit maximizers) (Angelsen and Kaimowitz, 1999)
	Accessibility	Nightlights	-	-	Can proxy “human development” at the local level	(Bruederle and Hodler, 2018)
		Roads	+	+	Partial endogeneity can be in part dealt with by also looking at time-lagged population density	Well established effect in the literature (Busch and Ferretti-Gallon, 2017) although Angelsen and Kaimowitz (1999) argue that some of the effect is overstated due to partial endogeneity
	Agricultural input prices	Ag input price index	+/-	...	Increase in input prices could lead to adoption of more extensive production with more land using fewer inputs, or could to reduction in ag by reducing profitability	(Angelsen and Kaimowitz, 1999)
	Off-farm opportunities	Area of “urban space” within Tier 2 unit	+/-	-	This could proxy a couple of different things. Could proxy off-farm opportunities, in which case the lit would find a -effect (Angelsen and Kaimowitz, 1999; Busch and Ferretti-Gallon, 2017). Or, and more likely, could proxy high human pop dens areas and greater pressure to deforest	(Angelsen and Kaimowitz, 1999; Busch and Ferretti-Gallon, 2017; Webb and Kiyoshi, 2007)
		Off-farm wages	-	...	Movement away from agricultural or other work that results in clearing forest land	Well established reduction in deforestation documented in the literature (Busch and Ferretti-Gallon, 2017)
	Domestic prices	Local timber price	+/-	...	Specific focus on acacia, eucalyptus, rubber. See Timber Price in Underlying causes; effect of timber prices is contested in the literature but is a key component	Looking at context-specific literature and fluctuations based on local land tenure regimes, policy, socio-cultural factors, as local price fluctuations can

		Local coffee price	+	...	Coffee price may not have as visible an effect on tree cover loss due to shade-growing practices that may not come through on a district-level aggregation of tree cover loss	incentivize more or less deforestation but ultimately limited by ownership and underlying policy (Angelsen and Kaimowitz, 1999; Busch and Ferretti-Gallon, 2017; Chomitz, 2006, Webb and Kiyoshi, 2007; Chi et al., 2013)
		Rubber price	+	...	Rubber has been identified by conservationists as a major current cause of deforestation. Fluctuations in price can effect extraction pressure on the ground.	
		Rice price	+	...	Rice is a food staple and key component of agricultural production	
		Metals price	+	...	Incentives for mining and illegal mining	
		Commodities prices	+	...	Overall market fluctuations	
	Education	Secondary school graduation rate by Tier 2 average	-	+/-	Could proxy ability to explore off-farm opportunities and improving human development	(Busch and Ferretti-Gallon, 2017; Chi et al., 2013)

The panel was constructed for the time period 2000-2017, and thus contains 1,658 observations for 144 variables across 95 Tier 2 (district) administrative units in the CAL from 2000-2017 (Figure 1; Summary Statistics, Appendix I). The CAL boundaries and districts used for analysis were delineated using boundaries provided by WWF-Vietnam. The Tier 2 administrative unit was selected as this provided enough site units within a panel for statistical purposes, but sufficient socioeconomic data was still available for this scale. The panel is unbalanced, as some variables are missing for certain districts or time periods. Nevertheless, the use of panel data enables analysis of temporal and spatial trends in addition to application of a variety of econometric techniques. We first performed simple correlative statistics and exploratory spatial data analysis, which was followed by spatial regression analysis which explored spatially autoregressive and spatial lag of x (SLX) models. Lastly, we utilized the panel nature of the dataset to perform site fixed effects modeling.

2.2 Spatial and econometric analysis

We first assessed district-level associations and spatial autocorrelation, particularly for the purposes of constructing spatial lags of key variables exhibiting spatial clustering or dispersion. Spatial autocorrelation was assessed using the univariate local Moran's I statistic in addition to bivariate local Moran's I for roughly establishing spatial association between socioeconomic variables and tree cover loss. We chose a row-normalized queen contiguity matrix,

$$W_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}$$

(Lee, 2001)

where $w_{ij}=1$ indicates contiguity between districts i and j . Clustering or dispersion is thus assessed using the average of a variable's values for districts directly bordering the district of interest. As a part of this pilot study, we experimented with a number of alternative spatial weighting options and thus in some instances also utilized a k-nearest neighbor (KNN) approach,

$$w(s^*, s_j) = \begin{cases} 1/k & \text{if } j \text{ is a nearest neighbor to location } s^* \\ 0 & \text{otherwise} \end{cases}$$

(Fotheringham et. al., 2000)

where $w(s^*, s_j)$, like W_{ij} above, is a scalar weight assigned to administrative unit j based on its distance from location s^* . For KNN weights, a number of "nearest neighbors" is selected to create the weight, essentially a moving average of "neighbor" values for the district of interest. The primary difference to note between the contiguity and nearest neighbor weights is illustrated in Figure 4. In a contiguity weight, the value of a variable at district E would not be included in the lag for district C, even though they are spatially near each other, because they do not share a border. However, in a KNN weight, E would be included in the lag for district C because E is among C's five nearest neighbors. In CAL, the most common range of frequencies of contiguous neighbors for a given Tier 2 district is 5-7, so we selected five neighbors as the number of averaged neighbors to establish a given district's weight. We do not use distance-band weights such as Euclidean or inverse distance weighting due to the aggregation of our dataset to administrative units (Anselin, 2018).

Figure 8. The primary relevant difference between weights structures is the inclusion of non-contiguous near neighbors in the KNN weight such as district E in district C's lag, unlike the contiguity weight, which would include only neighbor values of A, B, and D in the lag for C.

On a -1 to 1 scale, the univariate Moran's I statistic compares a variable to its own spatial lag (see Appendix III). The bivariate version of the Moran's I is mathematically equivalent to the spatial autoregression of y (Wy) on x , where the relationship between a non-lagged variable and the "neighborhood" of a lagged variable is assessed:

$$I_B = \frac{\sum_i (\sum_j W_{ij} y_j \times x_i)}{\sum_i x_i^2}$$

(Anselin, 2018)

Following the identification of spatial autocorrelation and bivariate spatial association, we then used correlative statistics, including Pearson's correlation matrices and scatter plot matrices with locally weighted scatterplot smoothing (LOWESS), to better understand the relationships and collinearity between variables used (Table 3; Figures 9 and 10). Next, we performed spatial regression analysis on two separate scales, following the Angelsen and Kaimowitz (1999) approach: for the underlying macroeconomic conditions driving deforestation, and the immediate microeconomic determinants. We separated these two sets of conditions as described by Angelsen and Kaimowitz (1999) and recommended by Scricciu (2007). As describes in the Literature Review and Figure 4, underlying causes and immediate causes are often endogenous or collinear, as many underlying macroeconomic conditions will ultimately create the conditions for the immediate ones.

We modeled underlying macroeconomic determinants of deforestation using the following simple linear OLS estimation with spatially lagged explanatory variables:

$$Loss_{at} = \alpha + \beta_1 X_{at} + \beta_2 WX_{at} + \beta_3 GP_t + \beta_5 NE_t + \varepsilon_{at}$$

(Bernhard et al., 2020; Gibbons et. al, 2012; Angelsen and Kaimowitz, 1999)

where $Loss_{dt}$ is tree cover loss by district d and year t . X_{dt} is a vector of biophysical controls and W indicates their spatial lags. GP is a vector of global commodity prices by year t . NE is a vector of conditions of Vietnam and neighboring and regional economies such as Laos, Cambodia, Myanmar, China, etc., including GDP per capita, GDP growth, and trade liberalization by trade as a percentage of GDP. α is the constant and ε_{dt} is the error term. Variables were tested for normal distribution using a Shapiro-Wilks test and abnormally distributed variables were log-transformed, retested, and skew-corrected if necessary (see raw variable distributions in Figure 9 and 10 histograms for those that required skew correction). This model skeleton was thus used for linear, log-linear, and log-log specifications.

We used temporal lags of one to two years where relevant in order to account for the time delay between a change in price or economic conditions and the resulting effect on deforestation. We also further expound upon our hypotheses for relationships between these variables and deforestation in Table 1. A simple spatial lag of x model with OLS linear estimation was used due to the time series nature of the global and regional trend variables. Without a site-level fluctuation, methods such as fixed effects were not usable. As such, this estimation is subject to limitations discussed in the following section but nonetheless provides a useful baseline for understanding key commodities and regional economies' effects on CAL deforestation.

Next, the more immediate social and economic determinants of deforestation were estimated using the following model. With district level data from 2000-2017, we used a site fixed-effect to account for time-invariant unobservables between districts. To select fixed effects over random effects we performed a Hausman test (threshold $p < 0.05$) without clustered standard errors and a Sargan-Hansen test (threshold $p < 0.05$) with clustered standard errors, both of which confirmed fixed over random effects. Because the panel is unbalanced due to missing data for certain districts in certain years, we were unable to perform space-time fixed effects and our sample size was reduced. However, like the underlying conditions model, we employed temporally lagged variables to account for time delay in effect on deforestation where relevant.

$$Loss_{dt} = \mu_d + \gamma_t + \rho_{dpt} + \varepsilon_{dpt}$$

(Bernhard et al., 2020; Hsiang and Sekar, 2016)

μ_d is the sector fixed effect, γ_t is time (year). ρ_{dpt} is a vector of time-varying socioeconomic conditions at district level d , and provincial level p , including poverty rates, education levels, proxy for off-farm wages (proximity to urban area), proxy for human development (nighttime lights) at time t . Time invariant biophysical controls such as slope and elevation are not included in this term as the fixed effect accounts for those differences, but temperature and precipitation are included as they are time variant. ε_{dpt} denotes the error term, unobservables across districts and provinces.

2.3 Scenarios of economic and socioeconomic conditions

We use the results from regression analysis in Phase 1 for scenario analysis. There are a number of different strategies that could be employed for scenario modeling, including forecasting. For the purposes of this baseline study, we began with simple methods for scenario analysis, using predictive margins based on fixed explanatory variable values. As is discussed further in the Results section, the following variables were identified as key determinants of deforestation: Chinese GDP and economic growth, rubber prices, poverty and/or local level of development at the district level, and education. Based on the Phase 1 models selected in the Results section, we

fixed the values of these four variables to a set level using predictive margins, to determine how that variable's effect on tree cover loss would change from the original model.

Our four key interest areas identified by Phase 1 modeling are displayed in Table 2. Table 2, which is also readapted in Table 13 to show how models were fit with fixed values for the variable of interest. Fixed values were selected using the existing historical minimum, maximum, and mean values for each raw variable. We followed the most successful model utilizing each variable, prioritizing R^2 values. X is a vector of non-interest variables (the remainder of variables run normally in the regression) for a log-log spatial OLS. ε_{dpt} is the error term for district d , province p , and at time t .

Table 2. Scenarios fit for Phase 2 using predictive margins based on results of Phase 1 modeling.

<p>Chinese economic growth</p> $\ln (Loss)_{dt} = \alpha + \beta_1 \ln [X]_{dt} + \beta_2 WX_{dt} + \beta_x \ln [China]_t + \varepsilon_{dt}$ <p>Table 6 M2</p> <p><u>Possible outcomes:</u> Economic contraction (<i>China</i> = 7.983889) Maintained growth (<i>China</i> = 8.912644) Economic boom (<i>China</i> = 9.729535)</p>	<p>Rubber price fluctuation</p> $\ln (Loss)_{dt} = \alpha + \beta_1 \ln [X]_{dt} + \beta_2 WX_{dt} + \beta_x \ln [Rubber]_t + \varepsilon_{dt}$ <p>Table 7 M2</p> <p><u>Possible outcomes:</u> Drop in price (<i>Rubber</i> = 3.261404) Increase in price (<i>Rubber</i> = 5.386853)</p>
<p>Poverty reduction</p> $Loss_{dpt} = \alpha + \beta_1 X_{dpt} + \beta_2 WX_{dpt} + \beta_x Pov_t + \varepsilon_{dpt}$ <p>Table 6 M4 (using poverty rate quintiles)</p> <p><u>Possible outcomes:</u> Worsening poverty (<i>Pov</i> = 2) Maintain current levels (<i>Pov</i> = 3) Continued improvement (<i>Pov</i> = 5)</p>	<p>Education level improvement</p> $Loss_{dpt} = \alpha + \beta_1 X_{dpt} + \beta_2 WX_{dpt} + \beta_x Edu_t + \varepsilon_{dpt}$ <p>Table 8 M1</p> <p><u>Possible outcomes:</u> Current secondary graduation rates are maintained (<i>Edu</i> = 91.42638) Graduation rates increase (<i>Edu</i> = 99.86)</p>

Predictive margins are essentially a simple tool for answering “what if”-type questions without complex forecasting requiring high R^2 values. Essentially, these margins enable controlling for compositional differences in the explanatory variables to be able to say, roughly, what the tree cover loss outcome could be in that case. For example, how would tree cover loss change in this model if rubber prices increased to their observed maximum? Or, how would tree cover loss change if all district secondary school graduation rates increased to the currently observed maximum? The predictive margins results are not interpretable alone and are be compared to the tree cover loss outcome in the original model on which the predictive margins are based.

2.4 Limitations

The methods used in this study are subject to standard issues in econometrics in addition to limitations specific to the uses of spatial data. Our dataset utilizes variables from a variety of sources, each with their own potential biases. Importantly, we note that tree cover loss data from (Hansen 2013) can be the result of a number of activities and is not always deforestation. However,

we assume that tree cover loss is a reasonable proxy for human-driven deforestation. Next, while we have sought to use robust econometric techniques that account for omitted variables and endogeneity, we note that these specifications are imperfect and the underlying conditions model in particular is vulnerable to these issues due to its use of coarse time series data without spatial variation. Lastly, standard errors are unlikely to be entirely independent across observations as policy and local governance variation, in addition to many other factors, effects spatial distribution of deforestation within provinces and districts. Throughout the analysis, we explore clustering standard errors to province level (Tier 1) and district level (Tier 2) to ensure that they are robust to heteroscedasticity and take the more conservative estimate when relevant in terms of p-value. (Stock and Watson, 2006). Our Phase 2 scenario modeling is particularly sensitive to these issues, and the linearity of models run in this study, so these are to be interpreted with extra caution and as a baseline.

This dataset also relies on data from multiple scales, from global macroeconomic conditions and national level statistics to finer-scaled raster data for tree cover loss. Problems with this movement between scales, known as the “ecological fallacy,” are discussed by Moulton (1990). Conclusions drawn about individual units based on aggregated groups are problematic as phenomena occurring at a local scale cannot be fully understood by coarse-scale trends. However, our research design and framework intentionally move from coarse to fine in an effort to understand the different dimensions at each level, from underlying to immediate determinants of deforestation.

As will be discussed further throughout this study, we used available secondary data for this study, which established a baseline for future research. Future studies should utilize full GSOV household survey data for socioeconomic conditions at the district level or finer, for which this study was limited by available online downloads (see Appendix II, Data Availability Assessment for further discussion of this). Additionally, our results isolate limited causality but could be confirmed by identifying instrumental variables or “shock” events to enable use of quasi-experimental and natural experimental methods. Using Lao data could also enable a regression discontinuity design study to a similar effect.

3. Results

3.1 Identifying potential determinants of deforestation

3.1.1 Exploratory spatial data analysis (ESDA)

ESDA revealed spatial autocorrelation of most key variables. While it is clear that biophysical variables would be spatially autocorrelated, as high elevation and high slope areas would be associated in mountainous regions of CAL, for example, identifying clustering of socioeconomic variables relative to deforestation is of standalone interest. Further, this spatial autocorrelation indicates that models failing to account for spatial sorting would be biased.

In addition to spatial autocorrelation, the correlation matrix in Table 3 and scatterplot matrices in Figures 9 and 10 further illustrate the collinearity of variables and bivariate relationships between key variables. Figures 11-16 show this clustering through coverage and statistical maps for several key variables, both biophysical and socioeconomic. Figures 17-19 present results of bivariate local Moran’s I for four potential immediate-scale socioeconomic determinants of deforestation, illustrating spatial association between these variables and deforestation in CAL equivalent to a simple spatial autoregression (see Appendix III). Our key finding from the bivariate statistics is that the hotspot areas of deforestation are not spatially

collocated with population-dense, heavily economic areas. Evidence suggests that higher poverty rates and smaller scale subsistence agriculture are more likely to be driving deforestation at the immediate scale. Figure 17 demonstrates that while high poverty rates are spatially associated with tree cover loss (0.1684), population density does not display this relationship (-0.2661). This provides some evidence that population pressures in the forest frontier areas may not be a key spatial driver of deforestation, but rather immediate determinants of deforestation are more likely to be poverty and expansion of subsistence agriculture. To further support this finding, Figure 19 shows that neither nighttime lights nor irrigated cropland is spatially associated with tree cover loss (-0.3452, -0.1392, respectively). These findings are also consistent with the existing literature focusing on determinants of deforestation in Vietnam (Müller and Munroe, 2005).

Larger scale agriculture and other intense economic activities that are a known cause of deforestation in the landscape may thus be driven by conditions further afield, such as broader economic conditions as will be discussed further in the next section, and therefore this deforestation would be mediated by land tenure and property rights institutions. With government or other institutional support, property rights regimes could be adjusted to disincentivize deforestation.

3.1.2 Underlying Causes Model

In this section we present the results from our assessment of underlying determinants of deforestation in CAL, macroeconomic trends linked to deforestation. First, Tables 4 and 5 are control specifications which highlight the relationships between biophysical control variables, cropland classifications, and tree cover loss in CAL (Table 10 also shows relationships between variables prior to accounting for spatial autocorrelation). Next, due to data inconsistencies in spatial and temporal variation in addition to high levels of collinearity in selected variables, we split similarly trending variables into separate specifications. To illustrate why this step was necessary, Table 3 presents a correlation matrix between prices and Vietnamese GDP. Not only are the selected prices and indices presented highly correlated with GDP, they are also correlated with each other. We expect some degree of this given the aggregation of dense time series to an annual time scale (where prices and indices fluctuate at extremely fine [daily] time scales) and lack of spatial variation. We also may expect that global prices and price indices will track, to some degree, on moderately similar trends at the annual scale given the interconnectedness of the global economy and commodity markets. However, the very high degree of collinearity is problematic for “kitchen sink” models that include all variables, hence our separated and restricted specifications.

Table 3. Pearson’s correlation coefficients between Vietnamese GDP and commodity prices/price indices. Values in red indicate correlated variables that are subject to collinearity and are therefore separated in models presented. This due to use of coarse temporal scale and global nature of prices. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Vietnam GDP	Rice price	Hardwood price index	Softwood price index	Metals price index	Rubber price	Tea coffee cocoa price index	Cereals price index
Vietnam GDP	1.000							
Rice price	0.6939***	1.000						
Hardwood price index	0.8399***	0.8873***	1.000					
Softwood price index	0.0358	-0.1402***	0.0017	1.000				
Metals price index	0.7424***	0.8071***	0.9499***	0.0553**	1.000			
Rubber price	0.4606***	0.7745***	0.8805***	-0.1457***	0.9336***	1.000		
Tea coffee cocoa price index	0.8975***	0.8359***	0.9418***	-0.1468***	0.8699***	0.8348***	1.000	
Cereals price index	0.7707***	0.9114***	0.9401***	-0.1380***	0.8877***	0.8578***	0.8616***	1.000

The results of the Underlying Causes model are presented in Tables 6 and 7, with a combined restricted model in Table 9. Contrary to EKC theories, and in alignment with the literature, it appears that economic growth and increasing trade are associated with tree cover loss in CAL (Bhattari and Hammig, 2001). Table 6 M1 provides evidence that each percent of annual Vietnamese economic growth is associated with approximately 1.99 ha of tree cover loss in CAL and increasing trade may be associated with up to 20.90 ha of tree cover loss annually. Each percent of Chinese GDP growth is associated with a loss of 3.96 ha of forest cover, and similar increases in Chinese trade may be associated with 1.72 ha of tree cover loss in CAL annually. These variables maintain the direction of the effect (in terms of increasing tree cover loss) but lose their significance in Table 7. Rising trade in neighboring economies Laos and Myanmar (also collinear with Cambodia) are associated with tree cover loss, although these were inconsistent from Table 6 to Tables 7 and 8. India and Japan are also strong regional economies significantly associated with tree cover loss. The results are less clear for other regional and neighboring economies. Importantly, many of these neighboring and regional economies are interwoven and therefore endogenous; the result for India and Japan may actually be driven by China's economic growth and its effect on the relationship between India and Japan's economies and deforestation.

Of commodity prices and price indices assessed, rubber and hardwood price present the greatest magnitude and consistency in their effect on tree cover loss. Increases in rubber price are consistently associated with increasing tree cover loss, both within the same year and within a one-year time lag. A one-unit increase in rubber price annually can result in a loss of between 2.7 and 10.5 ha of forest cover, and such an increase can result in 7.9 ha of loss in the following year. Hardwood price had a large relative magnitude of effect, as a one-unit increase in the value of the index was associated with a 49.79 ha loss in forest cover in Table 6. Increases in metals price were associated with increasing tree cover loss in specifications where it was statistically significant, but its significance was inconsistent across specifications. Food commodity prices such as cereals and rice were also less consistent, although the tea-coffee-cocoa index, like metals prices, maintained the direction of effect (increasing tree cover loss) in all specifications where it was statistically significant. For rice prices, the relationship changes between within-year and the temporally lagged rice price in Table 6.

3.1.3 Immediate Causes Model

The results of the Immediate Causes model are presented in Table 8, with a combined restricted model in Table 9. Both fixed-effects and standard spatial regression are included in the table. Population density is not significant in any of these specifications, likely due to lack of spatial collocation with forest cover. Poverty rate is consistently associated with tree cover loss, while urban area is consistently (though insignificantly) associated with reduction in tree cover loss. This may be due to a combination of lack of spatial collocation, like the population density variable, and increased off-farm wage and employment opportunities in urban areas. That said, while significant economic activity is not spatially associated with tree cover loss, those districts contributing more to overall Vietnamese GDP appear to have greater tree cover loss than those contributing less to overall GDP. This contribution to GDP could be in agriculture, but this result highlights a need for further data collection and research on each districts' primary economic activities that contribute to GDP, with spatial and temporal variation. Importantly, districts with higher secondary school graduation rates demonstrate reduced tree cover loss, and this trend was consistent and significant across all specifications. Table 8 M1 indicates that increasing the rate of secondary school graduation by 1% can reduce tree cover loss by as much as 12.7 ha. Of agricultural land types, rainfed agriculture and mosaic cropland appear to have the greatest effect

on tree cover loss. While Table 5 M3 shows that irrigated cropland was significantly associated with tree cover loss, this was not found in any specification that included biophysical control variables.

3.1.4 District and Provincial Sensitivity

Based on these results, particularly the combined restricted model in Table 9, we ran district and province-specific regressions to highlight the relative effects of key variables, and then created “sensitivity maps” that visualize which districts and provinces of CAL are the most sensitive to the effects of deforestation determinants by magnitude. Tables 11 and 12 present the results of those factor variable (district interaction) models and Figures 20 and 21 present the mapped visualizations of these regressions. The results indicate that the districts with the most sensitivity to fluctuations in underlying economic conditions are located in the provinces of Quảng Bình, Quảng Trị, and Thừa Thiên Huế. Returning to Figure 2, these are also provinces with a large number of frontier forest districts, where the areas of high forest cover begin to meet the most heavily economic and population-dense coastal areas. As such, both high tree cover loss and high sensitivity to economic conditions exist in districts of those CAL provinces. Importantly, Figure 20 (bottom right map) illustrates that a number of districts with high tree cover loss are also highly sensitive to improvement in immediate economic conditions such as educational attainment by secondary school graduation rate. The districts of Bắc Trà My, Trà Bồng, Cam Lộ, Triệu Phong, Hương Trà, and Đa Krông are strategically both higher tree cover loss and sensitive to this immediate driver.

We used the results of these Underlying and Immediate Determinants models, and the district-specific sensitivity specifications, first to create a “deforestation risk map,” and then to select models for Phase 2, using predictive margins for scenarios and target feasibility analysis. The deforestation risk map in Figure 22 is a weighted overlay of the following reclassified rasters: tree cover loss, which is given 50% weight; inverse distance weight (IDW)-interpolated rasters of underlying economic conditions from Table 9 M4 regressions and their subsequent sensitivity maps, given 10% weight; and immediate socioeconomic conditions variables from Table 9 M4 regressions, given 5-10% weight depending on magnitude and significance. The reclassifications were based on a 1-9 scale, separating the values of the sensitivity regression coefficients and interpolation values into integer values 1-9 in equal breaks. The layers with these values are then weighted as previously discussed, with actual historical tree cover loss as primary (50%), and overlaid using ArcGIS’s Weighted Overlay tool. It is evident from this mapping that the districts in the province of Quảng Trị, including Hướng Hóa, Đa Krông, and Cam Lộ, and other districts such as A Lưới, Minh Hóa, Hương Khê, and Đắk Glei display notable deforestation risk based on history of tree cover loss, sensitivity to various underlying economic conditions, and recent values for immediate socioeconomic conditions.

Next, Table 13 illustrates the results of predictive margins for scenarios for tree cover loss in four areas: Chinese economic growth, rubber price fluctuations, poverty rate changes, and educational attainment. These are to be cautiously interpreted as the models used had already low R^2 values for any kind of predictive modeling. These methods were also used to fix the most recent economic conditions and thus estimate approximate expected minimum deforestation rate assuming that these economic conditions continue. This can assist WWF in determining the feasibility of targets for reducing deforestation. These results are discussed further in the following section, Phase 2.

Figure 9. Scatterplot matrices for selected immediate determinants of deforestation variables. Variables on the vertical axes (left-hand side) are the Y-axis variables for each row. The purple line on each scatterplot shows a linear correlation fit, the magnitude and significance of the slope is the number above each plot (e.g., 564.298**). The blue line is a locally weighted scatterplot smoothing (LOWESS) which illustrates the nonlinearities in the relationships between each variable, which can also explain the $<0.5 R^2$ values in later regressions. Variables along the bottom are the x-axis variables for each column. Histograms are the distribution of each variable, which can also be helpful for determining which variables require skew-correction to obtain a roughly normal distribution to satisfy Gauss-Markov assumptions for OLS regression. The top row of correlations thus shows the bivariate correlations between forest cover loss and existing forest cover, population density, nighttime lights, secondary school graduation rate, poverty rate, and district contribution to GDP.

Figure 10. Scatterplot matrices for selected biophysical constraints on deforestation. Variables on the vertical axes (left-hand side) are the Y-axis variables for each row. The purple line on each scatterplot shows a linear correlation fit, the magnitude and significance of the slope is the number above each plot (e.g., 15368.455**). The blue line is a locally weighted scatterplot smoothing (LOWESS) which illustrates the nonlinearities in the relationships between each variable, which can also explain the $<0.5 R^2$ values in later regressions. Variables along the bottom are the x-axis variables for each column. Histograms are the distribution of each variable, which can also be helpful for determining which variables require skew-correction to obtain a roughly normal distribution to satisfy Gauss-Markov assumptions for OLS regression. The top row of correlations thus shows the bivariate correlations between forest cover loss and precipitation, temperature, slope, elevation, irrigated cropland, and mosaic cropland. Comparison of the irrigated cropland and mosaic cropland shows no association between forest cover loss and irrigated cropland, but a significant correlation between forest cover loss and mosaic cropland.

Coverage and univariate local Moran's I: Precipitation

Figure 11. Mean coverage map by district and univariate local Moran's I of precipitation in CAL. Displays significant spatial autocorrelation (Moran's I: 0.3416) and so a spatially lagged precipitation variable was constructed.

Univ. Moran's I of:
Temp.

LISA Cluster Map
 Not Significant (71)
 High-High (15)
 Low-Low (3)
 Low-High (0)
 High-Low (2)
 Neighborless (2)

LISA Significance Map
 Not Significant (71)
 p = 0.05 (12)
 p = 0.01 (1)
 p = 0.001 (7)
 Neighborless (2)

Slope

LISA Cluster Map
 Not Significant (68)
 High-High (15)
 Low-Low (6)
 Low-High (1)
 High-Low (1)
 Neighborless (2)

LISA Significance Map
 Not Significant (68)
 p = 0.05 (11)
 p = 0.01 (4)
 p = 0.001 (8)
 Neighborless (2)

Elevation

LISA Cluster Map
 Not Significant (67)
 High-High (17)
 Low-Low (7)
 Low-High (0)
 High-Low (0)
 Neighborless (2)

LISA Significance Map
 Not Significant (67)
 p = 0.05 (11)
 p = 0.01 (6)
 p = 0.001 (7)
 Neighborless (2)

Figure 12. Univariate local Moran's I of temperature, slope, and elevation in CAL. All display significant spatial autocorrelation (Moran's I: 0.2297, 0.4411, 0.7346, respectively) and so a spatially lagged variables were constructed.

Coverage and univariate local Moran's I: Forest cover loss (mean, for construction of spatial lag)

Figure 13. Mean coverage map by district and univariate local Moran's I of forest cover loss in CAL. Averaged annually to construct the spatial weight. Displays significant spatial autocorrelation (Moran's I: 0.354) and so a spatially lagged precipitation variable was constructed for spatially autoregressive modeling.

Provincial-level poverty rate

Provincial-level proportion of households with electricity access

Figure 14. Mean coverage maps at the provincial level for poverty rates and proportion of households with electricity access. Used as categorical variable rather than continuous in modeling due to provincial level grouping by GSOV rather than district level .

Coverage and univariate local Moran's I: Population density

Figure 15. Mean coverage map by district and univariate local Moran's I of population density in CAL. Displays significant spatial autocorrelation (Moran's I: 0.4102) and so a spatially lagged population density variable was constructed.

Univariate local Moran's I: District contribution to GDP

Univariate local Moran's I: nighttime lights

Figure 16. Mean coverage maps by district and univariate local Moran's I of human development variables, district contribution to national GDP and nighttime lights, in CAL. Both display significant spatial autocorrelation (Moran's I: 0.485, 0.608, respectively) and so spatially lagged variables were constructed.

Bivariate local Moran's I: poverty rate and lagged forest cover loss

Bivariate local Moran's I: population density and lagged forest cover loss

Figure 17. Bivariate local Moran's I of poverty rate and population density relative to lagged forest cover loss (essentially, the “neighborhood” of forest cover loss), in CAL. While high poverty rates are spatially associated with forest cover loss (0.1684), population density is not (-0.2661). This provides some evidence that population pressures in the forest frontier areas may not be driving deforestation, but poverty and expansion of subsistence agriculture are, from a spatial standpoint, more likely determinants.

Bivariate local Moran's I: nighttime lights and lagged forest cover loss

Bivariate local Moran's I: irrigated cropland and lagged forest cover loss

Figure 18. Bivariate local Moran's I of nighttime lights and irrigated cropland relative to lagged forest cover loss (essentially, the “neighborhood” of forest cover loss), in CAL. Neither is spatially associated with forest cover loss (-0.3452, -0.1392, respectively). From Figure 8 and evidence from the univariate local Moran's I in previous figures, the hotspot areas of deforestation are not spatially collocated with population-dense, heavily economic areas. This figure illustrates that they are neither potentially collocated with intense irrigated agriculture, suggesting that (and corroborating Figure 8) higher poverty rates and smaller scale subsistence agriculture may be driving deforestation at the immediate scale.

Bivariate local Moran's I: mosaic cropland and lagged forest cover loss

Bivariate local Moran's I: rainfed cropland and lagged forest cover loss

Figure 19. Bivariate local Moran's I of mosaic cropland and rainfed cropland relative to lagged forest cover loss (essentially, the "neighborhood" of forest cover loss), in CAL. Mosaic cropland is spatially associated with forest cover loss (0.3033). Rainfed cropland is not spatially associated with forest cover loss but is less dispersed than irrigated cropland relative to forest cover loss in Figure 9 (-0.055 compared to -0.139).

Table 4. Results of initial specification assessing the role only of biophysical control variables. Forest cover loss is the response variable. Spatial lags of slope and temperature are excluded due to collinearity with other biophysical lags. ***p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses (clustered to district).

Tree cover loss	M1 Log-Linear	M2 Log-log	M2 Linear, unlagged	M3 Linear, lagged
<i>Elevation</i>	-0.0000664 (0.0003499)	0.4358527 (0.3089644)	-61.48903 (994.8899)	564.8219 (3202.348)
<i>Elevation (spatial lag)</i>	-0.0003219 (0.0003549)	-0.0007396 (0.0007395)		-675.0184 (2251.421)
<i>Slope</i>	-0.1505185*** (0.0193649)	-0.6848749** (0.3196085)	-143393.2** (68268.06)	136226.5 (144568.4)
<i>Precipitation</i>	-0.0022168** (0.0010802)	-0.364951 (0.2822432)	-12917.43*** (2729.952)	-15838.73** (4485.264)
<i>Precipitation (spatial lag)</i>	0.0018992 (0.0019875)	0.0062531 (0.0037532)		10683.24 (10280.5)
<i>Temperature</i>	0.1395395*** (0.0308188)	2.761337** (1.36159)	344775.5*** (79448.26)	368600.3** (150707.5)
<i>Existing forest cover</i>	0.0001128*** (0.0000143)	0.5519738*** (0.095402)	305.9007*** (65.50001)	305.54** (137.344)
<i>Year</i>	0.1540747*** (0.0114614)	0.1599431 (0.01125)	566752.9*** (44585)	572376.1*** (62757.12)
<i>Proportion in protected area</i>	-0.0004213** (0.0001436)	-0.0001883 (0.0002434)	-1106.464* (627.2555)	-932.3636 (1065.383)
<i>Constant</i>	-299.4455*** (22.90498)	-320.0846*** (20.61693)	-1.14e+09*** (8.93e+07)	-1.16e+09*** (1.25e+08)
<i>n</i>	680	653	680	680
<i>R-sq</i>	0.3809	0.5335	0.3033	0.3064

Table 5. Results of initial specification assessing the role types of agricultural croplands. Tree cover loss is the response variable. Spatial lags of slope and temperature are excluded due to collinearity with other biophysical lags. *** p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses (clustered to district).

Tree cover loss	M1 Linear	M2 Linear	M3 Log-log	M4 Log-log
<i>Irrigated cropland</i>	-833.3261 (906.9169)	2077.859 (1899.037)	0.201797*** (0.0556102)	-0.0183835 (0.0588528)
<i>Irrigated cropland (spatial lag)</i>	-9644.399*** (1377.468)	2453.251 (2151.895)	-0.000017 (0.0003011)	-0.0030648*** (0.0004495)
<i>Mosaic cropland</i>	877.1*** (230.8801)	1064.529** (371.0898)	0.6772165*** (0.0374728)	0.4813517*** (0.0523451)
<i>Rainfed cropland</i>	945.9721** (319.3324)	201.0895 (772.5546)	0.3115648*** (0.0708857)	0.0315687 (0.0737679)
<i>Year</i>	563692.1*** (42369.38)	558299*** (55657.13)	0.1547289*** (0.0084733)	0.1612482*** (0.0095476)
<i>Elevation</i>	2023.826 (1733.584)			0.0005811** (0.000315)
<i>Elevation (spatial lag)</i>	-6371.642*** (1270.518)			-0.0025902*** (0.0003977)
<i>Slope</i>	303864.9*** (65659.05)			0.1181852*** (0.018947)
<i>Precipitation</i>	-15802.01*** (3180.814)			-0.0033265*** (0.0008038)
<i>Precipitation (spatial lag)</i>	6710.539 (7024.947)			-0.0010668 (0.0018185)
<i>Temperature</i>	582777.2*** (106160.1)			0.1334604*** (0.0373387)
<i>Existing forest cover</i>	187.5387** (64.55037)	297.2574** (95.79903)	0.0000329** (0.0000119)	0.0000303** (0.0000146)
<i>Proportion in protected area</i>	-2600.75*** (737.3091)	1795.763 (1412.523)	-0.0000232 (0.000146)	-0.0005778*** (0.0001547)
<i>Constant</i>	-1.14e+09*** (8.49e+07)	1.12e+09*** (1.11e+08)	-299.7585*** (16.98636)	-314.0616*** (18.92626)
<i>n</i>	680	757	718	656
<i>R-sq</i>	0.3921	0.3437	0.5531	0.6000

Table 6. Results of the Underlying Causes model, for macroeconomic conditions that may drive deforestation. Forest cover loss is the response variable. Square brackets indicate temporal lag (e.g., [L1] = one year time lag). Parentheses in variable name indicate natural log or skew correction. M1 and M2 are economies specifications (trade was excluded from M2 as all were collinear with the inclusion of temporal lags), while M3 to M5 are price models. Variables omitted due to collinearity are indicated with “omit,” while blank cells indicate that that variable was not included in the specification due to previously identified collinearity or for the purposes of a given specification. The same values of biophysical variables and control variables between M1 and M2 and between M3 and M5 should be explored in further research. *** p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses (clustered to district). Square-meter values are converted into hectares in the interpretation.

Tree cover loss	M1 NRE with trade	M2 NRE temporal lags, no trade	M3 Prices	M4 Prices and temporal lags	M5 Prices and temporal lags
<i>VN per capita GDP</i>	19994.16** (11262.26)	omit	-816.4136 (874.4272)		
<i>VN per capita GDP [L1]</i>		4951.827 (13162.37)		5610.994*** (707.564)	
<i>VN trade as component GDP</i>	208987.8*** (45761.97)				
<i>China per capita GDP</i>	4264.225 (7058.629)	39629.34*** (5493.504)			
<i>China per capita GDP [L1]</i>		17249.27** (5433.028)			
<i>China trade as component GDP</i>	207368.9 (180703.3)				
<i>Lao per capita GDP</i>	-55682.33** (20717.61)	-11798.15 (9308.188)			
<i>Lao per capita GDP [L1]</i>		35640.25*** (8916.897)			
<i>Lao trade as component GDP</i>	318663.6** (96782.06)				
<i>Cambodia per capita GDP</i>	20311.78** (8304.602)	omit			
<i>Cambodia per capita GDP [L1]</i>		12010.87 (6702.015)			
<i>Cambodia trade as component GDP</i>	omit				
<i>Thai per capita GDP</i>	2739.385* (1538.123)	7083.602 (2386.173)			
<i>Thai per capita GDP [L1]</i>		2219.646 (1281.933)			
<i>Thai trade as component GDP</i>	51391.73** (20270.9)				
<i>Thai trade as component GDP [L1]</i>	omit				
<i>Myanmar per capita GDP</i>	9773.863 (10102.32)	-45919.89*** (8489.44)			
<i>Myanmar per capita GDP [L1]</i>		omit			
<i>Myanmar trade as component GDP</i>	637335.5*** (144350.7)				
<i>India per capita GDP</i>	10479.59** (3270.766)	61132.85*** (8880.916)			
<i>India per capita GDP [L1]</i>		16821.04** (8180.191)			
<i>India trade as component GDP</i>	omit				
<i>Japan per capita GDP</i>	-7899.611*** (1796.776)	-2277.945** (1017.623)			
<i>Japan per capita GDP [L1]</i>		-889.8979 (572.7266)			
<i>Japan trade as component GDP</i>	omit				

<i>Rice price (ln k)</i>			-43332.92*** (7706.772)		-24900.05*** (3385.241)
<i>Rice price (ln k) [L1]</i>				10280.74** (5020.901)	
<i>Hardwood price index (ln)</i>					497920.3*** (87459.06)
<i>Hardwood price index (ln) [L1]</i>			omit		
<i>Softwood price index</i>				omit	
<i>Softwood price index [L1]</i>					omit
<i>Metals price index (ln k)</i>					164161.6* (83019.28)
<i>Metals price index (ln k) [L1]</i>			omit		omit
<i>Rubber price</i>			27962.46*** (6975.312)		104805.3** (48247.48)
<i>Rubber price [L1]</i>				78992.96*** (11961.03)	
<i>Tea coffee cocoa price index</i>			115584.3*** (21341.72)		-18657.42 (15094.74)
<i>Tea coffee cocoa price index [L1]</i>					omit
<i>Cereals price index</i>			131355.4*** (27119.04)		
<i>Cereals price index [L1]</i>				-15861.78 (14451.38)	
<i>Poverty rate</i>			154409.4 (93337.66)	231338** (74927.3)	154409.4 (93337.66)
<i>Irrigated cropland</i>			4024.161* (2348.825)	2341.974 (2139.536)	4024.161* (2348.825)
<i>Urban area</i>	-8833.604 (16563.66)	-8833.604 (16563.66)	-18960.44 (15022.94)	-9331.049 (15196.83)	-18960.44 (15022.94)
<i>District contribution to GDP</i>	-1378815** (560398.4)	-1378815** (560398.4)			
<i>Elevation</i>	1985.772 (3046.041)	1985.772 (3046.041)		-1909.749 (2107.021)	
<i>Elevation (spatial lag)</i>	-1722.745 (2218.22)	-1722.745 (2218.22)	-1985.333 (2140.133)		-1985.333 (2140.133)
<i>Slope</i>	-57185.41 (145566.2)	-57185.41 (145566.2)	-1985.333 (2140.133)	588075.5** (166730.3)	-1985.333 (2140.133)
<i>Precipitation</i>	1060.3 (5574.843)	1060.3 (5574.843)		19347.79 (15543.34)	
<i>Precipitation (spatial lag)</i>	-2063.274 (10643.47)	-2063.274 (10643.47)	21717 (16820.62)		21717 (16820.62)
<i>Temperature</i>	391184.4** (144237.4)	391184.4** (144237.4)	200748.7 (241258.6)		200748.7 (241258.6)
<i>Existing forest coverage</i>	229.3596* (132.8239)	229.3596* (132.8239)			
<i>Proportion in protected area</i>	-624.7059 (1292.044)	-624.7059 (1292.044)	960.1925 (1589.235)	1089.264 (1491.94)	960.1925 (1589.235)
<i>Constant</i>	1.32e+08** (5.39e+07)	-8.99e+07*** (1.77e+07)	2.01e+07** (7348026)	2.30e+07*** (6010935)	2.01e+07** (7348026)
<i>n</i>	680	680	0.3433	0.2984	0.3433
<i>R-sq</i>	0.3822	0.3823	305	256	305

Table 7. Restricted model for Underlying Causes, for macroeconomic conditions that may drive deforestation. This restricted specification includes key variables identified from Table 6 and is based on the best specification from Tables 4 and 5. This table is thus log-log only, based on the R^2 values from Table 4 M2 and Table 5 M4. Forest cover loss is the response variable. Square brackets indicate temporal lag (e.g., [L1] = one-year time lag). Parentheses in variable name indicate natural log or skew correction. M1 and M2 are economies models, while M3 to M5 are price models. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses (clustered to district).

	M1	M2
Tree cover loss	Log-log	Log-log
No Vietnam GDP		
<i>China GDP</i>	3.650173 (5.465476)	8.817803 (8.091506)
<i>China GDP</i> [L1]	6.953713 (5.577235)	29.67114* (22.42131)
<i>Laos trade</i>	-1.465076* (0.8965466)	-0.3950595 (1.259421)
<i>Myanmar trade</i>	-0.0492431 (0.1620198)	0.7134936 (0.5837525)
<i>Vietnam GDP</i>		1036.421** (511.3178)
<i>Vietnam GDP</i> [L1]		435.4749** (198.9324)
<i>Rubber price</i> [L1]	1.358209** (0.5339335)	56.0753* (28.06919)
<i>Hardwood price index k</i> [L1]	-0.8730011 (2.114281)	3.732241* (3.560287)
<i>Softwood price index</i> [L1]	-0.4684012 (1.346725)	81.67841* (44.19615)
<i>Rice price (k)</i>	-1.154796 (1.124695)	-49.42573** (23.35843)
<i>Metals price index (k)</i>	1.959699 (3.761136)	248.7556* (131.8583)
<i>Cereals price index</i>	-3.214041** (1.006481)	-8.763446*** (2.219172)
<i>Tea coffee cocoa price index</i>	1.127153 (1.223328)	24.76529* (14.21868)
<i>Elevation</i>	-0.4987483 (0.4562561)	-0.473074 (0.4594785)
<i>Elevation (spatial lag)</i>	0.000286 (0.0009187)	0.0002656 (0.0009248)
<i>Slope</i>	0.0577436 (0.1013438)	0.0516856 (0.1020776)
<i>Precipitation</i>	0.2078277 (0.5584836)	0.4964275 (0.6295477)
<i>Precipitation (spatial lag)</i>	0.0051998 (0.0051545)	0.0039969 (0.0052747)
<i>Temperature</i>	2.441207 (1.894591)	2.367465 (1.894285)
<i>Existing forest cover</i>	0.9601489*** (0.1691587)	0.9517192*** (0.1697263)
<i>Proportion in protected area</i>	-0.0003777 (0.0002925)	-0.0003966 (0.0002895)
<i>Year</i>	0.6005714 (2.691594)	9.392731** (4.634885)
<i>Constant</i>	18.21104 (31.20094)	18050.69** (8868.397)
<i>n</i>	653	653
<i>R-sq</i>	0.6214	0.6240

Table 8. Results of the Immediate Causes model, for micro and socioeconomic conditions that may drive deforestation. Forest cover loss is the response variable. Fixed effects were selected over random effects following a Hausmann test ($p=0.0087$) without clustered standard errors and a Sargan-Hansen test with clustered standard errors. Variables omitted due to collinearity are indicated with “omit,” while blank cells indicate that that variable was not included in the specification due to previously identified collinearity or for the purposes of a given specification. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses (clustered to district).

Tree cover loss	M1	M2 Temporal lags	M3 (No FE)	M4 (No FE)
<i>Population density (spatial lag)</i>	-35341.38 (33808)	-29057.39 (34522.63)	-363.88 (673.7428)	-801.613 (1047.978)
<i>Poverty rate (categorical)</i>	365447.3* (45168.57)			
<i>District contribution to GDP [L1]</i>	Omit		172395.4*** (529051)	1044925*** (233935.9)
<i>Secondary school graduation ratio</i>	-126960.5* (74337.03)			
<i>Secondary school graduation ratio [L1]</i>		-36425.6 (31534.99)	-29339.55*** (21921.96)	-124201.5*** (23544.69)
<i>Irrigated cropland</i>	86690 (126354.4)			
<i>Irrigated cropland [L1]</i>		221580.3 (152534.1)	475.4306 (921.3167)	1860.187 (1317.673)
<i>Mosaic cropland</i>	22678.12** (7209.053)			
<i>Mosaic cropland [L1]</i>		29060.38** (9003.066)	2036.557** (579.7422)	2142.061** (609.2791)
<i>Rainfed cropland</i>	16745.82 (44247.01)			
<i>Rainfed cropland [L1]</i>		60999.82* (34972.42)	370.7381*** (326.2099)	237.5954*** (300.1541)
<i>Nighttime lights (spatial lag) [L1]</i>	Omit	Omit	-5133.074 (49774.15)	260478.3 (194932.1)
<i>Precipitation (spatial lag)</i>	Omit	Omit	5620.537* (4219.364)	3653.445 (8258.201)
<i>Temperature (spatial lag)</i>	1664024* (758860)	494880 (577849.2)	56545.33* (139695.1)	79941.23 (196275.9)
<i>Constant</i>	-7.52e+07 (5.29e+07)	-1.20e+07 (4.60e+07)	3024997 (4058081)	-3856935 (4769536)
<i>n</i>	143	143	143	437
<i>Within R-sq</i>	0.2994	0.2885		
<i>R-sq</i>			0.2892	0.3413

Table 9. Results of a combined restricted model, using key variables identified from the results of both underlying and immediate determinants models. Economies variables and price variables are not included in the same specifications. Forest cover loss is the response variable. For M3/M4, fixed effects were selected over random effects following a Hausmann test ($p=0.0087$) without clustered standard errors and a Sargan-Hansen test with clustered standard errors. Time-invariant variables (such as elevation and slope) are excluded. Low statistical significance is expected in M3/M4 due to low sample size for FE. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses (clustered to district).

Tree cover loss	M1 Log-log	M2 Log-log	M3 FE Log-log	M4 FE Log-log
<i>China GDP</i>	489.8559* (261.4354)		1.913143 (2.57554)	
<i>China GDP [LI]</i>	737.2273* (360.8071)			36.18231* (24.44091)
<i>Vietnam GDP</i>	171.5475** (49.55133)			
<i>Vietnam GDP [LI]</i>	891.0766** (379.5819)			72.89299* (44.33401)
<i>Laos trade</i>				0.092908 (1.501932)
<i>Myanmar trade</i>				0.5858476 (0.3869134)
<i>Rubber price [LI]</i>		2.3599 (1.745511)		1.715393** (1.802417)
<i>Metals price k [LI]</i>		-0.2207012 (8.775313)		31.3177* (15.15001)
<i>Hardwood price [LI]</i>			-4.514747 (3.431119)	1.13195* (3.033319)
<i>Softwood price [LI]</i>			11.81184 (9.27918)	15.26349 (8.157866)
<i>Cereals price</i>		-3.962368** (1.575998)	-0.3931254 (2.659267)	-8.977499** (3.434027)
<i>Tea coffee cocoa index</i>		0.2075338 (1.418538)	0.9087376 (2.18815)	6.594833* (3.267695)
<i>Irrigated</i>	0.2429514 (0.3421191)	0.2434128 (0.3387617)	0.589293 (2.057668)	-0.374122 (1.513447)
<i>Mosaic</i>	0.3091218** (0.136097)	0.309942** (0.1352856)	2.646207 (3.008562)	0.9138157* (0.9640797)
<i>Rainfed</i>	-0.0363915 (0.4653968)	-0.0532371 (0.4200218)	-7.335116 (6.236651)	2.822467** (5.015623)
<i>Poverty rate (categorical)</i>	-0.0072743 (0.0308634)	-0.0062369 (0.0295392)	-0.0015723 (0.0991845)	
<i>Secondary school graduation ratio</i>	-0.0818358** (0.030287)	-0.0812551** (0.0320565)	-0.077436 (0.0573953)	-0.9262519* (0.9452198)
<i>Secondary school graduation ratio [LI]</i>	0.0157979 (0.0184579)	0.015478 (0.0182308)		
<i>Urban</i>	0.0015461 (0.0102265)	0.0488753 (0.0954481)	-0.05189811** (0.1880017)	
<i>Urban (spatial lag)</i>	0.029458 (0.1767216)			
<i>Elevation</i>	-0.35525 (0.5910799)	-0.3574898 (0.5869297)		
<i>Elevation (spatial lag)</i>	-0.0002213 (0.0011455)	-0.0002194 (0.0011418)		
<i>Slope</i>	-0.0086958 (0.1161231)	-0.0094793 (0.1157413)		
<i>Precipitation</i>	-0.5999727 (0.7764093)	-0.580415 (0.7867241)	-0.0775087 (1.635832)	0.0007873 (0.0027666)
<i>Precipitation (spatial lag)</i>	0.0083723 (0.0072913)	0.008261 (0.0072989)		
<i>Temperature</i>	1.678452 (2.905509)	1.694951 (2.872866)	35.1846 (24.59051)	0.1535764 (0.1467935)

<i>Existing forest cover</i>	0.9033585*** (0.1816531)	0.9023991*** (0.1828198)	5.043829 (9.512993)	
<i>Proportion in protected area</i>	-0.0001503 (0.000421)	-0.0001115 (0.0002858)		
<i>Year</i>	18.08574** (9.887542)	0.2481379** (0.0970283)		
<i>Constant</i>	-32559.37* (17941.68)	-480.3277** (179.8624)	-148.1936 (131.7449)	73.55838 (69.05033)
<i>n</i>	162	162	162	162
<i>Within R-sq</i>			0.8340	0.8390
<i>R-sq</i>	0.7871	0.7870		

Table 10. Results for spatially unlagged (spatial) regression specifications combining underlying and immediate determinants. This was an exploratory specification to observe variables prior to adjustment for spatial autocorrelation and to observe combined effect of global prices and district level immediate determinants of deforestation. These results are strictly correlative as coefficients are biased by spatial autocorrelation and important biophysical variables are not included. *** p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses (clustered to district).

<i>Tree cover loss (ln)</i>	<i>M1</i>	<i>M2</i>	<i>M3</i>
<i>Hardwood price index (ln) [L1]</i>	3.50893* (1.808863)	3.678235* (1.794384)	3.917797* (3.663841)
<i>Rice price index (ln k) [L1]</i>	2.06058** (0.7598934)	2.11341** (0.7572451)	-0.7726332 (2.003817)
<i>Metals price index (ln k) [L1]</i>	11.27949** (4.407276)	12.22363** (3.79163)	-3.154828 (3.734816)
<i>Year</i>	0.1723919* (0.0897471)	0.189211** (0.0758756)	.2633807** (0.1072396)
<i>Poverty rate</i>	0.010297 (0.044294)	0.0070413 (0.0351854)	0.0378121 (0.041743)
<i>Secondary school graduation ratio</i>	0.0314856 (0.0226285)	0.0329211 (0.0198737)	-0.073652* (0.0395249)
<i>Irrigated cropland</i>		0.026733** (0.0069316)	0.0188936 (0.0104945)
<i>Urban area</i>			0.1292283* (0.2613022)
<i>Constant</i>	-400.7552* (176.4774)	-447.3243** (142.8636)	-485.9** (192.6157)
<i>n</i>	499	499	381
<i>R-sq</i>	0.8963	0.9019	0.8757
<i>Adj R-sq</i>	0.8728	0.8794	0.8472

Table 11. Based on the model and results of Table 9, we ran a district-specific model, using a Tier 2/district factor variable (interaction) for each key explanatory variable identified in Table 9. The corresponding maps for each district's deforestation “sensitivity” to the economic variable (the magnitude of the interaction coefficient) are shown in Figures 20 and 21. Districts with insufficient data for estimation are not included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses (clustered to district).

Tree cover loss	Rainfed (%)	Mosaic (%)	Hardwood price (%)	Rubber price (%)	Tea coffee cocoa index (%)	China GDP (%)	Vietnam GDP (%)	District contribution to GDP (sq m)	Secondary School Grad. (%)
<i>A Lưới</i>	-0.0055296	0.0096012	4.977562**	1.386006*	2.312154**	2.007103*	3.183553*	28.55903	-0.1114018*
<i>Ba Tơ</i>	-0.0257917	-0.0005876	4.276189**	1.467484**	2.257592**	1.826458**	2.776632*	193.5583	-0.1028075**
<i>Bắc Trà My</i>	-0.5342821	-0.0035121	3.87703*	-0.9077472	1.665313*	1.575744	2.698512	742.9277	-0.1008695**
<i>Bố Trạch</i>	0.0245919	0.0067656	-0.2134797	0.0089485	0.0646617	-0.0006023	0.086852	473.5828*	-0.1567863
<i>Cam Lộ</i>	0.0116437	-0.0024099	6.916393**	1.748169**	3.103002**	2.902379**	4.897764**	23.08629	-0.2263029**
<i>Cẩm Xuyên</i>	0.030055*	-0.0050792	6.1229**	1.974133**	3.10753**	2.746997**	4.525864**	87.14736	-0.1744113**
<i>Duy Xuyên</i>	0.0200904	0.0548045**	4.902058**	2.123358**	2.92805**	2.306641**	3.531889*	56.38022	-0.1646367***
<i>Giáo Linh</i>	-0.0063118	-0.0022909	5.887006**	1.541942**	2.67131**	2.499903**	4.217241**	2.450015	-0.2004655**
<i>Hòa Vang</i>	0.0989841	0.0618727**	10.14438**	3.422376***	5.334731***	4.584775**	7.322452**	49.70241	-0.3837742**
<i>Hương Khê</i>	0.0132966**	0.0039997	4.667627*	1.281541**	2.185621**	1.950634*	3.22476*	145.2512	-0.1314555**
<i>Hương Sơn</i>	0.0032959	-0.0102035	1.568203	0.157121	0.4088468	-0.4640666	1.140587	148.5163	-0.0331908
<i>Hương Thủy</i>	0.0429754	0.0064346	8.269545**	2.192193**	3.748395**	3.485467**	5.92732**	26.31392	-0.1988761*
<i>Hương Trà</i>	0.0625429***	0.0019763	8.33034**	2.363069**	3.90498**	3.595275**	6.068822**	53.14526	-0.2217568**
<i>Huỳnh Hóa</i>	0.0055573	0.0014006	-3.025777	-0.8126672	1.379818	1.161165	1.809828	158.1133	-0.1212161**
<i>Hải Lăng</i>	0.0192195	0.00996**	7.605939**	2.191699**	3.661127**	3.355265**	5.544565**	24.31438	-0.2273371**
<i>Hội An</i>	0.5839337**	2.553211	-0.208743	-0.5523133	-0.5842502	-0.0908026	0.1996868	24.53398	0.0308012
<i>Kon Plông</i>	0.0294619	0.0010538	0.4430092	0.3370682	0.3833162	0.2792036	0.35293	2123.62	-0.0151831
<i>Kon Rẫy</i>	0.0182243	-0.0034478	0.3688323	0.0696703	0.1913796	0.1810036	0.2437314	665.3592	-0.0162974
<i>Kỳ Anh</i>	0.0096044**	0.0002692	-3.100152	1.156326*	-1.654653	1.242635	1.86354	33.06557	-0.1111618**
<i>Lệ Thủy</i>	0.0137382*	0.0019029	-3.226262	-0.765853	-1.420954	1.280253	2.100264	31.5604	-0.1103004**
<i>Minh Hóa</i>	0.0387019	0.016605	-2.684521	0.7494212	1.269005	1.007958	1.511628	-117.6445	-0.076188
<i>Nam Giang</i>	0.0167828	0.0074506	-0.4520235	0.0304588	0.1230805	-0.0885994	-0.1671158	599.3738	-0.064119**
<i>Nam Trà My</i>	-0.8361193	-0.0161655	-0.9741126	0.1075491	-0.2567861	-0.3087508	-0.5654861	-14988.56	-0.0257408
<i>Nam Đàn</i>	0.0527135*	0.0456715***	-1.789889	-0.8488683	1.137005	-0.5843881	0.6822095	17.57124	-0.1388389**
<i>Nam Đông</i>	0.1820495	0.0216919	-5.243133*	1.393871*	2.338915*	2.161367*	3.689293*	-1863.329	-0.0788165
<i>Ngọc Hồi</i>	0.0101192	-0.0020768	0.0151596	0.4063518*	0.2466227	-0.0085732	-0.1771287	-44.51271	-0.0426157
<i>Núi Thành</i>	0.0165594	0.0134996	6.600831**	2.446782***	3.616746**	2.986676**	4.604296**	8.803865	-0.2372313***
<i>Phong Điền</i>	-0.0045976	0.0108339	5.812997**	1.582901**	2.697788**	2.385723**	3.837162*	-123.0854	-0.1634875**
<i>Phủ Lộc</i>	0.0518513*	-0.0345527	6.640637**	2.030104**	3.230344**	-2.955888**	5.006397*	-11.39038	-0.1909949**
<i>Phủ Ninh</i>	0.0326912	0.0022439	5.209235**	1.820589**	2.74024**	2.291651**	3.669184*	45.29143	-0.1612004**
<i>Phủ Vang</i>	0.0365082	-0.5168058	7.328402**	2.101768**	3.438737**	3.38682**	5.947778**	15.59231	-0.1705146
<i>Phước Sơn</i>	-0.8793152	0.003829	-0.318337	0.2788709	0.2250521	0.0834867	0.1322972	-657.0093	-0.0421839
<i>Quảng Ninh</i>	0.036021	0.0017343	4.501927*	1.418886**	2.219359**	1.915962*	-3.181481*	274.01*	-0.1162054**
<i>Quảng Điền</i>	0.1180377**	1.217244*	4.575082*	1.303342*	2.168377*	1.795778	-2.825321	-10.17178	-0.0938405
<i>Tam Kỳ</i>	0.477972***	2.079787**	0.0935248	0.4008688	0.3292605	0.7914173	2.218059	17.22469	-0.0158529
<i>Thanh Chương</i>	0.0074207	0.0032014	-0.9860427	0.2335211	-0.4869513	0.3613476	-0.5415998	134.9077*	-0.1264625
<i>Triệu Phong</i>	0.0003148	0.0171844**	7.430785**	2.071621**	3.523442**	3.255696**	-5.383324*	9.120539	-0.2533305**
<i>Trà Bồng</i>	0.0254043	-0.0142089	5.832295**	1.927306**	3.040016**	2.557657**	-4.002658*	223.5002	-0.1150367**
<i>Tu Mơ Rông</i>	0.0316293	0.007237	0.38809	0.4364942	0.3618406	0.1910884	0.3126439	13405.95	-0.0325012
<i>Tuyên Hóa</i>	0.018086	0.0150341**	2.965774	0.9087773	1.451652	-1.193121	-0.956359	488.7762	-0.0588759
<i>Tây Giang</i>	0.1962391	0.0102519	3.40432	0.6702145	1.358493	-1.316076	-2.211456	omit	-0.0730449
<i>Vĩnh Linh</i>	0.024611*	0.0079336**	3.256841	0.5551889	1.24647	-1.140753	-1.946332	omit	-0.1202072**
<i>Vũ Quang</i>	0.0217151*	0.0459206	-1.916495	0.0272266	-0.3296068	-0.5435466	-1.444926	omit	-0.0192984
<i>Đa Krông</i>	0.0660854*	0.0007779	5.675494**	1.67747**	2.732148**	2.414359**	3.916356**	omit	-0.1703747**
<i>Đông Giang</i>	0.0379407	0.0111161	4.956894**	1.283122*	2.284811**	2.056754**	3.348268*	omit	-0.0779507
<i>Đại Lộc</i>	-0.0255404	-0.0211678	4.334122*	1.40782**	2.228314**	1.860935*	-2.92747*	omit	-0.1152837**
<i>Đắk Glei</i>	0.0154591	0.0000156	1.07374	0.8685195**	0.9009062*	0.6315844	0.9617059	omit	-0.0182167
<i>Đắk Hà</i>	-0.0046769	0.0048006*	omit	omit	omit	omit	omit	omit	-0.0439481
<i>Irrigated</i>	2.719056 (1.660667)	0.7725651 (0.7063274)	0.6163446 (0.5126875)	0.448398 (0.4640008)	0.5591077 (0.4878958)	0.6283461 (0.5080824)	0.6678265 (0.5257782)	0.523072 (0.4471357)	-0.1321531 (0.3992986)
<i>Mosaic</i>	3.509372** (1.60156)		0.1031849 (1.573181)	0.5220679 (0.678263)	0.4009773 (1.056906)	0.2536609 (1.557759)	0.0894109 (1.825654)	3.880169** (1.540845)	0.8950758 (1.29343)
<i>Rainfed</i>		1.528125 (3.555957)	-0.3181875 (2.073881)	-0.3814398 (0.9553958)	-0.3894117 (1.324197)	-0.188203 (1.861663)	-0.8126256 (2.46303)	2.126643 (3.608103)	0.574893484 (1.225673)
<i>Secondary school grad.</i>	-0.0306245* (0.0177873)	-0.0242876 (0.0180318)	-0.0449089** (0.0185132)	-0.0479459** (0.0156558)	-0.0470741** (0.0165057)	-0.0450352** (0.0178518)	-0.0438986** (0.0191826)	-0.0279643* (0.0170398)	
<i>Population density</i>	-0.011954 (0.0292273)	0.0396281** (0.0125932)	-0.0269304** (0.0125729)	-0.0075623* (0.0042774)	-0.0131426* (0.0068694)	-0.0138952 (0.0181521)	-0.0156212 (0.0231612)	-0.0030602 (0.0422094)	-0.0251287 (0.0193568)
<i>Nighttime lights (spatial lag)</i>	23.64441*** (5.596949)	5.647756** (2.643698)	2.91764*** (0.7293887)	1.070615*** (0.1926958)	1.690228*** (0.3312447)	2.953666** (1.106924)	4.669294** (1.883518)	-6.302542 (11.3263)	2.317237** (1.109345)

<i>Elevation</i>	-0.0368827 (0.0405967)	0.0213864 (0.0127983)	-0.0037974 (0.0093095)	0.0005996 (0.0028295)	0.0000786 (0.00436)	-0.0020954 (0.007428)	-0.0075413 (0.0126273)	-0.0437236 (0.090098)	0.0039571 (0.0062757)
<i>Elevation (spatial lag)</i>	-0.0932094 (0.1173254)	-0.0109688 (0.0109147)	-0.0195508 (0.014866)	-0.0064216* (0.0038125)	-0.0099303 (0.0064241)	-0.0137823 (0.0118256)	-0.0179982 (0.0204799)	-0.0087013 (0.0213211)	-0.0159079** (0.0061951)
<i>Slope</i>	10.61469 (11.15159)	0.9953401 (1.468462)	0.5175755 (1.07029)	0.2000193 (0.3751836)	0.3023531 (0.5809185)	0.270537 (0.9471347)	0.1937308 (1.461391)	3.908348 (9.977752)	-0.4052213 (0.7410893)
<i>Precipitation</i>	-0.0155842** (0.0067303)	-0.0200244** (0.0078164)	0.0104986 (0.0101843)	0.0138* (0.0077081)	0.0124873 (0.0085538)	0.0101261 (0.0099209)	0.008527 (0.0108)	-0.0159352** (0.0055665)	-0.009196 (0.0083586)
<i>Precipitation (spatial lag)</i>	-0.5897012 (0.5901891)	-0.2183445* (0.1183169)	0.1741821** (0.0842577)	0.0343186 (0.0250974)	0.0753112 (0.0403274)	0.1355953* (0.0676566)	0.2077783* (0.1096358)	-0.0469549 (0.6619679)	0.0016738 (0.075274)
<i>Temperature</i>	2.988588** (0.7726935)	2.073584** (1.018468)	2.225843** (0.6289743)	1.047171** (0.3267053)	1.59009 (0.4528153)	2.1272** (0.586543)	2.373835** (0.6865528)	1.974537** (0.8140824)	1.940586** (0.6546358)
<i>Existing forest cover</i>	0.0015395 (0.001075)	0.0009489 (0.0012188)	0.0003783 (0.000694)	0.0000802 (0.0002188)	0.0001623 (0.0003469)	0.0003077 (0.0005896)	0.0005037 (0.0008996)	0.0022648 (0.0013692)	0.0000737 (0.0004568)
<i>Proportion in protected area</i>	-0.0640856 (0.0601072)	-0.0107051 (0.013841)	-0.0060059 (0.0049018)	-0.0018672 (0.0018355)	-0.00294 (0.0027956)	-0.0046465 (0.004416)	-0.0072628 (0.0065574)	-0.0446056 (0.0370968)	-0.001438 (0.0071636)
<i>Year</i>	0.4475198** (0.1485951)	0.3241106** (0.1574964)	-0.1238075 (0.1378149)	0.033718 (0.0930576)	0.0363495 (0.105845)	-0.1284366 (0.1306298)	-0.1718987 (0.1516494)	0.2812051** (0.1411197)	-0.0941613 (0.109981)
<i>Constant</i>	1054.79** (376.5401)	650.2258** (288.4237)	205.2785 (259.2445)	-80.53449 (182.953)	45.61251 (203.8758)	217.5041 (245.8418)	302.8241 (280.8667)	599.2046** (284.0896)	175.1894 (216.0337)
<i>n</i>	95	95	95	95	95	95	95	95	95
<i>R-sq</i>	0.8599	0.8535	0.8700	0.8711	0.8715	0.8709	0.8696	0.8411	0.8723

Table 12. Based on the model and results of Table 9, we ran a province-specific model, using a Tier 1/province factor variable (interaction) for poverty rate. This done at the Tier 1 level, rather than Tier 2 (as in Table 11) because the poverty rate data was available only at the province level. The corresponding map for each province's deforestation "sensitivity" to poverty rate (the magnitude of the interaction coefficient) is shown in Figure 20, bottom left. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses (clustered to province).

Tree cover loss	Poverty rate
<i>Hà Tĩnh</i>	1.007326
<i>Kon Tum</i>	0.3969085
<i>Nghệ An</i>	0.5396337*
<i>Quảng Bình</i>	0.7059566
<i>Quảng Nam</i>	0.2146395
<i>Quảng Ngãi</i>	0.7935853
<i>Quảng Trị</i>	0.5407757
<i>Thừa Thiên Huế</i>	omit
<i>Đà Nẵng</i>	omit
<i>Irrigated</i>	0.07268 (0.3792964)
<i>Mosaic</i>	0.367212 (0.2304697)
<i>Rainfed</i>	-0.0238662 (0.4537143)
<i>Secondary school grad.</i>	0.1360462 (0.5836899)
<i>Population density</i>	-0.0018322 (0.0036137)
<i>Nighttime lights (spatial lag)</i>	-0.2268567 (0.1747108)
<i>Elevation</i>	-0.0000388 (0.0021281)
<i>Elevation (spatial lag)</i>	-0.0023419 (0.0027679)
<i>Slope</i>	0.2603684 (0.1692937)
<i>Precipitation</i>	0.0058308 (0.0116442)
<i>Precipitation (spatial lag)</i>	0.0039216 (0.0123261)
<i>Temperature</i>	0.0595969 (0.2602446)
<i>Existing forest cover</i>	5.88e-06 (0.0000784)
<i>Proportion in protected area</i>	-4.21e-07 (0.00006538)
<i>Constant</i>	-3.931278 (49.42009)
<i>n</i>	47
<i>R-sq</i>	0.7337

Figure 20. Sensitivity of deforestation to explanatory variables, based on district-level regression coefficients from Tables 11 and 12. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Generally, the districts most sensitive to economic conditions are located in the provinces of Quảng Bình, Quảng Trị, and Thừa Thiên Huế.

Figure 21. Sensitivity of deforestation to explanatory variables, based on district-level regression coefficients from Tables 11 and 12. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Generally, the districts most sensitive to economic conditions are located in the provinces of Quảng Bình, Quảng Trị, and Thừa Thiên Huế.

Figure 22. Deforestation “risk map” based on a weighted overlay of historical forest cover loss (50% weight), reclassified inverse distance weight (IDW)-interpolated rasters of underlying economic conditions from Table 9 M4 regressions and their subsequent sensitivity maps (10% weight), and immediate socioeconomic conditions variables from Table 9 M4 regressions (5-10% weight depending on magnitude and significance). Districts in the province of Quảng Trị, including Hướng Hóa, Đa Krông, and Cam Lộ, and other districts such as A Lưới, Minh Hóa, Hương Khê, and Đắk Glei display notable deforestation risk based on history of forest cover loss, sensitivity to various underlying economic conditions, and recent values for immediate socioeconomic conditions. Higher values (red) indicate greater “deforestation risk.” The value “1” is not included as this contained “no data” values (white, outside of the extent of the forest cover loss raster).

3.2 Scenarios of economic and socioeconomic conditions

Following the Phase 1 modeling, we selected Chinese economic growth, rubber price fluctuation, poverty reduction, and secondary school graduation for use with predictive margins. We estimated the models presented in Table 13. The specific model used is indicated underneath the equation in each cell, along with the fixed values to predict the possible effect on deforestation for each scenario. Fixed values were selected using the minimum, maximum, and median values for each variable. We followed the most successful model utilizing each variable, prioritizing R^2 values. X is a vector of non-interest variables (the remainder of variables run normally in the regression) for a log-log spatial OLS. ε_{dpt} is the error term for district d , province p , and at time t .

Table 13 provides a simple classification of these results, based on whether the coefficient for that variable's effect on tree cover loss increased, decreased, or remained the same from the coefficient for that variable in the original model (indicated under the equation in each Table 13 cell). **Green** indicates that the given scenario does **is not associated with tree cover loss**. **Red** indicates that the given scenario **is associated with tree cover loss**. **Yellow** indicates that a significant relationship could not be reliably determined with a P-value under 0.10 using predictive margins. The results are in line with the regression results overall, with the exception of increase in rubber price. Surprisingly, fixing the rubber price at its maximum observed value in the dataset reduced the effect of rubber price on tree cover loss, contrary to the results of the other regression models. We recommend further research to explore this inconsistency.

Table 13. Results for estimating forest cover loss using fixed values (predictive margins) for Chinese economic growth, rubber price fluctuation, poverty reduction, and secondary school graduation results for each possible outcome is indicated with green, red, or yellow.

<p>Chinese economic growth</p> $\ln (Loss)_{dt} = \alpha + \beta_1 \ln [X]_{dt} + \beta_2 WX_{dt} + \beta_x \ln [China]_t + \varepsilon_{dt}$ <p>Table 7 M1</p> <p>Possible outcomes:</p> <ol style="list-style-type: none"> 1. Economic contraction (<i>China</i> = 7.983889) 2. Maintained growth (<i>China</i> = 8.912644) 3. Economic boom (<i>China</i> = 9.729535) 	<p>Rubber price fluctuation</p> $\ln (Loss)_{dt} = \alpha + \beta_1 \ln [X]_{dt} + \beta_2 WX_{dt} + \beta_x \ln [Rubber]_t + \varepsilon_{dt}$ <p>Table 7 M2</p> <p>Possible outcomes:</p> <ol style="list-style-type: none"> 1. Drop in price (<i>Rubber</i> = 3.261404) 2. Increase in price (<i>Rubber</i> = 5.386853)
<p>Poverty reduction</p> $Loss_{dpt} = \alpha + \beta_1 X_{dpt} + \beta_2 WX_{dpt} + \beta_x Pov_t + \varepsilon_{dpt}$ <p>Table 6 M4</p> <p>Possible outcomes:</p> <ol style="list-style-type: none"> 1. Worsening poverty (<i>Pov</i> = 2) 2. Maintain current levels (<i>Pov</i> = 3) 3. Continued improvement (<i>Pov</i> = 5) <p>(poverty rate quintiles)</p>	<p>Education level improvement</p> $Loss_{dpt} = \alpha + \beta_1 X_{dpt} + \beta_2 WX_{dpt} + \beta_x Edu_t + \varepsilon_{dpt}$ <p>Table 8 M1</p> <p>Possible outcomes:</p> <ol style="list-style-type: none"> 1. Current secondary graduation rates are maintained (<i>Edu</i> = 91.42638) 2. Graduation rates increase (<i>Edu</i> = 99.86)

We estimated that maintained growth of Chinese GDP may predict increased deforestation, although we were unable to estimate a reliable prediction ($p > 0.10$) for economic contraction or economic boom. In contrast to the previous econometric outputs, we found that increase in rubber price may predict reduced deforestation, but this predictive margin was of borderline significance

($p=0.1000$). Next, a worsening of poverty rate was estimated to predict worsening deforestation, while increase in secondary school graduation rates beyond current levels was estimated to predict reduced deforestation.

3.2.1 Target feasibility analysis

We used predictive margins to further estimate minimum overall tree cover loss across CAL, given the most recent observed values for a number of economic variables. For strategic planning purposes, this can be useful for determining expected deforestation in the absence of interventions. We estimated

$$\ln (Loss)_{dt} = \alpha + \beta_1 \ln [X]_{dt} + \beta_2 WX_{dt} + \beta_x \ln [Z]_t + \varepsilon_{dt}$$

based on Tables 9, 10, and 13, where X is a vector of the same controls included in those models. We selected key economic variables also identified in those models and fixed their values in the vector Z , which included:

- China GDP (16806.742 [PPP] 2017; log transformed) =9.65057
- Vietnam GDP (6775.83219 [PPP] 2017; log transformed) =8.747604
- Rubber price (90.790599 [USD per metric tonne] 2016; log transformed) =4.310578
- Hardwood price index (124.791 [index 2005=100] 2016; log transformed) =4.82664
- Metals price index (126.574 [index 2005=100] 2016; log transformed) = 4.940827
- Cereals price index (148.98 [index 2005=100] 2016; log transformed) =5.003812
- Tea-coffee-cocoa index (172.587 [index 2005=100] 2016; log transformed) = 5.150901
- Secondary school graduation (95.8 [% average] 2017; log transformed) =4.5

Our resulting predictive margin given these fixed values was 7.547895** ($p=0.004$) with an original model R^2 of 0.652, and our 95% confidence interval predicted a range of [2.591826: 12.50396]. Using the conservative end of the range, current market conditions would predict minimum tree cover loss of approximately 2.59% with 95% confidence. This would not be uniform across CAL, and as Figures 20 and 21 illustrate, districts have varying degrees of sensitivity to the economic conditions included in this model. There are also certainly other non-economic conditions, such as changes in local or national forest policy or governance that could greatly impact this result. Further, variation in institutions such as land tenure and property rights can greatly affect land use change in CAL districts.

4. Discussion

In this section, we summarize overall findings from our exploration of socioeconomic determinants of deforestation across Vietnam's Central Annamites Landscape, and tie together the distinct-specific and spatial elements of this pilot study.

First, the data availability assessment is a fundamental component of our study (Appendix II), as it has shaped variables used to perform these analyses. With more temporal and spatial variation in just a few of the variables, much of the constraints resulting in coarse and correlative analysis, required by collinearity, could be avoided. An important next step for future research will be determining the approach for acquiring data that is online but currently inaccessible, such as GSOV household living conditions survey data or VNForest data available for visualization at a fine scale online.

For underlying macroeconomic determinants of deforestation in CAL, we found evidence to suggest that Vietnamese GDP growth and that of neighboring and regional economies, such as China, have been associated with increased tree cover loss in CAL. Overall economic and development pressures are thus underlying drivers of activities that result in deforestation; we did not find evidence to support a deforestation Environmental Kuznets Curve in CAL, in alignment with the Bhattari and Hammig's findings (2001). While we found through spatial analysis that these economic pressures, including those from population-dense and urban areas, are not spatially collocated with the areas of highest tree cover loss, we did find that districts in the provinces of Quảng Bình, Quảng Trị, and Thừa Thiên Huế are sensitive to fluctuations in underlying economic determinants of deforestation, and these districts are strategically located along the transition from urban/economic coastal areas to the western forest frontier (Figures 1, 20, 21).

Our findings suggest that components of contribution to GDP and land conversion taking place within these districts needs to be broken down further to better understand the implications for deforestation. The results from district individual contribution to GDP were not in alignment with those from districts with higher urban area. District contribution to GDP was associated with increased tree cover loss in several specifications, while higher urban area was consistently associated with reduced tree cover loss. This is potentially due to the increased off-farm wages and employment opportunities in urban areas but indicates a need to understand which specific contributions each district makes to national GDP. Future research can thus pinpoint which specific activities within each district are contributing to the tree cover loss associated with that district's contribution to GDP. Considering agricultural contributions to GDP, this could help to further isolate specific commodities driving land use change and tree cover loss (Table 6 M1 and M2; Figure 17, Figure 18).

Next, we used commodity prices that provide an underlying signal for land use change, and therefore can proxy for land use change as an underlying driver of deforestation. Rubber, and hardwood prices had the most consistent and greatest magnitude of effect on tree cover loss, with indices for tea-coffee-cocoa and metals prices also associated with increased tree cover loss where significant. Food commodities such as rice and cereals were inconsistent. This aligns with findings of Cochard et al. (2013) and provides evidence to support current conservation efforts focusing on rubber, acacia, and artisanal mining. However, our finding of spatial collocation of deforestation with poverty and low population density, combined with the possible benefits of improved educational attainment across spatial analysis, economic modeling, and scenario analysis, indicates that immediate socioeconomic factors cannot be forgotten. Returning to the Angelsen and Kaimowitz (1999) conceptual framework, the agent decision-making parameters are shaped

by immediate socioeconomic conditions. In the simple scenario modeling, we found some evidence that lower poverty rates may predict more deforestation, while higher education attainment may predict reduced deforestation. This was also supported by the results of the district-specific specifications in Tables 11 and 12, and mapping in Figures 20 and 21. Deforestation particularly in the districts of Bắc Trà My, Trà Bồng, Cam Lộ, Triệu Phong, Hương Trà, and Đa Krông could be reduced through targeted efforts to improve immediate socioeconomic conditions. This evidence thus supports programs that encourage off-farm, off-plantation, and off-mine employment and educational opportunity.

Finally, the feasibility of conservation targets and associated programs can be determined in the context of continuing economic determinants of deforestation. We cautiously predict annual deforestation of approximately 2.59% in CAL, with spatial inconsistency, given 2017 economic conditions. This figure is quite high, but these conditions undoubtedly fluctuate, and the actual deforestation rate is mediated by property rights regimes, land tenure and concessions, and the local and national institutional and regulatory environment which are not included in this model. However, we emphasize that economic conditions, particularly regional and global-scale, underlying economic forces, will continue to drive deforestation at a fairly high rate. Strategic planning for forest conservation will take this into account in setting targets, prioritizing spatial locations, allocating program and project funding to areas of greater risk. Overall, Figure 1 illustrates that the greatest average tree cover loss by magnitude in CAL is in Quảng Nam province, but Figure 22 provides evidence that the forests in Quảng Trị province are more at risk due to sensitivity to economic conditions, both underlying and immediate. Quảng Trị districts, particularly Hướng Hóa, Đa Krông, and Cam Lộ, in addition to Quảng Bình and Thừa Thiên Huế districts A Lưới, Minh Hóa, Hương Khê, and Đăk Glei, display notable deforestation risk which should be incorporated into strategic planning.

5. Conclusion

In conclusion, this study has used tiered spatial regression analysis to correlatively identify which commodities, economic development activities, and social conditions have historically had the greatest effect on forest cover by magnitude in CAL. Based on first phase results, we selected a subset of these determinants for scenario modelling to predict possible deforestation outcomes given certain economic scenarios.

The results, among others, indicate that terms of spatial collocation, high poverty rates and smaller scale agricultural land conversion are key immediate determinants of deforestation. That said, further research is required to specifically break down the types of agricultural land use that are responsible. In the absence of land use data with spatial and temporal variation, we break down these land use types using the proxy of commodity prices. Price or index value increases for rubber, hardwoods, metals, and a tea-coffee-cocoa index are consistently associated with tree cover loss, ranging from approximately 2 to 16 ha of tree cover loss annually. This therefore provides evidence to support programs targeting rubber, acacia harvesting, artisanal mining, and land conversion for cash crop plantations. Education is also key immediate socioeconomic factor, as poverty rate is consistently associated with tree cover loss and increased educational attainment is consistently associated with reduced tree cover loss, at approximately 12 ha per percentage increase in secondary school graduation. On the macro level, economic growth in China and Vietnam are correlatively associated with tree cover loss, as are rising trade in Myanmar and Laos. This study

provides a methodological contribution to the current academic literature identifying socioeconomic dimensions of deforestation through spatial econometric analysis and scenario modelling at the landscape level. For practitioner work, we hope that this pilot study provides a tool for strategic planning of conservation interventions in light of economic conditions and factors, not only through the empirical analyses conducted but also through our systematic assessment of currently available data for understanding these issues. As a pilot, we hope to highlight areas of need for data collection and future research.

Based on the outcome of the data availability assessment and spatial and quantitative results of this baseline study, we recommend future research focusing on immediate drivers of deforestation (micro/socioeconomic determinants) should utilize currently available data by focusing on small scale target areas within landscapes. This will necessitate working with GSOV to acquire the household-level data from the household living conditions surveys that is not available online, or collecting new data by administering a small household living conditions survey focusing on off-farm wage opportunities and opportunities outside of timber, rubber, and other heavily-deforesting harvesting. With the baseline established in this study, future work can focus on using quasi-experimental methods such as border or regression discontinuity based on spatial policy differences or temporal “shocks.” Instrumental variables could also be explored. Lao data could be added for CAL, enabling border discontinuity design for economic and conservation programs on the Vietnamese side. Further, the economic growth contraction in Cambodia in 2009 could be a source of “shock” for a natural experiment, but this needs to be explored.

A key missing element in this analysis is local-level governance, which can vary across provincial and district levels. Following Khuc et al. (2018), a provincial level governance variable could be included in future studies. Additionally, property rights regimes and land tenure systems must be incorporated into this analysis as well to account for institutions which are immediate determinants of deforestation (Angelsen and Kaimowitz, 1999; Angelsen, 2010; Agarwal et al., 2008).

We also note some key strategy considerations for organizations involved in integrated conservation-development in Vietnam. Economic conditions and their effect on deforestation varies across districts in the CAL, and continuing rates of deforestation driven by these factors must be taken into account throughout project and program planning and implementation. Evidence supports conservation efforts focusing on rubber, acacia/hardwood logging, and artisanal mining, but socioeconomic conditions of agricultural communities near forests cannot be forgotten.

6. References

*indicates data source

- *AidData Geoquery. (2018). Precipitation, elevation, slope, and population density (Gridded Population of the World (GPWv4) at Vietnam GADM(3) and (4). Zip download. <<http://geo.aiddata.org/>>. See (Willmott and Matsuura, 2001) (Friedl et al., 2010) (Defourny, 2017) and (CIESIN, 2016).
- Angelsen, Arild, and David Kaimowitz. 1999. "Rethinking the Causes of Deforestation: Lessons from Economic Models." *World Bank Research Observer* 14 (1): 73–98.
- Anselin, Luc. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics*.
- Anselin, Luc. 1988. "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity." *Geographical Analysis* 20 (1): 1–17.
- Barbier, EB. 2017a. "Poverty, Rural Population Distribution and Climate Change." *Environment and Development Economics*.
- Barbier, Edward B, Philippe Delacote, and Julien Wolfersberger. 2017b. "The Economic Analysis of the Forest Transition: A Review." *Journal of Forest Economics* 27: 10–17. <https://doi.org/10.1016/j.jfe.2017.02.003>.
- Bernhard, K. P., Smith, T. E., Sabuhoro, E., Nyandwi, E., & Munanura, I. E. 2020. "Effects of integrated conservation–development projects on unauthorized resource use in Volcanoes National Park, Rwanda: a mixed-methods spatio-temporal approach." *Oryx*, 1-12. <https://doi.org/10.1017/S0030605319000735>.
- Bhattarai, Madhusudan, and Michael Hammig. 2001. "Institutions and the Environmental Kuznets Curve for Deforestation: A Crosscountry Analysis for Latin America, Africa and Asia." *World Development* 29 (6): 995–1010.
- Busch, Jonah, and Kalifi Ferretti-Gallon. 2017. "What Drives Deforestation and What Stops It? A Meta-Analysis." *Review of Environmental Economics and Policy* 11 (1): 3–23.
- Chi, Vu Kim, Anton Van Rompaey, Gerard Govers, Veerle Vanacker, Birgit Schmook, and Nguyen Hieu. 2013. "Land Transitions in Northwest Vietnam: An Integrated Analysis of Biophysical and Socio-Cultural Factors." *Human Ecology* 41 (1): 37–50.
- Chomitz, Kenneth M., and Piet. Buys. 2007. *At Loggerheads? : Overview : Agricultural Expansion, Poverty Reduction, and Environment in the Tropical Forests*. World Bank.
- *CIESIN. 2016. Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision UN WPP Country Totals. Center for International Earth Science Information Network (CIESIN). Columbia University. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
- Cochard, Roland, Dung Tri Ngo, Patrick O. Waeber, and Christian A. Kull. 2016. "Extent and Causes of Forest Cover Changes in Vietnam's Provinces 1993–2013: A Review and Analysis of Official Data." *Environmental Reviews* 25 (2): 199–217. <https://doi.org/10.1139/er-2016-0050>.
- Costenbader, John, Jeremy Broadhead, Yurdi Yasmi, and Patrick B Durst. 2015 "Drivers Affecting Forest Change in the Greater Mekong Subregion (GMS): An Overview."
- Cropper, Maureen, and Charles Griffiths. 1994. "The Interaction of Population Growth and Environmental Quality." *The American Economic Review* 84 (2): 250–54.
- Defourny, P. 2017. ESA Land Cover Climate Change Initiative (Land_Cover_cci): Land Cover Maps, v2.0.7. Centre for Environmental Data Analysis, 7/2017.

- DeFries, Ruth S., Thomas Rudel, Maria Uriarte, and Matthew Hansen. 2010. "Deforestation Driven by Urban Population Growth and Agricultural Trade in the Twenty-First Century." *Nature Geoscience* 3 (3): 178–81. <https://doi.org/10.1038/ngeo756>.
- Fotheringham, A Stewart, Chris Brunsdon, and Martin Charlton. 2000. *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage Publications.
- Friedl, M.A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley and X. Huang. 2010. MODIS Collection 5 Global Land Cover: Algorithm refinements and characterization of new datasets, 2001-2012, Collection 5.1 IGBP Land Cover, Boston University, Boston, MA, USA.
- *GADM. 2018. Database of Global Administrative Areas. Administrative boundaries, Vietnam. (.shp download). Center for Spatial Sciences, University of California, Davis.
- *General Statistics Office of Vietnam (GSOV). 2018. Statistical data: "Population and employment," "Industry," "Agriculture, Forestry and Fishery," "Trade, Price and Tourism," "Education, Health, Culture, and Living Standard," National Accounts." Statistical Censuses and Surveys. http://www.gso.gov.vn/default_en.aspx?tabid=468/. (various .xls downloads).
- Gibbons, Stephen, and Henry G. Overman. 2012. "Mostly Pointless Spatial Econometrics?" *Journal of Regional Science* 52 (2): 172–91.
- *Global Forest Watch (GFW). 2018. Global Forest Watch Open Data Portal. Acquired through WWF-Vietnam. (raster files).
- Hansen, M.C., et al. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (6160): 850–53.
- Hsiang, Solomon, and Nitin Sekar. 2016. "Does Legalization Reduce Black Market Activity? Evidence from a Global Ivory Experiment and Elephant Poaching Data." NBER Working Paper .
- *International Monetary Fund (IMF). 2017. Selected commodity prices and price indices. World Economic Outlook Database. (.xls download).
- Khuc, Quy Van, Bao Quang Tran, Patrick Meyfroidt, and Mark W. Paschke. 2018. "Drivers of Deforestation and Forest Degradation in Vietnam: An Exploratory Analysis at the National Level." *Forest Policy and Economics* 90 (May): 128–41.
- Lambin, Eric F., and Patrick Meyfroidt. 2010. "Land Use Transitions: Socio-Ecological Feedback versus Socio-Economic Change." *Land Use Policy* 27 (2): 108–18.
- Lee, Sang-Il. 2001. "Developing a Bivariate Spatial Association Measure: An Integration of Pearson's r and Morans' I ." *Journal of Geographic Systems* 3: 369–85.
- Meyfroidt, Patrick, and Eric F. Lambin. 2009. "Forest Transition in Vietnam and Displacement of Deforestation Abroad." *Proceedings of the National Academy of Sciences* 106 (38): 16139.
- Meyfroidt, Patrick, and Eric F. Lambin. 2008a. "The Causes of the Reforestation in Vietnam." *Land Use Policy* 25 (2): 182–97.
- Meyfroidt, Patrick, and Eric F. Lambin. 2008b. "Forest Transition in Vietnam and Its Environmental Impacts." *Global Change Biology* 14 (6): 1319–36.
- Moulton, Brent R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics* 72 (2): 334.
- Müller, D., and D. K. Munroe. 2015. "Tradeoffs between Rural Development Policies and Forest Protection: Spatially Explicit Modeling in the Central Highlands of Vietnam." *Land Economics* 81 (3): 412–25.

- Müller, Daniel, and Manfred Zeller. 2002. "Land Use Dynamics in the Central Highlands of Vietnam: A Spatial Model Combining Village Survey Data with Satellite Imagery Interpretation."
- Munroe, Darla K., and Daniel Müller. 2007. "Issues in Spatially Explicit Statistical Land-Use/Cover Change (LUCC) Models: Examples from Western Honduras and the Central Highlands of Vietnam." *Land Use Policy* 24 (3): 521–30.
- Panayotou, Theodore. 2000. "Economic Growth and the Environment."
- Scrieciu, S. Serban. 2007. "Can Economic Causes of Tropical Deforestation Be Identified at a Global Level?" *Ecological Economics* 62 (3–4): 603–12.
- Stock, James H, and Mark W Watson. 2006. "Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression." NBER Working Paper.
- Thiha, Edward L. Webb, and Kiyoshi Honda. 2007. "Biophysical and Policy Drivers of Landscape Change in a Central Vietnamese District." *Environmental Conservation* 34 (2): 164–72.
- Vogelmann, James E., Phung Van Khoa, Do Xuan Lan, Jacob Shermeyer, Hua Shi, Michael C. Wimberly, Hoang Tat Duong, and Le Van Huong. 2017. "Assessment of Forest Degradation in Vietnam Using Landsat Time Series Data." *Forests* 8 (7).
- *VNForest. 2018. Forestry Data Sharing System. <http://maps.vnforest.gov.vn/>.
- *Willmott, C. J. and Matsuura, K. 2001. "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series." University of Delaware.
- *World Development Indicators (WDI). 2018. Various countries GDP (PPP), Economic Growth, Trade as % of GDP. World Bank DataBank. Washington, DC. (.xls download).
- *World Database on Protected Areas (WDPA). 2014. United Nations Environment World Conservation Monitoring Centre (UNEP-WCMC). (.shp download) <https://www.unep-wcmc.org/resources-and-data/wdpa>.
- *WWF-GM Forest Report. 2018. "Pulse of the Forest: The State of the Greater Mekong's Forests and the Everyday People Working to Protect Them." World Wide Fund for Nature (WWF) – Greater Mekong.

7. Acknowledgements

This pilot study was conducted and working paper drafted initially as a World Wide Fund for Nature (WWF) – Greater Mekong report in 2018-19. The authors are grateful to WWF Vietnam, and to Benjamin Rawson, Thibault Ledecq and Tam Le Viet for data collection support, review, and comments. This project received funding from WWF Greater Mekong and WWF Switzerland.

8. Author contributions

KPB: conceptualization, data curation, data management, methodology, formal analysis, writing original draft, writing review and editing, visualization
 SZ: conceptualization, project management, methodology, writing review and editing
 ACS: conceptualization, methodology, writing review

Appendix I: Summary statistics

<i>Variable</i>	<i>Name</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>	<i>UNIT</i>	<i>Source</i>
<i>hasc_2</i>	Tier 2 code	0					(calculation var not used in analysis)	-
<i>yr</i>	Year	1,658	2008.423	5.154266	2000	2017	Year	Each data source
<i>FID</i>	Field ID	1,488	48	26.85426	2	94	(calculation var not used in analysis)	-
<i>Rowid</i>	Row ID	1,488	358.3763	214.2557	0	671	(calculation var not used in analysis)	-
<i>SHAPE_AREA</i>	Shape area	1,488	.0450188	.0390419	0	.177289	(calculation var not used in analysis)	-
<i>SHAPE LENG</i>	Shape length	1,488	1.238417	.6526275	0	2.541338	(calculation var not used in analysis)	-
<i>rainfedcrop</i>	Rainfed cropland	1,488	1008.121	913.4818	0	3812	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>shrubland</i>	Shrubland	1,488	562.9906	658.6824	0	3048	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>urban</i>	Urban area	1,488	19.98858	39.1673	0	324	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>forest</i>	Forest area	1,488	3109.345	3950.718	0	17305	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>irrigatedc~p</i>	Irrigated cropland	1,488	277.4435	298.0154	0	1454	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>grassland</i>	Grassland	1,488	18.36761	43.5445	0	292	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>mosaiccrop</i>	Mosaic cropland	1,488	662.3011	725.3838	0	3355	ESA Land Cover classification sq km/ aggregated to Tier 2 unit	AidData ESA Classifications
<i>popdens</i>	Population density	372	825.2087	2074.002	0	18783.16	Avg persons per sq km/per Tier 2 unit	AidData, CIESIN
<i>nightlights</i>	Nighttime lights	1,302	5.31403	9.863291	0	62.6062	Mean intensity of nighttime lights /per Tier 2 unit. Values 0-63. Version 4 DMSP-OLS Nighttime Lights composites. The lights from cities, towns, and other sites with persistent lighting, including gas flares. Ephemeral events, such as fires have been discarded. Calibrated across sensors and years using Elvidge 2014 coefficients.	AidData, DMSP-OLS
<i>investment~t</i>	Investment	1,488	24.37981	.9892948	22.68	25.798	Percent of GDP	AidData - WB
<i>tradevol</i>	Trade volume	1,488	5.252312	5.406457	-10.464	12.524	Percent change	IMF, 2017
<i>importvol</i>	Import volume	1,488	5.245875	5.507662	-10.91	12.436	Percent change	IMF, 2017
<i>exportvol</i>	Export volume	1,488	5.259063	5.309419	-10.017	12.611	Percent change	IMF, 2017
<i>com~lnonfuel</i>	Commodity price index (fuel +nonfuel)	1,488	123.0669	47.37507	58.22	192.054	Index (2005=100)	IMF, 2017
<i>com~_nonfuel</i>	Commodity price index (nonfuel)	1,488	127.4549	37.30306	75.968	190.472	Index (2005=100)	IMF, 2017
<i>indusinput~x</i>	Industrial inputs price index	1,488	126.0261	39.59833	71.321	197.733	Index (2005=100)	IMF, 2017
<i>commod~_fuel</i>	Commodity price index (fuel)	1,488	120.4997	53.8545	47.193	194.152	Index (2005=100)	IMF, 2017
<i>natgaspric~x</i>	Natural gas price index	1,488	111.6899	41.90832	47.971	173.657	Index (2005=100)	IMF, 2017
<i>coalpricei~x</i>	Coal price index	1,488	138.3316	66.41041	52.876	269.86	Index (2005=100)	IMF, 2017
<i>foodpricei~x</i>	Food price index	1,488	128.7581	36.02934	80.854	180.871	Index (2005=100)	IMF, 2017
<i>cerealspri~x</i>	Cereals price index	1,488	151.3909	54.64471	80.124	239.648	Index (2005=100)	IMF, 2017
<i>wheatprice</i>	Wheat price	1,488	190.6983	63.16341	100.739	286.952	USD per metric tonne	IMF, 2017
<i>maizeprice</i>	Maize price	1,488	166.4994	68.90804	88.219	298.41	USD per metric tonne	IMF, 2017
<i>riceprice</i>	Rice price	1,488	387.7876	164.5332	172.705	700.2	USD per metric tonne	IMF, 2017
<i>barleyprice</i>	Barley price	1,488	142.5569	47.67708	77.235	238.229	Index (2005=100)	IMF, 2017

<i>vegoilprice~x</i>	Vegetable oil price index	1,488	142.4804	49.41642	71.491	215.715	Index (2005=100)	IMF, 2017
<i>palmoilprice</i>	Palm oil price	1,488	603.5602	249.8757	238.398	1076.502	USD per metric tonne	IMF, 2017
<i>sunflowoil~e</i>	Sunflower oil price	1,488	988.2836	397.515	379.327	1693.648	USD per metric tonne	IMF, 2017
<i>meatpricei~x</i>	Meat price index	1,488	108.8643	24.0006	79.045	160.483	Index (2005=100)	IMF, 2017
<i>sugarprice~x</i>	Sugar price index	1,488	122.8624	40.30335	69.539	210.768	Index (2005=100)	IMF, 2017
<i>bananaprice</i>	Banana price	1,488	732.5805	203.1176	375.191	984.282	USD per metric tonne	IMF, 2017
<i>teacoffcoco~x</i>	Tea coffee cocoa price index	1,488	129.8981	43.04712	65.626	205.544	Index (2005=100)	IMF, 2017
<i>coffeeprice~x</i>	Coffee price index	1,488	125.0121	51.14986	53.056	230.969	Index (2005=100)	IMF, 2017
<i>coffeeothe~e</i>	Coffee (other types) price	1,488	133.8929	57.6727	60.365	273.209	US cents per pound	IMF, 2017
<i>coffeeobu~e</i>	Coffee robusta price	1,488	73.89556	30.30685	27.32	115.996	US cents per pound	IMF, 2017
<i>agrawprice~x</i>	Agricultural raw materials price index	1,488	113.8035	18.34207	93.929	153.45	Index (2005=100)	IMF, 2017
<i>timberprice~x</i>	Timber price index	1,488	99.994	9.694661	81.363	110.82	Index (2005=100)	IMF, 2017
<i>hardwoodpr~x</i>	Hardwood price index	1,488	116.2979	24.83246	75.47	157.43	Index (2005=100)	IMF, 2017
<i>hardlogspr~e</i>	Hard logs price	1,488	253.0916	65.51914	159.876	390.512	USD per cubic meter	IMF, 2017
<i>softwoodpr~x</i>	Softwood price index	1,488	90.97581	6.143839	82.736	104.075	Index (2005=100)	IMF, 2017
<i>woodpricei~x</i>	Wood price index	1,488	138.2834	46.31793	75.971	233.331	Index (2005=100)	IMF, 2017
<i>rubberprice</i>	Rubber price	1,658	92.07549	49.11899	26.08613	218.5146	USD per metric tonne	IMF, 2017
<i>metalspric~x</i>	Metals price index	1,488	134.8546	56.66187	54.32	229.719	Index (2005=100)	IMF, 2017
<i>tier1</i>	Tier 1 code	1,488	5.935484	2.440662	2	10	(calculation var not used in analysis)	-
<i>percentele~c</i>	Percent electricity	736	96.94823	4.479194	67.5	100	Percent HH with electricity access/per Tier 2 unit	GSOV, 2018
<i>percgrad</i>	Percent secondary graduation	1,472	91.42638	7.620427	62.35	99.86	Secondary school graduation percent/per Tier 2 unit	GSOV, 2018
<i>povrate</i>	Poverty rate	736	16.39022	7.505114	.5	31.9	Percent HH under poverty rate /per Tier 1 unit	GSOV, 2018
<i>livingcost~x</i>	Living cost index	635	91.26474	4.041347	83.79	97.98	Index of spatial living cost relative to capital/per Tier 2 unit	GSOV, 2018
<i>Rowid</i>	Row ID	1,226	41.42985	23.70841	1	86	(calculation var not used in analysis)	-
<i>ZONECODE</i>	Zone code	1,226	45.56117	25.04138	1	88	(calculation var not used in analysis)	-
<i>wwfforcovl~t</i>	Tree cover loss total	1,396	5136.196	8327.433	1	61003	(calculation var not used in analysis)	GFW, 2018, calculated in Stata
<i>wwfforcovl~a</i>	Tree cover loss area	1,396	3777490	6124531	735.4645	4.49e+07	Area tree cover loss in sq m /per Tier 2 Unit	GFW, 2018, calculated in Stata
<i>wwfforcovl~m</i>	Tree cover loss count	1,396	65327.26	126523.4	1	976048	(calculation var not used in analysis)	GFW, 2018, calculated in Stata
<i>hasc2code</i>	Tier 2 code	1,658	49.05609	26.91738	3	95	(calculation var not used in analysis)	Calculated in Stata
<i>lnwwfforloss</i>	Ln(tree cover loss)	1,396	13.37584	2.576517	6.600502	17.61918	Natural log	Calculated in Stata
<i>lnwwfk</i>	Ln(tree cover loss) k	1,396	13.95682	1.71748	11.41361	17.62118	Natural log, skew-corrected	Calculated in Stata
<i>lnirrigated</i>	Ln(irrigated cropland)	1,402	4.976443	1.41992	0	7.282073	Natural log	Calculated in Stata

<i>lnhardwood</i>	Ln(hardwood price index)	1,488	4.731606	.2259987	4.323735	5.058981	Natural log	Calculated in Stata
<i>lnriceprice</i>	Ln(rice price)	1,488	5.863631	.4493755	5.151585	6.551366	Natural log	Calculated in Stata
<i>lnmetalspr~e</i>	Ln(metals price)	1,488	4.7961	.4900977	3.994893	5.436857	Natural log	Calculated in Stata
<i>lnurban</i>	Ln(urban area)	930	2.385006	1.59144	0	5.780744	Natural log	Calculated in Stata
<i>lnhardwoodk</i>	Ln(hardwood price index k)	1,488	4.874066	.1870528	4.522005	5.15894	Natural log, skew-corrected	Calculated in Stata
<i>lnricepricek</i>	Ln(rice price k)	1,488	6.267818	.3024311	5.819008	6.761726	Natural log, skew-corrected	Calculated in Stata
<i>lnmetalspr~k</i>	Ln(metals price k)	1,488	6.191242	.115429	5.983548	6.349567	Natural log, skew-corrected	Calculated in Stata
<i>quintile</i>	Poverty quintile	736	2.967391	1.436747	1	5	Poverty quintile (group variable) by Tier 1 unit	GSOV, 2018, calculated in Stata
<i>asdf_id</i>	Spatial ID	1,334	428.4205	157.8965	0	599	(calculation var not used in analysis)	-
<i>precip</i>	Precipitation	1,184	198.1581	44.83779	94.58334	364.15	Mean annual millimeters precipitation/per Tier 2 unit	AidData
<i>temp</i>	Temperature	1,184	23.62242	1.74047	17.35176	26.99167	Mean annual temp in degrees Celsius/per Tier 2 unit	AidData
<i>slope</i>	Slope	1,334	4.527104	3.387298	.1338127	12.6958	Mean slope in degrees/per Tier 2 unit	AidData
<i>elevation</i>	Elevation	1,334	251.4279	305.839	4.304589	1339.736	Mean elevation in meters/per Tier 2 unit	AidData
<i>gdp</i>	Contribution to GDP	1,334	1.460543	3.308371	.0000252	21.15763	Tier 2 admin unit's contribution to national Vietnamese GDP (economic output)	AidData
<i>disttoroads</i>	Distance to roads	1,334	3396.703	3659.561	190.7539	25533.94	Mean distance to main roads in m/per Tier 2 unit	AidData, calculated in ArcGIS
<i>wdpa_count</i>	Protected area pixel count	1,334	470.4678	387.0547	3	1770	(calculation var not used in analysis)	WDPA, 2014, Calculated in ArcGIS
<i>wdpa_unpro~d</i>	Unprotected area pixel count	1,334	376.4198	304.751	0	1179	(calculation var not used in analysis)	WDPA, 2014, Calculated in ArcGIS
<i>tier1code</i>	Tier 1 code	1,658	5.931846	2.444964	2	10	(calculation var not used in analysis)	WDPA, 2014, Calculated in ArcGIS
<i>(protected area) intarea</i>	Protected area	968	182.4457	313.4021	.0176924	1800.889	Percentage of Tier 2 unit within protected area	WDPA, 2014, Calculated in ArcGIS
<i>wirrigated</i>	Spatial lag irrigated cropland	1,626	311.7327	182.7647	26.05	928.9375	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>wgdp</i>	Spatial lag Tier 2 contribution to GDP	1,626	1.248833	2.241556	.0153948	11.98844	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>wprecip</i>	Spatial lag precipitation	1,626	176.1176	46.06031	26.45322	256.969	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>welevation</i>	Spatial lag elevation	1,626	255.0623	255.9283	7.268671	1119.758	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>welectric</i>	Spatial lag electricity access	1,626	97.02775	1.829247	93.0625	99.7625	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>wwwfloss</i>	Spatial lag tree cover loss	1,626	3890341	2176173	9898.127	8173287	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>wnitelite</i>	Spatial lag nighttime lights	1,626	4.758083	7.417117	.0858728	38.18291	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa
<i>wpopdens</i>	Spatial lag population density	1,626	705.6348	1275.894	27.74868	5966.179	Spatial lag using queen contiguity matrix (not interpreted alone)	Calculated in GeoDa

<i>cambodia_g~h</i>	Cambodia GDP growth	1,658	6.039449	2.611541	-1.402854	11.48502	Annual percentage growth rate of GDP at market prices based on constant local currency. Aggregates are based on constant 2010 U.S. dollars.	WDI, 2018
<i>cambodia_gdp</i>	Cambodia GDP (PPP)	1,658	2354.375	875.9583	1097.756	4009.063	GDP per capita based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. Data are in current international dollars based on the 2011 ICP round.	WDI, 2018
<i>cambodia_t~e</i>	Cambodia trade	1,658	124.8362	10.4624	105.1385	144.6145	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>china_growth</i>	China GDP growth	1,658	8.693015	1.988203	6.123804	13.63634	Annual percentage growth rate of GDP	WDI, 2018
<i>china_gdp</i>	China GDP (PPP)	1,658	8574.22	4351.825	2933.315	16806.74	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>china_trade</i>	China trade	1,658	48.78495	8.856586	37.03382	64.47888	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>bangladesh~h</i>	Bangladesh GDP growth	1,658	4.511124	1.135794	1.937755	6.164763	Annual percentage growth rate of GDP	WDI, 2018
<i>bangladesh~p</i>	Bangladesh GDP (PPP)	1,658	2307.471	772.1446	1301.463	3868.822	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>bangladesh~e</i>	Bangladesh trade	1,658	37.76405	6.658623	26.85823	48.11092	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>indonesia~h</i>	Indonesia GDP growth	1,658	3.911159	.6995297	2.210344	4.906815	Annual percentage growth rate of GDP	WDI, 2018
<i>indonesia~p</i>	Indonesia GDP (PPP)	1,658	7864.375	2383.619	4601.849	12283.62	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>indonesia~e</i>	Indonesia trade	1,658	53.21463	9.298553	37.43889	71.43688	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>lao_growth</i>	Lao GDP growth	1,658	5.632572	.9323945	4.051386	6.889552	Annual percentage growth rate of GDP	WDI, 2018
<i>lao_gdp</i>	Lao GDP (PPP)	1,658	3972.428	1558.201	1969.702	7023.372	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>lao_trade</i>	Lao trade	1,658	79.42402	11.58429	60.5721	99.05974	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>myanmar_gr~h</i>	Myanmar GDP growth	1,658	9.211449	2.754738	4.761976	12.68542	Annual percentage growth rate of GDP	WDI, 2018
<i>myanmar_gdp</i>	Myanmar GDP (PPP)	1,658	3239.776	1597.176	1035.524	6160.703	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>myanmar_tr~e</i>	Myanmar trade	1,658	13.06653	18.84569	.1674176	47.95317	Trade is the sum of exports and imports of goods and services	WDI, 2018

							measured as a share of gross domestic product.	
<i>thai_growth</i>	Thai GDP growth	1,658	3.4442	2.250113	-1.189801	6.990585	Annual percentage growth rate of GDP	WDI, 2018
<i>thai_gdp</i>	Thai GDP (PPP)	1,658	12319.44	3187.598	7283.511	17872.22	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>thai_trade</i>	Thai trade	1,658	128.032	7.934509	114.9697	140.437	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>japan_growth</i>	Japan GDP growth	1,658	.9445324	1.928018	-5.404594	4.173028	Annual percentage growth rate of GDP	WDI, 2018
<i>japan_gdp</i>	Japan GDP (PPP)	1,658	34444.29	4874.833	26838.87	43278.99	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>japan_trade</i>	Japan trade	1,573	28.38769	5.648063	19.79813	37.54577	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>hk_growth</i>	Hong Kong GDP growth	1,658	3.232382	2.787158	-2.668881	7.855551	Annual percentage growth rate of GDP	WDI, 2018
<i>hk_gdp</i>	Hong Kong GDP (PPP)	1,658	43673.26	11058.39	26962.56	61540.16	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>hk_trade</i>	Hong Kong trade	1,658	356.2186	61.29775	240.849	442.62	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>korea_growth</i>	Korea GDP growth	1,658	3.566257	1.902555	.1905173	8.017418	Annual percentage growth rate of GDP	WDI, 2018
<i>korea_gdp</i>	Korea GDP (PPP)	1,658	28091.49	5900.229	18083.08	38335.27	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>korea_trade</i>	Korea trade	1,658	83.16543	15.69702	60.15371	110.0001	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>india_growth</i>	India GDP growth	1,658	5.535079	2.033808	2.015559	8.763184	Annual percentage growth rate of GDP	WDI, 2018
<i>india_gdp</i>	India GDP (PPP)	1,658	4010.561	1560.462	1977.645	7059.324	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>india_trade</i>	India trade	1,658	43.08632	9.355209	26.27484	55.79372	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>singa_growth</i>	Singapore GDP growth	1,658	3.316941	4.154112	-3.588249	13.21649	Annual percentage growth rate of GDP	WDI, 2018
<i>singa_gdp</i>	Singapore GDP (PPP)	1,658	65622.17	17001.88	40413.86	93905.42	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>singa_trade</i>	Singapore trade	1,658	374.29	34.68102	310.2561	441.6038	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>phil_growth</i>	Philippines GDP growth	1,658	3.439368	1.740815	-.462891	5.903122	Annual percentage growth rate of GDP	WDI, 2018
<i>phil_gdp</i>	Philippines GDP (PPP)	1,658	5317.924	1491.208	3348.152	8342.804	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>phil_trade</i>	Philippines trade	1,658	81.0987	16.73236	60.24529	104.7299	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>nepal_growth</i>	Nepal GDP growth	1,658	2.88099	1.80279	-1.528662	6.71925	Annual percentage growth rate of GDP	WDI, 2018
<i>nepal_gdp</i>	Nepal GDP (PPP)	1,658	1843.573	454.3585	1220.425	2696.694	GDP per capita based on purchasing power parity (PPP)	WDI, 2018

<i>nepal_trade</i>	Nepal trade	1,658	47.77292	4.120536	41.82825	55.79992	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018
<i>viet_gdp</i>	Vietnam GDP growth	1,658	4046.994	1420.109	2030.797	6775.832	Annual percentage growth rate of GDP	WDI, 2018
<i>viet_growth</i>	Vietnam GDP (PPP)	1,658	6.41405	.6791641	5.247367	7.547248	GDP per capita based on purchasing power parity (PPP)	WDI, 2018
<i>viet_trade</i>	Vietnam trade	1,658	148.4803	24.62612	111.4171	200.3846	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.	WDI, 2018

Appendix II. Data availability assessment

Reliable methods for estimation using spatial panel econometrics require data from multiple specific administrative units or spatial subareas, which can be difficult to obtain in the absence of annual, or at least periodic, raster data or consistent government data collection. We sought data at the Tier 2 administrative level in order to establish sufficient observations in a panel for statistical robustness. In addition to the need for multiple spatial units, these methods become more reliable with more periods of time series. We thus sought data for the time period 2000-2017, with data only available in waves (e.g., 2005, 2010, 2015) imputable for the entire period. In this regard, we summarize our findings of key areas of missing, incomplete, or available but inaccessible data at the landscape level with the following:

- **Socioeconomic household survey data** available in a usable **disaggregated** format online
 - Particularly data for off-farm wages and poverty rates at the spatial and temporal scale required
 - GSOV data from household living conditions surveys is available in aggregated format for online download, however the disaggregated data is not available for download. The aggregated versions usually have one but not both of the spatial and temporal variation necessary. For example, we use poverty rate as a group categorical variable at the provincial level.
- **Commodity-specific spatial land use area with multiple temporal waves**
 - E.g., raster image of rice agricultural area 2010 and 2015; raster image of timber production area 2010 and 2015
 - FORMIS data (from the Forestry Data Sharing System) is currently an excellent resource for fine-scale visualization, but the Data Catalogue is still under construction. Due to a technical problem on the website, we were also unable to download the statistical reports at the district level for forest function sub-class. Therefore, while the online visualization of land use categories (Forest function sub-class: Watershed, protection forest for tide shielding, wind and sand shielding forest, environmental protection, national park, nature reserve, scientific research, historical and landscape area, big timber production, small timber production, bamboo production, other production) is useful, the data downloads are unusable for the purposes of this study. Furthermore, even if available, these would only be available for 2016 and 2017. The full raster or vector datasets would need to be made available for even the one-time-period use. However, this could still be useful for a much finer-scaled study, rather than at landscape level.
 - Product GFSAD30SEACE (Southeast Asia) in the Global Croplands project from NASA's Land Processes Distributed Active Archive Center has only three classifications: water, non-cropland, and cropland.
 - This would enable calculation of land conversion rate and comparison to price as a signal
- **District level economic activity data** including import/export volume, manufacturing/industrial production, local (Tier 2) commodity production and prices with **multiple temporal waves**
 - GSOV data for this is available but usually either at district level OR over time, not both (over time is usually aggregated to country or provincial level); in this study we have

extracted variables that do have both, however these were limited and constrained our analysis.

As a pilot study, we took a broad approach in secondary data collection and assessment of data availability. While we took care to identify and incorporate necessary control and explanatory variables, we used acceptable proxies and linear imputation when key variables were unavailable for the desired time period or entirely unavailable. Table 1 details the variable types used and theory backing those variable types in the literature, including literature justifications for any proxies used. See Appendix I for units and summary statistics for specific variables.

We encourage future research to move beyond this baseline with finer-scaled or newly collected data. For example, this study could be reproduced at the district level, using the methodology applied here in addition to following Müller and Zeller (2002) and Munroe and Müller (2007). Tree cover loss data and microeconomic data can be aggregated to 1 km-grid cells, and a small micro-survey of households should be conducted if the fine-scaled GSOV household survey data is not acquired. This would also enable more effective use of the Forestry Data Sharing System database.

Appendix III: Understanding the Moran's I Outputs

Univariate and bivariate cluster and significance maps are presented in this study. Here we further describe the interpretation of these maps.

a. Univariate (*Local Indicator of Spatial Association: LISA*)

For the univariate statistic, a variable is compared to its own spatial lag. The spatial lag is essentially an average of the values of the variable for the districts surrounding a given district, based on contiguous neighboring districts in this study. The first “High” or “Low” is thus the x-axis of the Moran’s I plot, which is the variable, and the second “High” or “Low” is the y-axis, which is its spatial lag. For example, for the univariate local Moran’s I of precipitation, High-High (dark red) indicates a district of high precipitation surrounded by districts of low precipitation. Low-Low is a district of low precipitation surrounded by districts of low precipitation. Low-High is a district of low precipitation surrounded by districts of high precipitation. Finally, High-Low is a district of high precipitation surrounded by districts of low precipitation. The green significance maps show the level of statistical significance of these clusters. Only plots with an overall local univariate Moran’s I statistic (presented above the x-y plot) significant at a threshold of 5% were included in this study.

b. Bivariate (*Bivariate Local Indicator of Spatial Association: BiLISA*)

For the bivariate statistic, a variable is compared to the spatial lag of another variable, which is based on district contiguity (or k nearest neighbor) in this study. Similar to the univariate Moran’s I, the first “High” or “Low” is the x-axis of the Moran’s I plot, which is the non-spatial variable for a given district, and the second “High” or “Low” is the y-axis, but in the bivariate version this is the spatial lag of a different variable. For example, for the bivariate local Moran’s I of irrigated agriculture relative to tree cover loss, High-High (dark red) indicates a district of high irrigated agriculture surrounded by districts of high tree cover loss. Low-Low is a district of low irrigated agriculture surrounded by districts of low tree cover loss. Low-High is a district of low irrigated agriculture surrounded by districts of high tree cover loss. Finally, High-Low is a district of high irrigated agriculture surrounded by districts of low tree cover loss. The nature of the spatial lag is such that the bivariate statistic incorporates the neighborhood of districts but not the specific district itself, and as such must be cautiously interpreted; these statistics, however, are nonetheless useful for understanding the spatial distribution of economic and socioeconomic variables relative to tree cover loss. The green significance maps show the level of statistical significance of these clusters. Only plots with an overall local univariate Moran’s I statistic (presented above the x-y plot) significant at a threshold of 5% were included in this study.

For both map types, numbers in parentheses (e.g., Not Significant (23) or High-High (5)) indicate the number of observations within that category.

See Anselin (1988; 2018), Anselin et al. (2002), Anselin et al. (2008), Lee (2001), and Fotheringham et al. (2000) for further information on these statistics and their methods.