FUNDAMENTAL RESEARCH

*For. Sci.* 64(2):129–138
doi: 10.1093/forsci/fxx002


economics

# Forest Cover, Agricultural, and Socio-Economic Development: A Weighted Beta-Logistic Approach with Ratio Response

**Francisco X. Aguilar and Nianfu Song**

A beta-logistic generalized linear mixed model was used to study the association of national-level forest cover with selected covariates from a sample of 158 nations over the 1992–2013 period. The model avoided the improper assumption of normally distributed forest cover data and used covariates representing land, economic, and social factors. Agricultural land expansion was the most important factor associated with declines in forest cover. Relationships with other covariates are more nuanced. We found no support for an environmental Kuznets curve; however, a Kuznets-like trajectory with population and education variables was discerned. Population density had a negative marginal effect on forest cover up to 220 people/square kilometer ($km^2$) but changed to positive after this threshold. A turning point for proportion of rural population was found at 22% when the association changed from positive to negative. The threshold for education level was 93% when its association switched from negative to positive. Economic variables inclusive of per capita income and 10-year lagged GDP growth rate had a weak but statistically significant association. In the future, high-income nations are expected to continue moderate growth in forest cover; a few fast-urbanizing developing countries are predicted to keep increasing forest area, but most will be likely to see a decline.

**Keywords:** forest cover, deforestation, development, low and middle-income countries, rural, environmental Kuznets curve, beta-logistic

## Introduction

Human activities have been the driving factors behind changes in the world's forest cover, composition, and structure for over a century (Liu et al. 2007, Vogt et al. 2007). In recent years, a few countries have successfully expanded their forest land area, but most are still converting forests to other uses (FAO 2015). Recent estimates show that deforestation trends worldwide have slowed down but remain a major concern to attain sustainable development (MacDicken 2015). Past studies at multiple geo-political scales have improved our understanding of the complexity behind deforestation trends, but their results are far from conclusive (Damette and Delacote 2012, Leblois et al. 2017).

Global deforestation rates are not constant, and within the same country trends might reverse over time (World Bank 2016). It is generally agreed that at early stages of economic development extractive forest practices may offer a sizable contribution to a country's economy, but at the expense of widespread deforestation and forest degradation (Zeder 2008, Edelson 2007). Technological modernization can even exacerbate the deforestation process in low- or middle-income countries by lowering the cost of forest extraction, thus accelerating the process of land-use change (World Bank 2016). As a country moves into a higher-income category, a reversal of this trend might be expected following a trajectory commonly referred to as the environmental Kuznets curve (Stern et al. 1996). A trend toward environmental protection illustrated by the restoration of forest cover has been observed in the United States as the country concluded its agricultural expansion, in Europe following World War II, and more recently in rising forest acreages in China and Vietnam. However, there are numerous exceptions to an environmental Kuznets curve trajectory (Stern 2004). One notable example is Brazil, which experienced fast deforestation over the past two decades in spite of concurring economic expansion (World Bank 2016).

Manuscript received March 30, 2017; accepted November 13, 2017; published online March 19, 2018.

**Affiliation:** *Francisco X. Aguilar (aguilarf@missouri.edu), Associate Professor, The School of Natural Resources, 203 Anheuser-Busch Natural Resources Bldg., University of Missouri, Columbia, MO 65211, Phone: (573) 882–6304. Nianfu Song (songn@mail.missouri.edu), Research Associate, The School of Natural Resources, 203 Anheuser-Busch Natural Resources Bldg., University of Missouri, Columbia, MO 65211, Phone: (573) 875–5341 x 235*

**Acknowledgments:** *This work was partly supported by the US Department of Agriculture's McIntire-Stennis Project MO-MCNR005 and Joint Venture Agreement with the US Forest Service 09-JV-11242311-064. Professor Christopher K. Wikle of the Department of Statistics at University of Missouri assisted with the statistical analysis.*

Multiple drivers of deforestation have been identified, but conclusions about the role of socio-economic and other forces are mixed and far from universally accepted (Geist and Lambin 2002). Settlement and agricultural development pressures have been identified as major forces behind deforestation by many, for example, Assefa and Bork (2014), De Sá et al. (2013), Giri et al. (2015), and Tadesse et al. (2014). Economic policies (Bhattarai and Hammig 2001, Bonilla-Moheno et al. 2012, Tadesse et al. 2014) and trade liberalization (Meyfroidt et al. 2013, Leblois et al. 2017) have been linked to changes in forest cover too. Rates of economic and population growth as well as poverty levels in rural areas have been associated with deforestation (Ehrhardt-Martinez 1998). The relationship between income level and forest land changes has been a central research subject, but no consensus on a definitive association has been reached (Ehrhardt-Martinez 1998, Caviglia-Harris and Sills 2005, Swinton et al. 2003, Angelsen 2010, Culas 2012). Some support an environmental Kuznets curve trajectory (e.g., Bhattarai and Hammig 2001); some do it tenuously (e.g., Barbier and Burgess 2001); and others suggest that income effects have been overestimated or misspecified (e.g., Koop and Tole 1999, Stern 2004, Damette and Delacote 2012). Numerous other factors ranging from timber prices to agricultural yields have been reported in the literature, with no consistency over roles and statistically significant effects on changes in forest cover.

The literature examining factors associated with changes in forest area is ample in its geographic and temporal scope. Most recent studies have focused on developing countries or self-specified regions where deforestation is prevalent (Leblois et al. 2017). For instance, Ehrhardt-Martinez (1998) used panel data that excluded developed countries to find an association between deforestation and selected socioeconomic indicators. Koop and Tole (2001) estimated a model that analyzed the association between deforestation and gross domestic product (GDP), population, growth rates of these two variables, and distribution of income in developing countries only. Scrieciu (2006) used panel data from 50 tropical countries to analyze causes of tropical deforestation. Rudel (2007) estimated a logistic model based on limited data information on deforestation in tropical countries too to suggest that deforestation has shifted from a state-sponsored to an enterprise-driven process. Van and Azomahou (2007) used panel data from 59 developing countries to conclude that quadratic terms improved estimation of deforestation levels. Unfortunately, only two coefficients in that estimation were significant, and none of them were quadratic terms. Culas (2012) used panel data from 43 tropical countries to estimate deforestation in three different continents, and GDP per capita was found to be significant in all. However, Culas (2014) combined seemingly unrelated regression and ordinary least squares models to estimate deforestation in three continents to conclude that each region needed different parameters for their equations. Leblois et al. (2017) used a random-effects panel data model for deforestation in developing countries and concluded that agricultural trade and forest transition stage were two major factors associated with deforestation.

This study investigated the systematic association between national-level forest cover rates and selected land and socioeconomic covariates. Its aim was to quantify the systematic relative importance of selected covariates on forest cover rates using the most recently available information. Our analytical approach relied on a beta-logistic generalized linear mixed model (GLMM) to overcome issues of violation of normally distributed data (Bonnor 1967, Johansson 1985, Jennings et al. 1999, Rautiainen et al. 2005). The beta distribution of national-level forest rates enabled estimation of a common model for countries at all development stages. Thus, we offer a comprehensive analysis of changes in forest cover rates not limited to one-directional deforestation or tropical developing countries. In this paper, we describe our analytical approach to studying changes in national-level rates of forest cover, present our findings, and offer implications for individual covariates in the context of the extant literature.

## Analytical Framework

A nation's forest cover rate represents the ratio between forest extension and territorial land area. Forest cover rates ranging from 0 to 1 offer comparable relative measures of forest extension across nations irrespective of total land area. Forest cover values have been a metric commonly used to capture deforestation in several cross-country studies, with the shortcoming that many had erroneously assumed normally distributed error terms (Bonnor 1967, Johansson 1985, Jennings et al. 1999, Rautiainen et al. 2005). Ferrari and Cribari-Neto (2004) and Korhonen et al. (2007) have suggested the use of beta-logistic regression when dealing with ratio data in ecological applications and specifically in forest land changes.

A beta-logistic GLMM allows modeling of heterogeneous observations by including random intercepts designed to capture idiosyncratic conditions (Lindstrom and Bates 1990, Van and Azomahou 2007). Let $Y_{it}$ be the forest cover rate of country i at year t, and $Y_{it} \sim \mathrm{Beta}(\alpha_1, \alpha_2)$ is assumed to have a beta distribution that can take a variety of shapes depending on the values of the parameters $\alpha_1$ and $\alpha_2$ (Casella and Berger 2002). The expected value $E(Y_{it})$ is a logistic function of a vector of covariates $X_{it}$:

$$E(Y_{it}) = \frac{e^{X_{it}\beta}}{1+e^{X_{it}\beta}}, \tag{1}$$

where $\beta$ is a vector of estimated parameters. $X_{it}\beta$ is commonly referred to as the systematic component of a GLMM (Schabenberger

### Management and Policy Implications

Analysis of changes in country-level forest covers over 20 years shows that associations with land, economic, and social conditions are complex. Agricultural land expansion was identified as the most important factor associated with declines in forest cover rates. But these trends are not necessarily uniform. Improvements in education had an initial perverse effect by reducing forest area, but that trend was reversed once most of the population attained higher education levels. Increased in population density and the proportion of rural dwellers had similar associations with forest cover. Measures of economic development were correlated with recovering forest cover rates, but their association was weaker than that of other factors. Public policies that support agricultural intensification, formal education, and higher population density in urban areas will likely facilitate conservation and expansion of forests. However, any recovery in forest cover might only occur in the middle to long term, often taking longer than 10 years to be observed. Policymakers and other forest sector stakeholders will need to exercise adequate patience and enduring commitments.

and Pierce 2001). The intercept $\beta_{i0}$ in $\beta$ is assumed to be a random variable for country i so that the model is generalizable. Specifically:

$$\beta_{i0} = \beta_0 + b_{i0} \quad (2)$$

where $b_{i0}$ is assumed to be a normally distributed random variable with mean 0. The random component $b_{i0}$ adjusts the intercept of the model for country i. The jth estimated coefficient $\beta_j$ represents the change in the logarithm of odds ratio $\log\left(\frac{E(Y_{it})}{1-E(Y_{it})}\right)$ as a result of a unit change in $X_{itj}$, the jth variable for country i in period t (Breslow and Clayton 1993). A country shares the same fixed coefficients with other countries but has its own intercept $\beta_{i0}$. Consequently, this model provides same coefficients for covariates conditional on the random intercept for each country.

Statistical estimation included weights, and the model was specified to capture nonlinear associations. Each country observation was weighted by its total land area to reflect the relative greater importance of large-territory countries in the analysis. Squared covariates were included in the systematic expression following Van and Azomahou (2007), and $X_{it}\beta$ was specified as a polynomial function of covariates to increase model flexibility. The systematic component of the generalized linear model is then an additive polynomial of individual covariates. $X_{it}\beta = \sum_{j=1}^{8}(\beta_{i0} + Z_{itj}\beta_{j1} + Z_{itj}^2\beta_{j2})$, where $Z_{itj}$ and $Z_{itj}^2$ are respectively the jth covariate and its square. $\beta_{j1}$, and $\beta_{j2}$, for j= 1 to 8, are common parameters for all countries denoting the eight covariates identified in our empirical estimation.

Definitions of the response, the eight covariates included in the model, and land area as the weight in the estimation are summarized in Table 1. The eight covariates were selected to capture salient land, economic, and social factors reportedly associated with forest cover rates in the literature and reflect data availability at the time of the study. Covariates corresponded to (1) agricultural land area per capita, (2) gross national income (GNI) per capita, (3) 10-year lagged GDP growth rate, (4) population density, (5) population growth rate, (6) proportion of rural population, (7) rate of secondary school enrollment, and (8) 15-year lagged rate of school enrollment. Agricultural land area as a proportion of total land area of a country was included in this model to capture land changes linked to agricultural expansion or contraction (López-Carr and Burgdorfer 2013). GDP growth rate represents an indicator of national economic expansion that was lagged by 10 years to denote the delay between economic improvements and slowdown of deforestation (Leblois et al. 2017). GNI per capita represents the average income of citizens in a country (including domestic and foreign income) adjusted by purchasing power parity to denote current income levels comparable across nations. Population density and education were used as basic demographic descriptors. Population density can affect consumption of food, firewood, and fibers, among other land resources, thus increasing pressure on forest utilization and conversion to alternative land uses (Auty 2001). Education levels may help explain changes in people's ability and opinions about forests and a precursor to stronger social institutions that might influence future deforestation trends (Chevalier et al. 2004). The gross rate of enrollment in secondary schools, which helped capture cross-country differences in education levels, was lagged 15 years as gains in education levels do not have immediate but decadal lagged effects on forest cover (Ehrhardt-Martinez 1998). The percent of rural population in a country aimed to capture the segment of the population whose income is largely derived from the land (Bhattarai and Hammig 2004). Also, on average, an advanced economy has a smaller share of rural population (Moomaw 1988, Restuccia et al. 2008, World Bank 2016). Moreover, rural and urban populations exhibit different livelihoods that excise different types of impacts on a country's forest resources. Effects of other factors not included in our covariates are considered country-specific or global random effects, and hence are included respectively in random intercepts and error terms.

The marginal association[1] of variable $Z_{itj}$ at year t can be expressed as a function of rate $E(Y_{it})$ and coefficients $\beta_{j1}$, and $\beta_{j2}$ by taking the partial derivative of Equation (1) with respect to $Z_{itj}$ (Long and Freese 2006):

$$\frac{\partial E(Y_{it})}{\partial Z_{itj}} = \frac{e^{X_{it}\beta}}{1+e^{X_{it}\beta}} \cdot \frac{1}{1+e^{X_{it}\beta}}(\beta_{j1} + 2Z_{itj}\beta_{j2}),$$

$$\text{where } X_{it}\beta = \sum_{j=1}^{8}(\beta_{i0} + Z_{itj}\beta_{j1} + Z_{itj}^2\beta_{j2})$$

$$\frac{\partial E(Y_{it})}{\partial Z_{itj}} = E(Y_{it})[1 - E(Y_{it})](\beta_{j1} + 2Z_{itj}\beta_{j2}), \quad (3)$$

When $Y_{it}$ is known for an observation year, the marginal association with a covariate at year t can be estimated as:

$$\frac{\partial E(Y_{it})}{\partial Z_{itj}} = Y_{it}(1 - Y_{it})(\beta_{j1} + 2Z_{itj}\beta_{j2}), \quad (4)$$

The sign of the marginal association of a variable $Z_{itj}$ on forest cover $Y_{it}$ is determined by the sign of $\frac{\partial X_{it}\beta}{\partial Z_{itj}} = \beta_{j1} + 2Z_{itj}\beta_{j2}$. Since there are different intercepts for countries in the systematic component of the model, the marginal association represents the impact of a

**Table 1. Definitions of covariates and the weight used in the GLMM for the global forest cover rate model of country i in year t.**

| Variables | Definitions |
|---|---|
| $Y_{it}$ | Forest cover rate, ratio between forest area, and total land area at the end of year t |
| $ALP_{it}$ | Agricultural land area as proportion of total land area at the end of year t |
| $GDP10_{it}$ | 10-year lagged value of GDP growth rate |
| $GNIC_{it}$ | Gross national income per capita measured in PPP dollars at the end of year t |
| $LND_{it}$ | Land area at the end of year t, a variable for weight in the estimation |
| $POG_{it}$ | Population annual growth rate in year t |
| $POPD_{it}$ | Population density as number of people per km² at the end of year t |
| $RPO_{it}$ | Rural population as percent of total population at the end of year t |
| $SCH_{it}$ | Rate of secondary school enrollment in period t, enrollment regardless of age, as percent of population of official secondary education age. This value could be larger than 100% when students in other age group enroll in secondary schools |
| $SCH15_{it}$ | 15-year lagged values of rate of secondary school enrollment |

unit change in a covariate on the forest cover rate of country i at year t. Because the values of covariates and forest cover rate vary annually, marginal associations are expected to change over time. One interesting property of the marginal association is that it approaches 0 as the value of $Y_{it}$ nears 0 or 1. This property implies that the predicted forest cover rate will be bounded between 0 and 1 even if the value of a covariate is extremely large. Finally, the relative importance of the association of each covariate on countries' forest covers was of special interest. Hence, standardized coefficients were estimated to denote the variation captured by linearly independent variables when assumed to be independent (Kutner 2004).

GLMM estimation with beta distribution, logit link, and a weight variable for land area was conducted using the SAS procedure GLIMIX (Bolker et al. 2009). Degrees of freedom were determined by the DDFM=BETWITHIN SAS option that splits degrees between and within countries (Schluchter and Elashoff 1990).

## Data

Data were primarily retrieved from the World Bank Open Data (World Bank 2016). National school enrollment numbers for Brazil and Turkmenistan came from the United Nations Educational, Scientific, and Cultural Organization (UNESCO 2016) and the United Nations Statistics Division (2016). Annual data of 1778 observations for 158 nations from 1992 to 2013 were used in the estimation. Forest cover rates varied among nations, with some observing an increase in forest cover rates while others decreased from 1992 to 2013 (Figure 1), consistent with other recent global assessments (Keenan et al. 2015). Rates, however, remained relatively stable over the years.

The plot on the right of Figure 1 illustrates the association between forest cover and GNI. Brazil, China, Iran, and Mexico had similar values of GNI per capita: around 5,000 to 18,000 US dollars per person per year. Despite similar GNI per capita, Brazil and Mexico had declining trends while Iran and China denoted increasing trends. Their forest cover rates were also very different. Brazil's forest cover rates were around 0.6; Mexico's between 0.3 and 0.4; China's slightly above 0.2; and Iran's below 0.1. These different patterns illustrate the inherent complexity of associating forest cover rates with covariates and emphasize the importance of including country-specific coefficients (Van and Azomahou 2007).

Table 2 shows data descriptions over our study period. Forest cover rates of countries vary from <0.01 in Omen to 0.99 in Suriname. Covariates expressed in percent values vary greatly; for example, <0.57% to 85.28% for agricultural land area, and from –5.20% to 9.93% for population growth rate. Extreme values of land areas ranged from 30 thousand km$^2$ to 16 million km$^2$.

## Study Limitations

Our approach to assessing the association between national-level forest cover rates and selected covariates faced several limitations. Among them we highlight two of the most salient regarding the challenge of national-level information and empirical model specification. These caveats are not particular to our study but are relevant to other research examining forest area dynamics at a national level or evidence of an environmental Kuznets curve in deforestation trends (e.g., List and Gallet 1999, Roca 2003). Our model relies on information aggregated at the national level (e.g., country's forest cover, secondary school enrollment rate). The scale of observations inherently assumes that forest resources, institutions, and markets, among others, are homogeneous within a given country. Hence, our inferences about the association between covariates and forest cover rate are made for an average nation weighted by land area. By including a country's idiosyncratic conditions through random effects, the model controls for nationwide specific effects of variables not included as covariates (e.g., country's ownership patterns, national policies) and allows for a more flexible model specification (Stern 2004), but the specific effects of programs or conditions of regions within a country cannot be tested in this model. For example, it would be empirically impossible to include a variable that offers a consistent and standard measure of sub-national programs and policies across all 158 countries included in this study.

Arguably, the choice of explanatory variables remains an empirical issue. The literature offers numerous examples of how coefficients and levels of significance tend to vary as explanatory variables are dropped or added. Some have relied to the presentation of various model specifications to address this issue; for

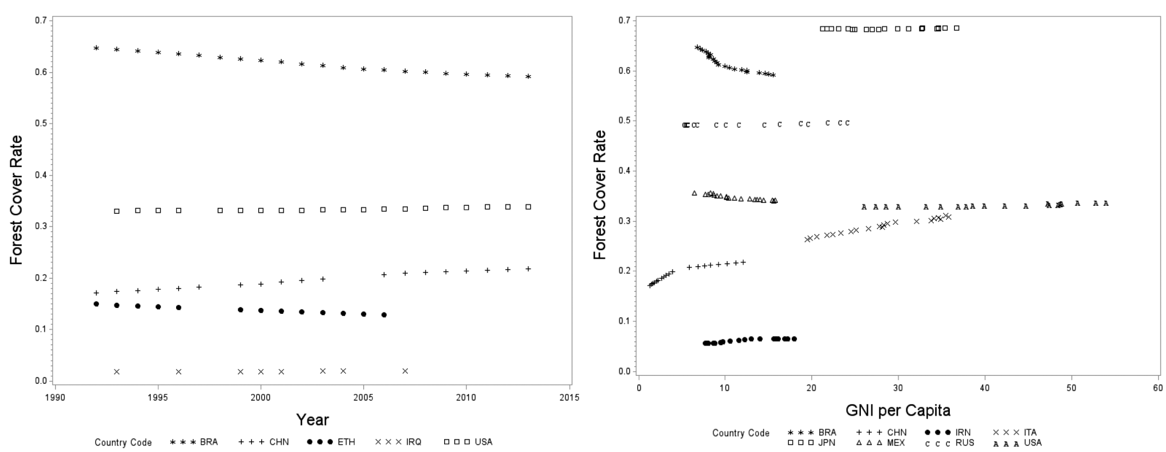


**Figure 1. Forest cover rates from 1992 to 2013 (left) and plotted against GNI per capita (right) for selected countries (BRA = Brazil, CHN = China, ETH = Ethiopia, IRQ = Iraq, IRN = Iran, ITA = Italy, JPN = Japan, MEX = Mexico, USA= United States). Data source: World Bank (2016).**

**Table 2. Sample means weighted by land area, minimum, maximum, and units of variables used in the estimation.**

| Variable | Mean in 1992 75 countries | Mean in 2012 62 Countries | Minimum of 158 countries | Maximum of 158 countries | Unit |
|---|---|---|---|---|---|
| $Y_{it}$ | 0.31 | 0.34 | <0.01 | 0.99 | Ratio: forest cover/total land |
| $ALP_{it}$ | 31.15 | 37.65 | 0.57 | 85.28 | % |
| $GDP10_{it}$ | 3.14 | 4.86 | -10.82 | 23.60 | % |
| $GNIC_{it}$ | 7.55 | 19.36 | 0.23 | 66.91 | International $1,000/capita |
| $LND_{it}$ | N/A | N/A | 0.03 | 16,376,87 | 1,000 $km^2$ |
| $POG_{it}$ | 1.78 | 1.05 | -5.20 | 9.93 | % |
| $POPD_{it}$ | 55.62 | 64.72 | 2.58 | 1,310.80 | People/$km^2$ |
| $RPO_{it}$ | 48.34 | 33.35 | 5.01 | 93.71 | % |
| $SCH_{it}$ | 58.97 | 91.97 | 6.16 | 135.54 | % |
| $SCH15_{it}$ | 49.51 | 81.97 | 1.91 | 151.00 | % |

*Note:* The values are not world average but weighted values of sampled countries; international $ dollar values are transformed by PPP conversion factors from World Bank; LND for land area was used as a weight but not a covariate in the model.

example, Ehrhardt-Martinez (1998) present seven different model specifications, Ehrhardt-Martinez et al. (2002) offer six, Jorgenson (2006) five, and Leblois et al. (2017) four. The use of institutional variables offers an example related to covariates whose in/exclusion is debatable. As a case in point, the literature is not consistent in terms of the inclusion and statistical significance of institutional variables or even what metrics should be used to capture associated effects. For instance, Bhattarai and Hammig (2001) controlled for political institutions using a rating scale derived from Freedom House, as did Van and Azomahou (2007), and later Leblois et al. (2017) used an institutional index derived from Freedom House ratings and other proxies. Bhattarai and Hammig (2001) and Van and Azomahou (2007) found significant effects of institutional variables on deforestation, while Leblois et al. (2017) did not. Others, such as Culas (2012) or Jorgenson (2006), did not include any institutional variables.

The number and nature of covariates in our model resemble those of Leblois et al. (2017), but our data cover a longer period and are not limited to developing countries, arguably better representing global trends. Our model takes an approach closer to that of Jorgenson (2006) and is derived from the theoretical foundation proposed by Ehrhardt-Martinez (1998), where the rate of forest area change is driven by competing land uses (e.g., agriculture), development (e.g., economic growth), and improvements in social capital (e.g., changes in basic education). We posit that investment in human capital (e.g., through secondary school enrollment) is at the root of improvements in national-level political and institutional conditions. Hence, it is omitted, as the simultaneous inclusion of institutional, educational, and economic variables could have derived in issues of multicollinearity and endogeneity (Dias and Tebaldi 2012).

## Results

GLMM estimation showed statistically significant coefficients for all covariates and their squared values at a 1% type-I error level (Table 3). When the largest absolute value of standardized (linear and squared) coefficients for each variable are used to denote relative covariate importance, agriculture land proportion explained most of the variation in annual forest cover rates from 1992 to 2012. It exhibited the largest absolute value at 39.689 for the standardized coefficients of ALP and $ALP^2$. The covariates for population density and rural population proportion were the second and third most important, followed by secondary school enrollment. Income per capita and lagged economic and population growth rates had the smallest standardized coefficients but nonetheless were still statistically significant.

**Table 3. Estimated fixed coefficients for model covariates.†**

| Covariates | Coefficients (β)* | Standardized coefficients |
|---|---|---|
| Fixed intercept | 0.551 | –1.139 |
| $ALP_{it}$ | –0.044 | –39.689 |
| $ALP^2_{it}$ | 2.775 | 20.182 |
| $GDP10_{it}$ | $0.045 \times 10^{-3}$ | 0.010 |
| $GDP10^2_{it}$ | 0.053 | 0.035 |
| $GNIC_{it}$ | 0.005 | 3.181 |
| $GNIC^2_{it}$ | –0.655 | –2.121 |
| $POG_{it}$ | –0.030 | –1.558 |
| $POG^2_{it}$ | 167.000 | 4.422 |
| $POPD_{it}$ | –0.0017 | –14.865 |
| $POPD^2_{it}$ | 0.038 | 36.840 |
| $RPO_{it}$ | 0.0070 | 6.660 |
| $RPO^2_{it}$ | –1.625 | –14.607 |
| $SCH_{it}$ | –0.0067 | –8.899 |
| $SCH^2_{it}$ | 0.362 | 6.769 |
| $SCH15_{it}$ | –0.005 | –7.174 |
| $SCH15^2_{it}$ | 0.235 | 4.106 |

†Squared values were divided by 10,000 to facilitate convergence. *All coefficients are significant at 1% Type-I error level. Log-likelihood= 352,340; Chi-square = 363,742.7.

Figure 2 plots predicted and actual observed forest cover rates for selected nations to illustrate the fitness of the data to the beta-logistic model. Predicted values correctly reflect an increase in forest cover in China, negligible change in the United States, and declines in Brazil and Uganda. The prediction for the United States in Figure 2 is typical of other high-income countries in our data set such as Finland and Germany. The predictions for Brazil and China are representative for large and medium-sized nations such as Ethiopia, India, Indonesia, and the Russian Federation. The prediction for Uganda is one of the worst in our model. This is partly explained by the model allocating a small weight to its observations due to the country's smaller land area.

Marginal associations calculated using Equation (4) for year 2012 are plotted in Figure 3. This year was chosen because it is the most recent year for which complete values were available for most countries (62) in our sample. These measured the mean association between changes in forest cover rates and a unit increase in covariates in 2012. Marginal values between the covariates in agricultural land proportion in 2012 are negative for all countries and have an upward trend. The signs of marginal associations for other covariates have both positive and negative values, reflecting the flexibility

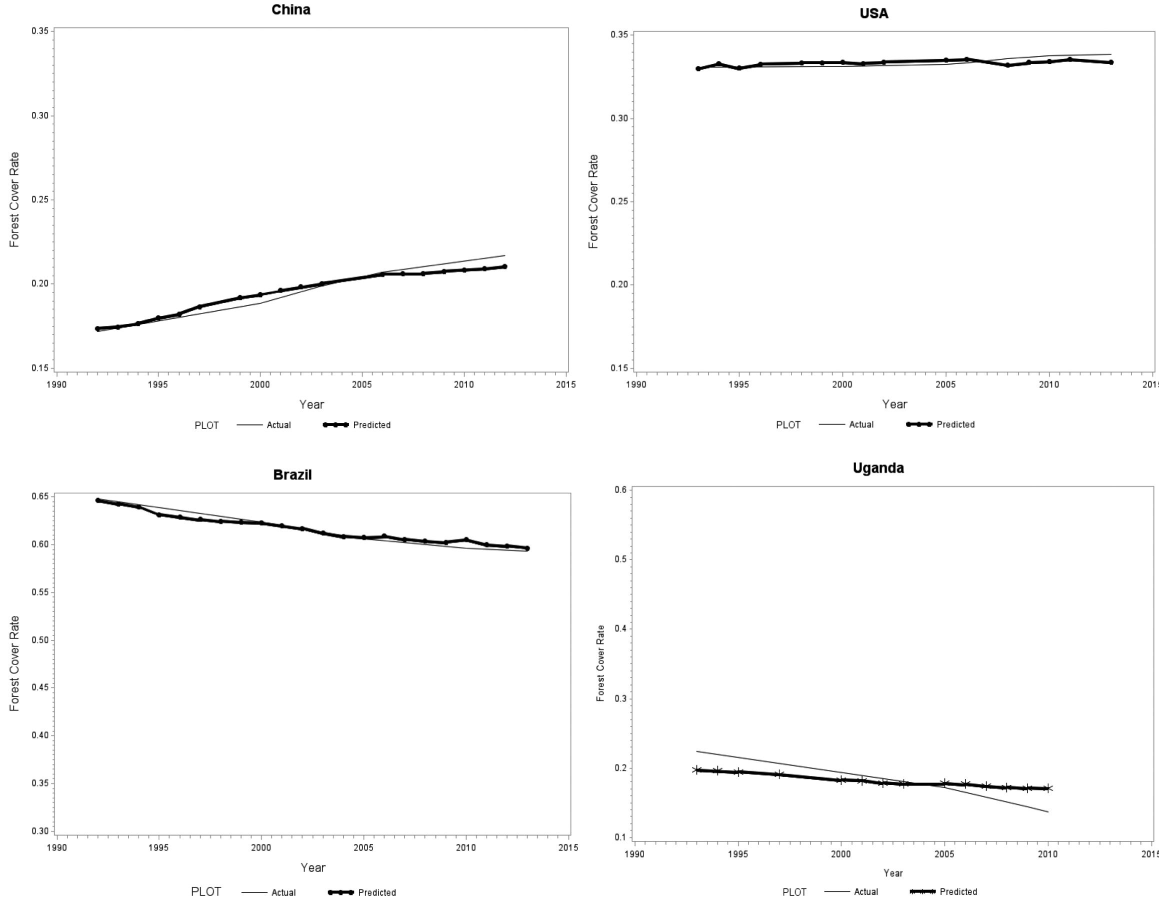


**Figure 2. Plots of reported and predicted forest cover rates for selected countries from 1992 to 2013. Data source: World Bank (2016) and authors' estimation.**

of the model. The marginal associations between forest cover rates and population density, secondary school enrollment rate, and its 15-year lagged variable all consistently had positive trends. To the contrary, rural population proportion had an inverse association. The marginal associations for income per capita, 10-year lagged GDP growth rate ago, and 15-year lagged population growth rate were weaker.

The marginal association for population density was negative for most nations with less than 220 people/km$^2$ and positive for the 10 most densely populated. The marginal association of rural population proportion was positive for countries whose respective value was below 22% and negative for others. The marginal association of secondary school enrollment rate was positive when the rate of enrollment of a country was higher than 93% but negative for other countries, and the switch point for the sign of marginal association to change was 110 for the covariate that represents secondary school enrollment rate lagged 15 years.

Average marginal associations weighted by land area, shown in Table 4, represent a global marginal effect of the eight covariates at the beginning and end of our data period. Negative average coefficients for agricultural land area, population density, rural population rate, and school enrollment rate lagged 15 years in 1992 and 2012 imply that nations' forest cover rates were negatively associated to these three covariates at a global scale over a 20-year period. The sign of the covariates capturing secondary education enrollment rate changed from negative in 1992 to positive in 2012. They could suggest a gradual change from an inverse to a direct association between forest cover and education levels as average rates of secondary schooling increased from 58.97% to 91.97%.

## Discussion

The flexibility of the beta-logistic model allowed fitting well the data and helped capture the complexity of forest cover dynamics. The beta-logistic model restrained the response values within 0 and 1 to allow predictions within reasonable boundaries. The model correctly predicted differing trends (e.g., a decline of forest cover in Brazil and growth in China) even for nations at similar stages of economic development. Random intercepts captured between-country effects, and fixed coefficients captured global associations between forest cover rate and covariates. For example, the coefficient for agriculture land proportion can be translated as a global average decline of 0.466% (Table 4) in forest cover associated to a 1% increase in new agricultural land in 2012. However, this value cannot be interpreted as differences between nations.

The negative association between the percent values of agricultural land areas and forest cover rates reflects competition between agricultural and forested lands. We found this important effect, although others (e.g., Leblois et al. 2017) did not find a direct association between agricultural land and deforestation when studying developing countries. Nevertheless, this conclusion is consistent with past studies that report conversion of forests to agriculture land as a major driver of forest losses (e.g., Bonilla-Moheno et al.

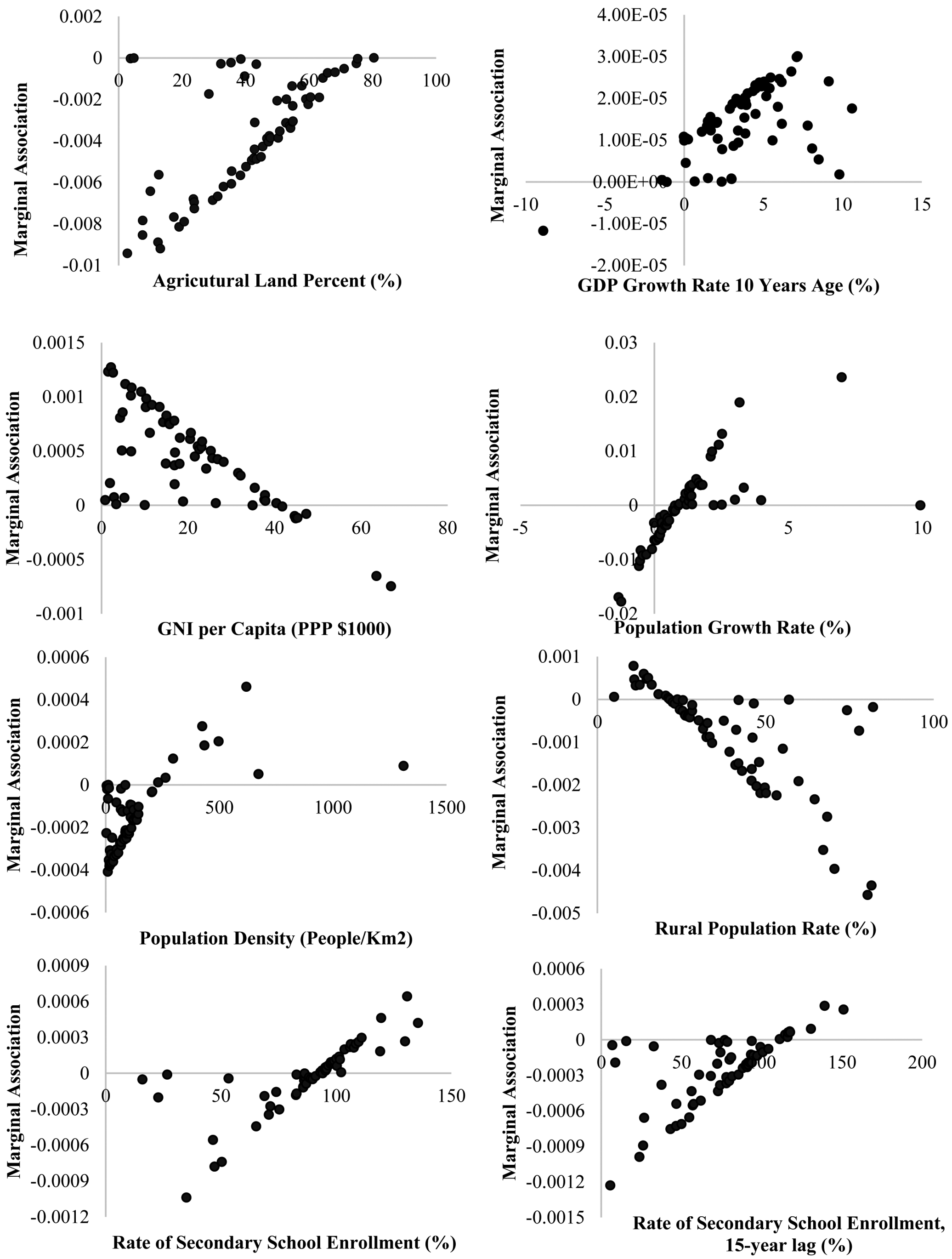


**Figure 3. Plot of estimated marginal associations (effects) of statistically salient covariates, based on the absolute values of standardized coefficients, and forest cover rates for 62 individual countries in the year 2012.**

**Table 4. Average values of covariates and their estimated marginal associations with forest cover rates weighted by land area for years 1992 and 2012.**

| Covariates | Weighted averages of marginal effects in 1992 (75 countries) | Weighted averages of marginal effects in 2012 (62 countries) |
|---|---|---|
| $ALP_{it}$ | –0.00593 | –0.00466 |
| GDP10 | 0.00001 | 0.00002 |
| $GNIC_{it}$ | 0.00063 | 0.00052 |
| POG | 0.00455 | –0.00067 |
| $POPD_{it}$ | –0.00023 | –0.00022 |
| $RPO_{it}$ | –0.00124 | –0.00053 |
| $SCH_{it}$ | –0.00018 | 0.00002 |
| $SCH15_{it}$ | –0.00045 | –0.00023 |

2012, Giri et al. 2015). The relatively weak association between forest cover rate and GNI per capita as well as that between forest cover and GDP growth rate means that income is not a salient factor behind changes in national-level forest rates. Coefficients in Table 3 represent a forest cover rate, as a concave function of income per capita does not support an environmental Kuznets curve trajectory (Stern 2004).

The quadratic terms in the model helped capture nonlinear relationships (Van and Azomahou 2007). This is illustrated by the case of the covariate for population density that had direct and inverse associations along its continuum. One potential explanation for this shift in directional association could be that the imperiled natural environments in countries of high population density alarm their people and encourage efforts to restore lost forestlands (Ehrhardt-Martinez 1998). Another potential explanation is that greater population density initially created new land pressures through forests converted to dwellings and agricultural lands, but once higher levels of population density were reached, for example, through urbanization, that trend was reversed (Barbieri and Carr 2005). The association between population and forest cover offers a parallel for a potential population density Kuznets trend as an element of development beyond income levels (Mather et al. 1999).

It also seems reasonable that education had a negative association with forest cover for countries with low secondary school enrollment and positive for others. We posit that countries with lower levels of education, reflecting human capital, saw an exacerbation of deforestation. Plausibly, as education levels increased, people were initially equipped with new knowledge and with the expectation of higher incomes that translated into greater conversion of forests into agricultural lands to achieve such goals. Eventually, once a country's population reached a high level of education, the trend reversed and forest covers once again increased. Moreover, as a country gets more developed, fewer people rely on agriculture as a major income source and might be more concerned about their conservation and sustainable management. Hence, countries of higher-tier education levels tend to reach greater forest preservation and regeneration (Ehrhardt-Martinez 1998). Better education may also lead to high productivity of agriculture, hence there would be less pressure to reallocate forests to agricultural lands (Brueckner 2000).

The estimated model could be used to predict future forest cover conditions. Agricultural land area is declining in most high-income countries but increasing in others. Population density of high-income countries grows at a much slower pace than others. Moreover, the share of rural population in high-income countries has remained relatively stable over the 1992–2012 period compared to those of some middle-income countries (e.g., China and Vietnam) that are experiencing fast urbanization and recover of forest land area. If these trends continue, our model predicts that the forest cover rates of wealthy countries will continue to increase slowly, and most other countries will continue to convert forest into other uses, except several fast-urbanizing nations.

Further studies are needed to find covariates that could help predict dramatic forest cover changes in some countries beyond country-specific fixed effects, for instance, plots of predicted values for some countries like Uganda (Figure 2) as well as large countries. The model also failed to predict the dramatic decline of forest cover for Indonesia and Uganda (Figure 3). One potential solution to better capture those effects would be the inclusion of random slopes for selected covariates or autoregressive terms in future analyses.

## Conclusion

A beta-logistic GLMM was estimated to quantify the associations between forest cover rates, land, and socioeconomic covariates selected from the literature. Historical changes of forest cover rates were predicted well for most countries in a data set covering 158 nations from around the world over the period 1992–2013.

The estimated model partly confirmed the conclusion of previous studies about associations between forest cover and agriculture land uses, population growth, urbanization, and education. Agricultural land area change was found to be the most closely associated covariate with forest cover rate. Population density, rural population proportion, and education for the current and the previous generation (through a 15-year lagged variable for secondary education enrollment) were found to be the second most important covariates. Income level in terms of GNI and GDP growth rate lagged 10 years, and population growth rate were comparatively less important. We found no support for an environmental Kuznets curve under the setting of our model. However, we did observe a Kuznets-like trajectory in the association of forest cover with population and education variables. Over the 1992–2013 period, higher levels of population density had a negative effect on forest cover up to 220 people/km$^2$, after which its association reversed to a positive one. The threshold for education level was 93% when its marginal effect changed from negative to positive.

We expect high-income countries to slowly increase their forest cover in the future, but most other countries will likely keep converting forestland to other uses. However, some fast-urbanizing developing nations can be expected to increase their forest cover at a faster pace than high-income countries.

## Endnote

1. We refer to marginal associations because coefficients cannot conclusively denote causal relationships.

## Literature Cited


Angelsen, A. 2010. Policies for reduced deforestation and their impact on agricultural production. *Proc. Natl. Acad. Sci*. 107(46):19639–19644.

Assefa, E., and H.R. Bork. 2014. Deforestation and forest management in Southern Ethiopia: Investigations in the Chencha and Arbaminch areas. *Environ. Manage*. 53(2):284–299.

Auty, R.M. (ed.). 2001. *Resource abundance and economic development*. Oxford University Press, Oxford.

Barbier, E.B., and J.C. Burgess. 2001. The economics of tropical deforestation. *J. Econ. Surv*. 15(3):413–433.

Barbieri, A.F., and D.L. Carr. 2005. Gender-specific out-migration, deforestation and urbanization in the Ecuadorian Amazon. *Global Planet. Change* 47(2):99–110.

Bhattarai, M., and M. Hammig. 2001. Institutions and the environmental Kuznets curve for deforestation: A cross country analysis for Latin America, Africa and Asia. *World Dev*. 29(6):995–1010.

Bhattarai, M., and M. Hammig. 2004. Governance, economic policy, and the environmental Kuznets curve for natural tropical forests. *Environ. Dev. Econ*. 9:367–382.

Bolker, B.M., M.E. Brooks, C.J. Clark, S.W. Geange, J.R. Poulsen, M.H. Stevens, and J.S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol. Evol*. 24(3):127–135.

Bonilla-Moheno, M., T.M. Aide, and M.L. Clark. 2012. The influence of socioeconomic, environmental, and demographic factors on municipality-scale land-cover change in Mexico. *Reg. Environ. Change* 12(3):543–557.

Bonnor, G.M. 1967. Estimation of ground canopy density from ground measurements. *J. For*. 65(8):544–547.

Breslow, N.E., and D.G. Clayton. 1993. Approximate inference in generalized linear mixed models. *J. Am. Stat. Assoc*. 88(421):9–25.

Brueckner, J.K. 2000. Urban sprawl: Diagnosis and remedies. *Int. Regional Sci. Rev*. 23(2):160–171.

Casella, G., and R.L., Berger. 2002. *Statistical inference* (vol. 2). Duxbury, Pacific Grove, CA.

Caviglia-Harris, J.L., and E.O. Sills. 2005. Land use and income diversification: Comparing traditional and colonist populations in the Brazilian Amazon. *Agr. Econ*. 32(3):221–237.

Chen, D., and K. Clements. 1999. *World consumption economics*. World Scientific Publishing Company Pte. Limited, Singapore.

Chevalier, A., C. Harmon, I. Walker, and Y. Zhu. 2004. Does education raise productivity, or just reflect it? *Econ. J.* 114(499):F499–F517.

Culas, R.J. 2012. REDD and forest transition: Tunneling through the environmental Kuznets curve. *Ecol. Econ*. 29:44–51.

Culas, R.J. 2014. Causes of deforestation and policies for reduced emissions (REDD+): A cross-country analysis. *J. Appl. Eco*. 13(4):7–27.

Dalla-Nora, E.L., A.P.D. de Aguiar, D.M. Lapola, and G. Woltjer. 2014. Why have land use change models for the Amazon failed to capture the amount of deforestation over the last decade? *Land Use Pol*. 39:403–411.

Damette, O., and P. Delacote. 2012. On the economic factors of deforestation: What can we learn from quantile analysis? *Econ. Model*. 29(6):2427–2434.

Debertin, D.L. 2012. *Agricultural production economics (2nd ed.)*. Amazon Createspace, Lexington, KY.

De Sá, S.A., C. Palmer, and S. Di Falco. 2013. Dynamics of indirect land-use change: Empirical evidence from Brazil. *J. Envir. Econ. Manage*. 65(3):377–393.

Dias, J., and E. Tebaldi. 2012. Institutions, human capital, and growth: The institutional mechanism. *Struct. Change Econ. D*. 23(3):300–312.

Edelson, S.M. 2007. Clearing swamps, harvesting forests: trees and the making of a plantation landscape in the colonial South Carolina low country. *Agr. Hist*. 81(3):381–406.

Ehrhardt-Martinez, K. 1998. Social determinants of deforestation in developing countries: A cross-national study. *Soc. Forces* 77(2):567–586.

Ehrhardt-Martinez, K., E. Crenshaw, and J.C. Jenkins. 2002. Deforestation and the environmental kuznets curve: a cross-national investigation of intervening mechanisms. *Social Science Quarterly*. 83(1):226–243.

FAO (Food and Agriculture Organization of the United Nations). 2015. *Global Forest Resources Assessment* 244 p.

Ferrari, S., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *J. Appl. Stat*. 31(7):799–815.

Geist, H.J., and E.F. Lambin. 2002. Proximate causes and underlying driving forces of tropical deforestation tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *BioScience* 52(2):143–150.

Giri, C., J. Long, S. Abbas, R.M. Murali, F.M. Qamer, B. Pengra, and D. Thau. 2015. Distribution and dynamics of mangrove forests of South Asia. *J. Envir. Manage*. 148:101–111.

Jennings, S.B., N.D. Brown, and D. Sheil. 1999. Assessing forest canopies and understory illumination: Canopy closure, canopy cover and other measures. *For*. 72(1):59–74.

Johansson, T. 1985. Estimating canopy density by the vertical tube method. *Forest Ecol. Manag*. 11:139–144.

Jorgenson, A. 2006. Unequal ecological exchange and environmental degradation: A theoretical proposition and cross-national study of deforestation, 1990–2000. *Rural Sociol*. 71(4):685–712.

Keenan, R., G. Reams, F. Achard, J. de Freitas, A. Gringer, and E. Lindquist. 2015. Dynamics of global forest area: results from the fao global forest resources assessment 2015. *Forest Ecol. Manag*. 352(7):9–20.

Koop, G., and L. Tole. 1999. Is there an environmental Kuznets curve for deforestation? *J. Dev. Econ*. 58:231–244.

Koop, G., and L. Tole. 2001. Deforestation, distribution and development. *Global Environ. Chang*. 11(3):193–202.

Korhonen, L., K.T. Korhonen, P. Stenberg, M. Maltamo, and M. Rautiainen. 2007. Local models for forest canopy cover with beta regression. *Silva Fennica*. 41(4):671.

Kutner, M.H., C.J. Nachtscheim, J. Neter, and W. Li. 2004. *Applied linear statistical models* (5th ed.), McGraw Hill/Irwin, Chicago.

Leblois, A., O. Damette, and J. Wolfersberger. 2017. What has driven deforestation in developing countries since the 2000s? Evidence from new remote-sensing data. *World Dev*. 92:82–102.

Lindstrom, M.J., and D.M. Bates. 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics* 46:673–687.

List, J., and C. Gallet. 1999. The environmental Kuznets curve: Does one size fit all? *Ecol. Econ*. 31(3):409–423.

Liu, J., T. Dietz, S. Carpenter, M. Alberti, C. Folke, E. Moran, A. Pell, P. Deadman, T. Kratz, J. Lubchenco, E. Ostrom, Z. Ouyang, W. Provencher, C. Redman, S. Schneider, and W. Taylor. 2007. Complexity of coupled human and natural systems. *Science* 317:1513–1516.

Long, J.S., and J. Freese. 2006. *Regression models for categorical dependent variables using Stata*. Stata Press, College Station, TX.

López-Carr, D., and J. Burgdorfer. 2013. Deforestation drivers: Population, migration, and tropical land use. *Environment: Science and Policy for Sustainable Development* 55(1):3–11. Available online at http://www.tandfonline.com/loi/venv20.

MacDicken, K.G. 2015. Global forest resources assessment 2015: What, why and how? *Forest Ecol. Manag*. 352:3–8.

Mather, A., C. Needle, and J. Fairbairn. 1999. Environmental kuznets curves and forest trends. *J. of the Geog. Assoc*. 84: 55–65.

Meyfroidt, P., E.F. Lambin, K.H. Erb, and T.W. Hertel. 2013. Globalization of land use: Distant drivers of land change and geographic displacement of land use. *Curr. Opin. Env. Sust*. 5(5):438–444.

Moomaw, R.L. 1988. Agglomeration economies: Localization or urbanization? *Urban Stud*. 25(2):150–161.

Neher, P.A. 1990. *Natural resource economics*. Cambridge University Press, Cambridge.

Rautiainen, M., P. Stenberg, and T. Nilson. 2005. Estimating canopy cover in Scots pine stands. *Silva Fennica* 39(1):137–142.

Restuccia, D., D.T. Yang, and X. Zhu. 2008. Agriculture and aggregate productivity: A quantitative cross-country analysis. *J. Monetary Econ*. 55(2):234–250.

Roca, J. 2003. Do individual preferences explain the environmental Kuznets curve? *Ecol. Econ*. 45(1):3–10.

Rudel, T.K. 2007. Changing agents of deforestation: From state-initiated to enterprise driven processes, 1970–2000. *Land Use Policy* 24(1):35–21.

Schabenberger, O., and F.J. Pierce. 2001. *Contemporary statistical models for the plant and soil sciences*. CRC Press, Boca Raton, FL.

Schluchter, M.D., and J.T. Elashoff. 1990. Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures. *J. Stat. Comput. Sim*. 37(1–2):69–87.

Scrieciu, S.S. 2006. Can economic causes of tropical deforestation be identified at a global level? *Ecol. Econ*. 62:603–612.

Stern, D. 2004. The rise and fall of the environmental Kuznets curve. *World Dev*. 32(8):1419–1439.

Stern, D.I., M.S. Common, and E.B. Barbier. 1996. Economic growth and environmental degradation: The environmental Kuznets curve and sustainable development. *World Dev*. 24(7):1151–1160.

Swinton, S.M., G. Escobar, and T. Reardon. 2003. Poverty and environment in Latin America: Concepts, evidence and policy implications. *World Dev* 31(11):1865–1872.

Tadesse, G., E. Zavaleta, C. Shennan, and M. FitzSimmons. 2014. Policy and demographic factors shape deforestation patterns and socio-ecological processes in southwest Ethiopian coffee agroecosystems. *Appli. Geogr*. 54:149–159.

United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics. 2016. Total secondary gross enrolment ratio. http://data.un.org/Data.aspx?q=enrollment+secondary+&d=UNESCO&f=series%3aGER_23, accessed August 10, 2016.

United Nations Statistic Division. 2016. UNdata. Available online at http://data.un.org/Data.aspx?d=POP&f=tableCode%3A22, accessed August 2016.

Van, P., and T. Azomahou. 2007. Nonlinearities and heterogeneity in environmental quality: An empirical analysis of deforestation. *Journal of Development Economics* 84(1):291–309.

Vogt, K., R. Gara, J. Honea, D. Vogt, T. Patel-Wynand, P. Roads, A. Fanzeres, and R. Sugurdardottir. 2007. Historical perceptions and uses of forests. P. 1–29 in *Forests and society: Sustainability and life cycles of forests in human landscapes,* CABI, London.

World Bank. 2016. Indicators. Available online at http://data.worldbank.org/indicator. Accessed August 2016.

Zeder, M.A. 2008. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc. Natl. Acad. Sci.* 105(33):11597–11604.