The Evolving Relationship between Market Access and Deforestation on the Amazon Frontier

Yu Wu
Department of Economics, Department of Agricultural and Resource Economics
North Carolina State University
ywu22@ncsu.edu

and

Erin O. Sills
Department of Forestry and Environmental Resources
North Carolina State University
erin_sills@ncsu.edu

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-7, 2018

Abstract

Recognition of tropical deforestation as a major source of carbon emissions has invigorated efforts to reduce deforestation in the Brazilian Amazon. Previous studies throughout the tropics have found that market access, one of the key drivers of deforestation, is positively correlated with deforestation. This relationship has been explained by the effect of market access on farm gate prices for agricultural and/or timber products. However, access to markets also means access to urban centers, which has multiple influences on household decisions. This study incorporates these multiple dimensions of "urban access" into a household production framework, showing that their combined effect on deforestation is ambiguous and could change over time. We test whether the effect of distance to cities has changed over time in a region typical of the agrarian reform settlements established by the Brazilian federal government throughout the Amazon. Specifically, we test for structural shifts in the relationship between distance to cities and deforestation by modeling deforestation over 24 years on approximately 8,800 farm properties. Estimation results for a two-part, "within-between" Random Effects model show that after controlling for forest stock and other drivers of deforestation, the relationship between distance to cities and deforestation changed from negative to positive around 2004. Further investigation suggests that the diminishing role of transportation cost in farm gate price combined with the increasing enforcement of environmental laws forced the effect of distance on deforestation from negative to positive. Regional heterogeneities underlying the general evolving relationship are consistent with observed variation in urban development.

1. Introduction

The deforestation of the Amazon rainforest is occurring mainly in Brazil, where about 62% of the rainforest is located, and where about 80% of deforestation has occurred (Hansen et al. 2013). From 1988 to 2004, Brazil had the world's highest rate of deforestation, averaging 18,400 km^2 per year (INPE 2017). After a peak in 2004, Brazil also had the largest reduction in deforestation (and carbon emissions) of any country through 2012; and then rates again increased to 7,893 km^2 in 2016 (INPE 2017)). As recognition has grown that tropical deforestation is a major source of carbon emissions, efforts have been made to reduce deforestation in the Brazilian Amazon in the 2000s. Yet there remain questions about whether these efforts have been effective at reducing deforestation and which efforts have been most effective. One challenge to crafting the best policy response is that decision-making and drivers of deforestation change over time as the frontier evolves.

There is a large body of research on the drivers of deforestation and land-use decisions by landowners in the Amazon. The standard theoretical framework posits that the probability of deforestation in any given location is a function of its biophysical characteristics (Ricardian theory) and its market access (von Thünen theory) (Pfaff et al. 2013). Von Thünen theory relates land rents to transportation costs, and to the costs of establishing land tenure (Von Thünen 1966; Angelsen 2007). In the context of the Amazon, this theory has been posited to explain observed patterns of deforestation through two channels: higher price of agricultural outputs (at the farmgate) increases the derived demand for agricultural land (which is obtained by clearing forest); and higher profitability encourages investments in protecting land tenure, including through deforestation (Schneider 1993).

The empirical literature has generallyused location as a proxy for farm gate price, and has represented market access using roads or the distance to urban center (Sills 2014). Many previous studies have found a negative relationship between this distance and deforestation (Chomitz and Gray 1996; Geoghegan et al. 2004; de Souza Soler et al. 2010; Caviglia-Harris and Harris 2011). This result is consistent with the von Thünen theory of land rent, which suggests that the rents to agricultural lands are highest closest to markets (Sills and Caviglia-Harris 2009), thus creating an incentive to clear those lands first. Some studies have tried improved proxies for market access; for example, instead of using distance to the nearest road, researchers used road distance to markets over the relevant transportation network (Aguiar et al. 2007; Mann et al. 2010).

The negative effect of distance to urban center on deforestation has typically been interpreted as showing the expected relationship between the farm-gate prices of outputs and the derived demand for agricultural land. However, access to urban centers may have multiple influences on household decisions. For example, proximity to urban centers could mean better off-farm employment opportunities, and therefore a higher opportunity cost for household labor used in deforestation or agriculture. Proximity to urban centers also means that environmental enforcement agents based in those cities have better access to the farm. As explored in the regional economics and rural development literature, researchers are recognizing new and growing connections between urban and rural space in the Amazon (Simmons et al. 2002; Padoch et al. 2008).

This paper incorporates these multiple dimensions of urban access into a household production framework, showing that their combined effect on deforestation is ambiguous and could change over time. We examine the evolution by modeling deforestation from 1985 to 2009 on approximately 8,800 farm properties typical of the agrarian reform settlements that have been established by the federal agency, the National Institute for Colonization and Agrarian Reform (INCRA), throughout the Amazon. To the best of our knowledge, this is the first study that analyzes the long-term evolution of deforestation at the property level in the Amazon. We test for structural shifts in the effect of urban access on deforestation in the agrarian reform settlements where both deforestation and the farming population are concentrated. Tenure insecurity is less for an issue in this region, which brings into sharper focus the other possible reasons that distance to urban center might matter.

2. Theoretical Framework

To place market access in the context of other drivers of deforestation and identify its possible roles in decisions about deforestation, this section presents a household production model. Farm households in agricultural settlements in the Brazilian Amazon are integrated production and consumption units that rely primarily on their own labor for production and consume at least some portion of their farm production. In this setting, the household production model is an appropriate framework for understanding integrated household decisions about production, consumption, and labor supply (Sadoulet and de Janvry 1995; Singh et al. 1986). We do not aim to set up a universal model of deforestation, but rather a model appropriate to agricultural settlements, where each settler household is allocated a given area of forest that they can put into agricultural production by allocating labor to deforestation.

We start with an intertemporal model of unitary decision-making. This is because deforestation is a fundamentally intertemporal process, with forest cleared in the present to obtain agricultural land used in multiple future periods. We assume unitary household decision-making, because the data available to estimate the model are from household surveys that make the same assumption. Thus, "household" decision-making refers either to the choices of the head of household (typically a man) or some other process of collective decision-making within the household that we do not explicitly model. Assuming the "household" cares about utility over an infinite time horizon (because he cares about future generations), he maximizes the utility from consumption of self-produced agricultural goods C_t^A, market goods C_t^M and leisure T_t^l over time, conditioned on household characteristics z_t^C that influence preferences, and subject to the production function, the time (labor) constraint, the accumulation of agricultural land, the land constraint, and the budget constraint[1]:

$$\max_{C_t^A, C_t^M, T_t^A, T_t^D, T_t^O, I_t^A} E_t \sum_{t=1}^{\infty} \beta^t U\left(C_t^A, C_t^M, T_t^l; z_t^C\right) \tag{1}$$

s.t.

$$Q_t^A = Q(T_t^A, L_t^A, I_t^A; z_t^L) \tag{2}$$

$$T_t = T_t^A + T_t^D + T_t^O + T_t^l \tag{3}$$

$$L_t^A = L_{t-1}^A + D_t \tag{4}$$

$$L_t^A \leq \overline{L} \tag{5}$$

$$(P_t^A - x_t^A)C_t^A + (P_t^M + x_t^M)C_t^M + f_t\left[\frac{L_t^A}{\overline{L}} - (1 - r_t)\right]$$

$$= \left[(P_t^A - x_t^A)Q_t^A - (P_t^I + x_t^I)I_t^A\right] + (W_t^O - x_t^O)T_t^O + S_t \tag{6}$$

Equation (2) represents the production function of agriculture good Q_t^A in the period *t*, which is a function of the household's time (labor) devoted to agriculture production T_t^A, the area of land on agriculture L_t^A, and other purchased inputs I_t^A, conditioned on land characteristics z_t^L. The time devoted to agricultural production T_t^A in the period *t* is limited by the household time

[1] We omit the subscript i for household from the presentation of the model.

constraint, equation (3). Besides agricultural production, the household allocates its total amount of time T_t to deforestation T_t^D, off-farm work T_t^O, and leisure T_t^l.

The agriculture land accumulation is demonstrated by equation (4). We treat the process of agricultural land accumulation as the evolution of capital, and deforestation is the investment. The area of land devoted to agricultural in the current period, L_t^A, is derived from the old agricultural land in the last period, L_{t-1}^A, plus the new created agriculture land through deforestation in the current period, D_t (Pendleton and Howe 2002), which is a function of labor allocated to deforestation, T_t^D.

Equation (5) represents the land constraint. Given the land tenure in our empirical case is secure and the total land size is fixed for each household, neither the area of agricultural land nor the area of deforestation can exceed the exogenously given property size $\overline{L}$. Also, deforestation is non-negative, since it measures the net loss of primary forest (old-growth forest) in our study.

Beyond the land constraint, deforestation is limited by policy enforcement[2]. To reflect this, we incorporate the possibility of a "fine" [3]for deforestation beyond the legal limit on each property. Specifically, the Brazilian Forest code requires that forestland owners retain a minimum percentage of their land in a "Legal Reserve" of forest. For each percent of land used for agriculture beyond the legal maximum $(1 - r_t)$, where r_t is the Legal Reserve requirement or percentage of

---

[2] The Brazilian Forest Code (law) has required landowners to maintain 20 to 80 percent of their rural properties under native vegetation in different time periods and different regions (Soares-Filho et al. 2014). The first Forest Code (Federal Decree No. 23793) was created and passed in 1934 requiring 25 percent of the property must be conserved with forest as a Legal Reserve (LR). The conservation requirements for LRs and the stringency of enforcement changed over our study period.

[3] The penalty for excessive deforestation includes confiscation of assets, monetary fines, or conditional access to credit and commercialization channels, but historically the monetary fines has been rarely paid (Börner et al. 2015). After 2004, under improved monitoring for enforcement, fines were issued with satellite images showing the property boundaries and the existing deforestation (Fearnside 2005). Under the mechanism of the Environmental Reserve Quota (CRA, Portuguese acronym) introduced nationally in 2012, landowners are required to georeferenced their property boundaries and remaining forests, and failed to do so will result in loss of access to credit and output markets (Azevedo et al. 2017).

property that must be conserved with forest in the period *t*, the household risks being fined for breaking the law and illegally clearing forest. Let f_t represent the expected monetary amount of the fine in the period *t* (i.e. the fine for excess deforestation multiplied by the household's expected probability of being fined), which is the shadow cost of expanding agriculture land beyond the legal allowed amount (Assunção et al. 2015). Therefore, the expected expense associated with illegal deforestation is equal to $f_t[\frac{L_t^A}{\bar{L}} - (1 - r_t)]$, which is included in the budget constraint equation (6). While there is an emerging market for forest conservation credits (Soares-Filho et al. 2016), there was no award for maintaining excess Legal Reserve in Rondônia during our study period. Thus, $f_t = 0$ when $[\frac{L_t^A}{\bar{L}} - (1 - r_t)] \leq 0$.

In the budget constraint equation (6), the household's total expenses equal its total income in the period *t*. In addition to any fines for expanding agricultural land beyond legal limits $f_t[\frac{L_t^A}{\bar{L}} - (1 - r_t)]$, the household uses its income to consume self-produced agricultural goods C_t^A and other market goods C_t^M. The sales price received by a household for any agricultural goods not consumed (the farm-gate price) equals the market price of the agricultural goods P_t^A minus a transaction cost x_t^A; and the purchase price of other market goods equals the market price of the market goods P_t^M plus a transaction cost x_t^M incurred in buying (Sadoulet and de Janvry 1995; Key, Sadoulet and de Janvry 2000). While transaction costs include all costs of measuring what is being exchanged and enforcing agreements (North 1994), costs associated with distance to market are among the most substantial in the agricultural settlements in the Amazon. These include not only the cost of transporting goods and people, but also potentially fewer market players (allowing intermediaries to capture higher margins) and less information available (leading to higher search and recruitment costs) further from market (Sadoulet and de Janvry 1995). The major household income source is agricultural product sales. The household sells Q_t^A amount of agricultural products at the farm gate price $(P_t^A - x_t^A)$, coincident with the expenses of purchased inputs[4] $(P_t^I + x_t^I)I_t^A$. The household could also earn income from off-farm jobs either on other farms or in the urban centers. In the case of work in urban centers, the household bears the transaction costs of traveling to work, and the take home wage is $(W_t^O - x_t^O)$. In the meanwhile, the household gains exogenous transfer S_t (e.g. government welfare payments).

Following von Thünen theory, the transaction cost is dominated by transportation cost and is assumed to be a function of distance to urban center (Chomitz and Gray 1996). Holding everything else constant, an increase in distance will decrease the farm gate price of agricultural product; and decrease the take-home wage of off-farm work:

$$P_t^A - x_t^A = p_t^A(d) \tag{7}$$

$$W_t^O - x_t^O = p_t^O(d) \tag{8}$$

where $\frac{\partial p_t^A(d)}{\partial\, d} < 0$, and $\frac{\partial p_t^O(d)}{\partial\, d} < 0$.

Furthermore, the policy stringency parameter or shadow cost of illegal deforestation, f_t, can be written as a function of distance as well. Because law enforcement personnel are typically based in urban centers and often face transportation constraints, landowners may expect a higher risk of being caught for illegal clearing or burning if their land is closer to an urban center. This suggests that the shadow cost of illegal deforestation f_t is negatively related to the distance:

$$f_t = p_t^f(d) \tag{9}$$

where $\frac{\partial p_t^f(d)}{\partial\, d} < 0$.

[4] The coincident costs of agricultural production also include own labor on agricultural production and on deforestation. However, wage payments and earning for own labor would cancel out in the budget constraint, and thus for simplicity, we only include the cost of purchased inputs in the equation.

Substituting (7-9) into the first order condition with respect to the labor allocated to deforestation yields the marginal agricultural output of deforestation in the period t:

$$\frac{\partial Q_t^A}{\partial D_t} = \frac{1}{\frac{\partial U}{\partial c_t^A}} \left\{ \left[\frac{\frac{\partial U}{\partial c_t^A}}{\bar{L}} \frac{p_t^f(d)}{p_t^A(d)} + \mu_t \right] + \beta E_t \left[\frac{\frac{\partial U}{\partial c_{t+1}^A}}{\bar{L}} \frac{p_{t+1}^f(d)}{p_{t+1}^A(d)} + \mu_{t+1} - \frac{\partial Q_{t+1}^A}{\partial D_t} \frac{\partial U}{\partial C_{t+1}^A} \right] + \frac{\frac{\partial U}{\partial c_t^A}}{\frac{\partial g(T_t^D)}{\partial T_t^D}} \frac{p_t^O(d)}{p_t^A(d)} \right\} \tag{10}$$

where μ_t is the shadow price for the land evolution, equation (4).

Equation (10) incorporates the multiple dimensions of distance to urban center and reveals their different effects on the agricultural land expansion (deforestation) decision. We assume that additional agricultural land converted from forest increases agricultural production at a decreasing rate (Pendleton and Howe 2002), that is, $\frac{\partial Q_t^A}{\partial D_t} > 0$, $\frac{\partial^2 Q_t^A}{\partial^2 D_t} < 0$. So, higher marginal agricultural output of deforestation implies lower deforestation. Holding all things equal, an increase in distance to urban center decreases the shadow cost of illegal deforestation and the farm gate price of agricultural goods at the same time. Therefore, the change of the ratio $\frac{p_t^f(d)}{p_t^A(d)}$ is ambiguous, and the change of deforestation level is uncertain. Similarly, simultaneously decreases in the off-farm wage and the farm gate price lead to an uncertain change in the ratio $\frac{p_t^O(d)}{p_t^A(d)}$ and the deforestation level. If the effect of distance on farm gate price of agricultural goods dominates, the ratios $\frac{p_t^f(d)}{p_t^A(d)}$ and $\frac{p_t^O(d)}{p_t^A(d)}$ will increase with an increase in distance (a decrease in farm gate price), and lead to a decrease in deforestation. That is, deforestation will be negatively correlated with distance to urban center, which is consistent with von Thünen theory and previous studies. However, if the effect of distance on policy enforcement (the shadow cost of illegal deforestation) or/and on the wage for off-farm work dominates, this would result in a positive relationship between deforestation and distance.

3. Empirical Strategy

The objective of the empirical analysis is to estimate the marginal effect of market access – distance to urban center – on deforestation; to identify whether and how the relationship between market access and deforestation has changed over time, specifically in the context of the Ouro Preto do Oeste region. In the empirical model, the dependent variable is deforestation, measured by annual loss of primary forest cover (deforestation of old-growth forest). Thus, the dependent variable is non-negative; and the larger the dependent variable, the more deforestation. As postulated by the theoretical framework, deforestation is a function of labor allocated to forest clearance, which is determined by the farm gate price of agricultural product, the shadow cost of illegal deforestation, the wage for off-farm work, land area, and characteristics of the land and household. All of the price variables are functions of distance to urban center, which is measured as the road distance to the nearest urban center[5]. Since the extent of the road system did not change for most of our study period and study region, distance is a time-invariant variable. However, farm gate prices of agricultural products still change over time due to variation of market prices, so we include the time-variant state-level market price and its interaction term with distance. In the Ouro Preto do Oeste region, pasture is the dominant agricultural land use, and milk production is the only income source which is regular and expected by households in every period (both wet and dry seasons). Thus, we include the state-level milk price as the most relevant output price. Guided

[5] The urban centers include (1) six municipalities in the Ouro Preto do Oeste region: Ouro Preto do Oeste, Vale do Paraíso, Urupá, Mirante da Serra, Nova União, and Teixeirópolis; (2) the two cities immediately outside the Ouro Preto do Oeste region: Jaru and Ji-Paraná; (3) one district under Ouro Preto do Oeste municipality: Rondominas.

by previous studies, the characteristics of land include age of the property[6], land slope, soil quality, water access, and shape of the property. Our preliminary empirical analysis suggests household characteristics will not affect the production decision making and could be eliminated in the deforestation model.[7] Furthermore, we control for forest stock in the last year and spatial dummy variables for municipalities. To sum up, a general form of the empirical model can be written as:

$$D_{it} = f(d_i, p_t, (d_i \cdot p_t), \overline{L}_i, Z_{it}^L, F_{i(t-1)}, Y_t, M_i, \varepsilon_{it}) \quad (11)$$

where D_{it} is deforested hectares for the property i of the year t; d_i is the distance to the nearest urban center for the property i; p_t is the state level milk price of the year t; $(d_i \cdot p_t)$ is the interaction term between distance and milk price; $\overline{L}_i$ is the lot size for the property i; Z_{it}^L includes the land characteristics mentioned above; $F_{i(t-1)}$ is primary forest cover for the property i in the year $(t-1)$; Y_t represents dummy variables for calendar years; M_i represents dummy variables for the six municipalities; and ε_{it} is the error term.

The dependent variable, deforestation, is semicontinuous, with a continuous distribution except for a probability mass at zero (Olsen and Schafer 2001). The high proportion of zero values makes the normal distribution inappropriate for modelling the data (Min and Agresti 2002). Zero deforestation may occur for two reasons: (1) the farm household chooses not to expand agricultural land in the current period by clearing forest; (2) there is no forest left to clear on the property. Obviously, the second reason for zero deforestation is not driven by the covariates in equation (11) except for last year's forest stock $F_{i(t-1)}$. Therefore, we first selected observations whose forest stock in the last year was not zero[8]. Then we model the household as making decisions in two

[6] Age of the property is years since first primary forest cleared. For properties that had forest cleared before 1985 (the first year of the study period), ages are assigned based on the official settlement records from INCRA.
[7] When we include household characteristics to estimate deforestation in the same model set up using the four-waves survey data, the household characteristics are insignificant in the regression results.
[8] We drop 4896 observations, which are 2.3 percent of the sample population. We also conducted all the regressions with the origin data before dropping these zero forest stock observations, and we received similar results.

steps: first, it determines whether to deforest or not; second, it determines how much area to deforest. These processes can be modeled by a two-part model, such as a double hurdle model or the truncated normal hurdle model introduced by Cragg (1971). In the two-part model, the first stage is a binary model for the dichotomous event of having zero or positive deforestation values. In the second stage, Cragg (1971) suggests a truncated normal distribution to make the dependent variable positive (Wooldridge 2010). However, a simpler version for economic interpretation is: conditional on a positive value of the dependent variable, the second stage follows a lognormal distribution (Duan et al. 1984). Beyond the simplicity, the lognormal distribution model has a well-behaved likelihood function and is typically more robust than the truncated normal model under the model selection test of Vuong (Min and Agresti 2002; Hsu and Liu 2008). Comparing with the selection model, while the two-part model makes the conditional independence assumption, it allows all covariates to appear in both parts and does not require exclusion restrictions (Olsen and Schafer 2001), which is more appropriate in our study case where the decision to deforest and the extent of deforestation are driven by similar processes.

Another common approach to deal with semicontinous data is the tobit model. Although it censors out the observations at zero and cannot interpret the true zeros, it overcomes the possible limitations of the two-part model. These limitations of the two-part model include the bias on estimating small values when zeros are eliminated in the second stage, and the conditional independence assumption. Hence, besides the two-part model, we estimate a tobit model to combine the zero observations regime and positive observations regime into a single log-likelihood function (Engel and Moffatt 2014).

We estimate these models with panel data: twenty-four years of data from 8,793 farm properties. Panel data allow us to track household choices over time and control for unobserved idiosyncratic differences across households. Fixed effects (FE) and random effects (RE) models are the two usual approaches to panel data problem. RE supports estimation of the effects of both time-invariant and time-variant factors, and is thus preferable to FE when its key assumptions hold: exogeneity of covariates and normality of residuals. In practice, the exogeneity of covariates is often rejected in empirical analyses, implying that the FE model is a better option. However, there are no sufficient statistics for probit or tobit model to fit conditioned fixed effects model, and unconditional fixed effects estimates are biased in the probit or tobit model (Greene 2004; StataCorp 2015). In addition, given that time-invariant effects are of central interest in this study (i.e. the distance to urban center), we need an alternative approach to estimate the effects of time-invariant variables and to control for heterogeneity bias at the same time. One well-known approach is the Mundlak adjusted RE model (Mundlak 1978). Mundlak proposed an adjusted RE framework model, which starts from the assumption that heterogeneity bias of the standard RE model is the result of attempting to model "within effects" and "between effects" in one term. So rather than correcting heterogeneity bias, it models the bias by separating the within and between effects. A hybrid model (Allison 2009) or "within-between" RE formulation (Bell and Jones 2015) rearranged the adjusted RE formulation given by Mundlak (1978) to be more interpretable for within and between effects and remove possible collinearity without the risk of heterogeneity bias.

The paper adopts the "within-between" RE model (Bell and Jones 2015) and integrates it into a two-part model. In the first stage, probit model and logit model are used to model the binary outcome. In the second stage, assuming that a zero correlation between the individual-specific error terms in the two stages models and an IHS-normal distribution[9], we conduct a conditional

[9] The inverse hyperbolic sine (IHS) is an alternative transformation approach to handle extreme values which can be defined at negative or zero (Burbidge and Magee 1988):
$$\ln\left(\theta D_{it} + (\theta^2 D_{it}^2 + 1)^{\frac{1}{2}}\right)/\theta = sinh^{-1}(\theta D_{it})/\theta$$

linear regression with the "within-between" RE model. Besides the two-part model, we conduct an adjusted random effects tobit model, which allows for both types of zeros. In the regressions, the dependent variable and all explanatory variables with skewed distributions or extreme values are transformed by inverse hyperbolic sine (IHS) except for dummy variables.

The theoretical framework demonstrates that distance to urban center affects deforestation through multiple channels. As the size and relative influence of these different channels changes over time, the net effect of distance on deforestation could also change, even reversing sign. Thus, we assess whether and when the direction and size of the effect has changed, and then relate that to exogenous factors associated with different channels or mechanisms. By interacting distance with year dummy variables, we start with a test of whether the effect of distance varies by year. If the test results indicate the effect of distance does not remain the same across the study period, then we test if there exist common patterns among years and test for possible break point(s). We utilize two approaches to test the observed (possible) break points. In the first approach, based on the marginal effects of distance reported in the regression, we calculate the average slope of deforestation with respect to distance in pre-break years (a linear combination of marginal effects) and the average slope in after-break years (another linear combination of marginal effects). Then we test the equality of these two multiple linear combinations at the five percent level. In the second approach, we estimate an interrupted model with potential structural break(s) breaking the effect of distance into different groupings. Then we test the equality between sets of coefficients (both intercept and slope) in two multiple regressions (Chow 1960; Lewis-Beck 1986).

---

where $\theta = 1$ in our transformation and the inverse sine is approximately equal to $\ln(2D_{it})$, except for small value (e.g. less than 1), so it can be interpreted in the same way as a natural logarithmic transformation (Pence 2006; Friedline et al. 2015).

4. The Study Area and Data Description

4.1 The Study Area

Our study site, the Ouro Preto do Oeste region of Rondônia, is located in northwestern Brazil on the western end of the "arc of deforestation" across the southern Brazilian Amazon (Figure 1). The climate of the region is humid tropical with average temperatures of 24℃ and annual precipitation of 2300 mm with a distinct dry season in July and August. The native vegetation includes both dense and open tropical forests (INPE 2000; Caviglia-Harris et al. 2009).

Ouro Preto do Oeste witnessed large scale migration and settlement beginning in the early 1970s driven by the federal government's programs of road-building and colonization. The federal government agency, the National Institute of Agrarian Colonization and Reform (INCRA), distributed land free of charge or for trivial loans at minimal or zero interest rates (Sills and Caviglia-Harris 2009). Unlike many other parts of the Amazon, land tenure is secure in the agricultural settlements in our study region (Jones et al. 1995; Sills and Caviglia-Harris 2009). Given these conditions, we model farm households as having a fixed area of land (which cannot be expanded through deforestation of additional land) with exogenous biophysical characteristics determined by INCRA's original allocation of lots.

Migration to the region continued in the 1980s, motivated by relatively fertile soils and easy access along the BR-364 inter-state highway, especially after it was paved in the mid-1980s. All of the major roads in the region were paved or in the process of being paved by 2005, but the side roads and especially bridges remain unimproved and difficult to travel during the rainy season (Shone and Caviglia-Harris 2006). In the early 1990s, the region was subdivided into four, and later six, municipalities: Ouro Preto do Oeste, Vale do Paraíso, Urupá, Mirante da Serra, Nova União, and Teixeirópolis. The total population was over 83,000 by 2010, with 46% rural and the rest in the central city of Ouro Preto do Oeste and towns in each of the other five municipalities (IBGE 2010). In 2010, the total (urban) population of the six municipalities varied from4,888 (1,716) to 37,928 (28,180) (IBGE 2010).

The cattle herd in Rondônia grew steadily from minimal levels in the 1970s to the second largest among the Amazonian states by 1991 (Faminow 1998). From 1997 to 2010, the growth rate of cattle herd in the Ouro Preto do Oeste region was 99% and in Rondônia was 173%; the growth rate of milk producing cows in the Ouro Preto do Oeste region was 218% and in Rondônia was 215% (IBGE 2016).In our study region, on average, 20.1 % of the land on a property was pasture in 1986, increasing to 77.6% in 2009. Creating this pasture has been the immediate motivation for most deforestation in the region. Farmers in the region (especially the 83% with 100 hectares or less of land) specialize in dairy cattle. Other sources of income for farm households include annual and perennial crops, honey and fish, off-farm labor and government payments such as pensions and school subsidy payments. Income from milk is the both the largest and among the most regular sources of income.

4.2 Data Description

The data[10] consist of (1) land cover on all 8,793 small farm properties (lot size less than 240 hectares [11]) in the Ouro Preto do Oeste region over 24 years (1985-2009), derived from annual Landsat imagery archive; (2) spatial data including farm boundaries, road networks, market locations, and biophysical characteristics (soil, terrain, and hydrology) of the 8,793 properties,

[10] The data set was made available through the project "Living with Deforestation: Analyzing Transformations in Welfare and Land Use on an Old Amazonian Frontier," Jill Caviglia Harris, Erin Sills, and Dar Roberts, NSF Project SES-0752936 (2008-2013).

[11] National policy in Brazil (11.326/2006 Política Nacional da Agricultura Familiar e Empreendimentos Familiares Rurais) defines family farmers as those owning less than 4 fiscal units of land, which is equivalent to 240 HA in Rondônia.

obtained from multiple Brazilian governmental agencies supplemented with spatial data collection using GPS; and (3) four waves of survey data on a stratified sample of farm households living on these properties (Caviglia-Harris, Roberts, and Sills 2014).

Table 1 provides variable definitions and descriptive statistics. Land cover is classified into 7 categories: primary forest, pasture and green pasture, secondary forest, bare soil or urbanized areas, rock or savanna, water, and obscured by cloud (Roberts et al. 2002). Deforestation is defined as loss of primary forest, which is mostly dense tropical forest in this region (RADAMBRASIL 1978). The property boundaries were obtained mainly from INCRA settlement maps, which included 9392 properties. After excluding properties that are (1) very small (less than 1 hectare), (2) very large (larger than 240 hectares), (3) urban, or (4) large forest reserves, the sample retained includes 8793 properties (Figure 1). The landscape covered by 8793 farm properties contains 14 whole or partial INCRA settlements, and 90% of the Ouro Preto do Oeste region's area. The distance to the nearest urban center was calculated as the minimum travel distance along the road network in the Ouro Preto do Oeste region, which was mapped using data collected during 2005 and 2009[12] with a mapping grade global positioning system (GPS) receiver and roads digitized from an INCRA settlement map (Caviglia-Harris and Harris 2008). While road distance is a significant improvement on the Euclidean distances used in much of the literature, it still does not account for variable travel time or wear and tear related to the conditions of roads and bridges. The data on milk price at state level are taken from the Municipal Livestock Research (PPM) of Brazilian Institute of Geography and Statistics (IBGE) and adjusted for inflation[13].

---

[12] Although road conditions may change (especially the condition of side roads and bridges in the rainy season), the road network remained almost the same during the study period (Google earth 1984-2009). The main exceptions are the side roads around properties in settlements laid out in pie shapes (7.5% of the studied properties), which were only established in 1996.

[13] The Brazilian currency changed 4 times between 1986 to 2009. All prices reported in this paper were converted into the current currency: the Brazilian real (BRL). The formula used to calculate real price: $Real\ Price_t =$

5. Results

Table 2 reports average marginal effects (AME) from the estimated models for a two-part model and a Tobit model, both estimated using the "within-between" RE (WBRE) approach. As a comparison, results based on the standard RE and FE estimators [14] are presented as well. In the WBRE approach, the within and between effects are estimated separately. In contrast, a single estimated coefficient in the standard RE accounts for both within and between effects; and the coefficients in FE represent within effects only. The coefficients of within-effects in WBRE are nearly the same as the corresponding coefficients in FE, as are the standard errors. This indicates that WBRE performs at least as well as FE (Bell and Jones 2015). Estimations with the standard RE estimator differ more from the others, which could be due to heterogeneity bias resulting from modeling time-invariant and time-variant factors in one term. For the probit and logit models in the part 1, the marginal effects are calculated for the probability of a positive outcome, $\Pr({D_{it}}^{*} > 0)$. For the tobit model, two types of marginal effects are presented: a) marginal effect on the latent variable, ${D_{it}}^{*}$; b) marginal effect on the observed (censored) dependent variable, D_{it}.

5.1 Drivers of Deforestation

As reported in tables 2, all hypothesized drivers of deforestation are statistically significant after controlling for forest stock in the previous year, year fixed effects and municipality fixed effects. Among land characteristics, the time-variant regressor age of the property has non-zero within-effects and between-effects. Within-effects measure variation within individuals over time, and between-effects capture variation between individuals. The positive signs of property age in both

$(Nominal\ Price_t) \times (\frac{CPI_{2009}}{CPI_t})$, where t represents the current year, 2009 is the base year and CPI represents consumer price index. Source of CPI: Federal Reserve Bank of St. Louis: Consumer Price Index for Brazil, Annual.

[14] Fixed effects were not estimated for probit model and tobit model.

within-effects and between-effects in models (1) - (9) indicate that the older the property is, the higher the probability of deforestation and the extent of deforestation on the property, relative to younger properties in both the spatial and temporal dimensions. The other land characteristics are time-invariant regressors and demonstrate between effects. The marginal effect of average slope of the lot is negative in all models, suggesting that it is costly to clear land for agricultural use or lower agricultural profits are expected on lots with steeper average slope. The positive sign on soil quality (percentage of the lot classified as good soil quality for agricultural production) means deforestation is more likely to occur and that a larger deforested area is more likely on lots with a larger percentage of good soil. The indicator of easy access to surface water on the lot is positively correlated with the probability of deforestation as well as the extent of deforestation in the two-part model, while it is less statistically significant in the model of the probability of deforestation. This result suggests that land is more likely to be used for agricultural production with when there is good access to water. However, the Tobit model shows a negative correlation between water access and deforestation, although at a lower significance level. This may suggest that the law enforcement factor[15] dominates, i.e. households clear less forest when there are rivers or bodies of water on the lots. Lots laid out in a pie shape have lower probabilities of deforestation and less deforested areas. The other regressor is lot size, whose sign is positive in the part 1 model, but negative in the part 2 model and the Tobit model. This suggests that less forest would be cleared on larger properties, even after controlling for forest stock, but that the probability of any deforestation is positively related to property size.

[15] According to the Brazilian Forest Code (law), forest use is restricted in Areas of Permanent Preservation (APPs), which include Riparian Preservation Areas (RPAs) that protect riverside forest buffers (Soares-Filho et al. 2014). APPs were first written in Federal Decree in 1934.

The interaction term of milk price and distance to the nearest urban center is significantly positively related to both probability of deforestation and deforestation level[16]. Holding the effect of distance to the nearest urban center and other covariates constant, the average marginal effect of milk price at the state level on deforestation probability is positive, but the effect on deforestation level is negative. Specifically, from WBRE two-part model, a 10% increase in milk price is associated with an increase of 0.02% in probability of positive deforestation (the signs and marginal effects of logit and probit model are identical); and conditioning on the positive deforestation, results in 0.24 % decrease in deforestation level, about 0.004 hectares decreasing on average. In aggregate, the average marginal effect of distance to urban center is negatively correlated to both deforestation probability (decrease of 0.02% in probability) and deforestation level (0.16 percentage decrease in level, 0.003 hectares decreasing). This result is consistent with von Thünen theory and previous studies. However, it could also mask significant heterogeneity across years, which we investigate next.

5.2 Has the Influence of Market Access on Deforestation Changed Over Time?

By interacting the distance variable with dummies for each time period, we can investigate whether and how the effect of distance has changed across years. Holding milk price and other covariates constant, the average marginal effects of IHS transformed distance by different years are reported in Table 3. The marginal effects of distance are statistically significantly in most years with different signs, suggesting that the effect of distance has changed over time. The signs of the marginal effects (same as signs of coefficients) indicate whether there would be more deforestation

[16] Milk price is an individual-invariant regressor capturing within-effects and distance is a time-invariant regressor capturing between-effects, the WBRE models here did not separately estimate the within-effects and between-effects of their interaction term.

close to urban centers (negative sign) or far away from the urban centers (positive sign). Although different models did not result in exactly the same signs and magnitudes for every year, there is clearly a common trend (Figure 2): deforestation is significantly negatively related to distance in the years before 1999 and significantly positively related to distance after 2004; 1999 and 2001 are the two exceptional years when distance effects are significantly positive before 2004. Therefore, we propose three possible scenarios for the break point(s): (1) a single break point in 2004; (2) a single break point in 1998; (3) two break points in 1998 and 2004. The largest negative effect of deforestation on the probability of deforestation was estimated for 1986: the probability of deforestation decreases 1.2% when distance increase 10%. In part two, the largest negative distance effect on deforestation level was in 1995. The largest positive distance effect on deforestation was in 2007 in all models: deforestation probability increases by 0.42% when distance increases by 10%, and the deforestation level increases 2.4% when distance increase by 10% (based on the conditional WBRE).

Two approaches are utilized to test the possible break point(s) of distance effect (2004, 1999 or 1999 and 2004). In the first approach, we test the equality of these two/three linear combinations at the five percent level (Table 4). Panel A presents the test results for the break year in 2004. Under the alternative hypothesis that “differences between the average effects of 2004-and-pre and after-2004 are not equal to zero”, the p-value is less than 0.0001 (based on Wald test reporting its significance levels using a chi-squared distribution), so we can reject the null and conclude that the difference in means is statistically significantly different from zero. That is, the average effects of distance on deforestation are significantly different from each other in the pre-1999 and post-2004 periods. In addition, the mean of the pre-2005 marginal effects of distance is significantly negative (at 0.0001 significance level), and the mean of the post-2004 marginal effects of distance is significantly positive (at 0.0001 significance level). Prior to 2005, on average, the probability of deforestation would decrease 0.3% when distance increase 10%; conditional on positive deforestation, the deforestation level would decrease 0.67% with a 10% increase in distance. Starting in 2005, on average, with a 10% increase in distance, the probability of deforestation would increase 0.2; conditional on positive deforestation, the deforestation level would increase 2.1%. Panel B reports the test results for 1998 as a break year. While the p-value of χ^2 statistic is less than 0.001 which suggests the average effects of 1998-and-pre and after-1998 are statistically significantly different from each other, the mean of the after-1998 average marginal effects of distance is less significant (at 0.01 significance level in model (1); 0.001 significance level in model (2); 0.1 significance level in model (8)). Similarly, the mean of the 1998-to-2004 average marginal effects of distance reported in panel C is less significant (at 0.1 significance level in model (1) and (2); at 0.05 significance level in model (8)). Overall, these results suggest that 2004 is a more statistically significant break point.

In the second approach, we test the equality between sets of coefficients (both intercept and slope) in two/three regressions (Chow 1960). As reported in Table 5, the p-value of χ^2 statistic (based on Wald test) is less than 0.001 for all three scenarios of the structural break(s), which suggests that we reject the null hypothesis of no structural break(s). However, compared with the average marginal effect pre and after 2004 breakpoint (at 0.0001 significance level), the average marginal effect of distance in year after 1998 is less significant (at 0.01 significance level in model (1); 0.001 significance level in model (2); 0.1 level in model (8)); and the average marginal effect of distance in year between 1998 and 2004 is less significant as well (at 0.1 level in model (1) and (2); at 0.01 level in model (8)). Therefore, the break point of 2004 is relatively more statistically significant based on interrupted models as well.

6. Discussion

The analysis results revealed that the effect of distance to urban center on deforestation has changed over the study period, from significantly negative to significantly positive starting in 2005. According to the theoretical framework, if the effect of distance on farm gate price of agricultural goods dominates, the relationship between deforestation and distance is expected to be negative. This appears to apply prior to 2005. In more recent years, the effects of distance on policy enforcement, on the wage of off-farm work, or/and on factors influencing household utility through distance (e.g. getting health care and education for children) may dominate, explaining why the relationship between deforestation and distance becomes positive. Thus, we investigate (1) the development of milk (dominant agricultural product) markets and transportation cost; (2) environment legislation and protection policies over the study period; and (3) the development of urban centers which may indicate changes of off-farm wages and other changes of factors influencing household utility through distance.

6.1 Milk Markets

In early 1980s, no milk processing facilities were observed and milk production was mainly for self-consumption in the region (Leite and Furley 1985). In 1991, 70% of farmers produced milk and 80% of the milk was sold to milk plants in Ji- Paraná and Ouro Preto do Oeste (Pedlowski and Dale 1992). According to the household survey conducted in 1991 (Pedlowski and Dale 1992), in the Ouro Preto do Oeste region, the milk plants sent collection trucks to pick up milk at the "gate" of the farms daily, and farmers had to pay a premium for transporting the milk to the plants. The transportation cost paid by farmers amounts to 22% of the milk production. However, in the household survey of 2005 (Caviglia-Harris, Roberts, and Sills 2014), no farmers were directly charged for transportation. In fact, the number of milk processing plants had increased from 11 in prior-1996 years to 19 in 2005, and the expansion of milk plants increased competition among the plants so that they waved transportation premium to stimulate farmers to become their suppliers (Saha 2008). Since most milk plants were located in urban centers, the distance to urban centers could be a reasonable proxy for distance to milk markets. Therefore, the fading role of transportation cost indicates that the effect of distance to urban center on farm gate price of milk becomes less dominant along with increased competition over time.

6.2 Policy Enforcement

The requirement to maintain a "Legal Reserve" (LR) of forest on each property recorded in the Brazilian Forest Code (law) is the most significant regulatory constraint for land use decision making in the study region. Initiated in 1934 (Federal Decree No. 23793) and redefined in 1965 (Federal Decree No.7731), the Forest Code has required landowners to maintain 20 to 80 percent of their rural properties under native vegetation in different time periods and different regions, which depends on the vegetation type present, the property's location and size (Soares-Filho et al. 2014; Brancalion et al. 2016). While starting in 1989 environmental laws became more encompassing and scientific, these requirements were proved challenging to enforce and compliance is difficult to monitor (Drummond and Barros-Platiau 2006; Soares-Filho et al. 2014). Deforestation rates in Rondônia as well as the Brazilian Legal Amazon rose rapidly during the period from 1990 to 2004 (Figure 3), and Brazil had the world's highest rate of deforestation. In 2004, the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm) launched, which is an agreed policy turning point (Assunção et al. 2015; Dalla-Nora et al. 2014). Since peaking in 2004 at 27, 423 square kilometers per year, deforestation in the Legal Amazon and Rondônia have dropped by 73 percent by 2009.

The PPCDAm is the first time to request multiple different government institutions to cooperate and take actions to combat deforestation. The Plan includes the National Institute for Space Research (INPE), the Federal Police, the Federal Highway Police, and the Brazilian Army (IPAM, 2009). One significant change under the PPCDAm framework is the stronger monitoring power over deforestation. The Real-Time System for Detection of Deforestation (DETER) was introduced which improved remote sensing-based Amazon monitoring capacity. DETER was able to detect deforested areas larger than 25 hectares in 15-day intervals and issue an alert that an area is endangered (Assunção et al. 2014). That response would enable environmental police and enforcement personnel to go to the deforested spot and catch illegal loggers (Faleiros 2011; Popkin, 2016).

Under the mechanism of monitoring and control of deforestation after 2004, the closer to urban center the higher law enforcement power. Although the remote sensing monitoring treats land properties equally, the reaction time between the monitoring alert and law enforcement action varies according to the land's spatial location. If the law enforcement personnel were supposed to stay in the urban center where their offices are located and face transportation constraints, landowners would expect a higher risk of being caught for illegal clearing or burning if their land is closer to the urban center. The positive relationship between deforestation and distance to urban center in after-2004 years may be due to the dominant effect of the law enforcement.

6.3 Urban Development

Another possibility to explain the positive relationship between distance and deforestation is that the effect of off-farm wage dominates. The closer to urban center, the lower transportation cost of off-farm employment is, which may lead the household allocate family labor to off-farm employment rather than clearing forest. The distinction between wages of urban employment and incomes in rural sector affects the evolvement of deforestation frontier. According to standard urban economics theories, wages in urban centers will be higher due to the greater demand and cheaper inputs (Fujita 1989). Also, a large amount evidence suggests the positive relationship between urban size and wages (Glaeser and Mare 2001; Partridge et al. 2009). Therefore, the evolvement of urban extent could be used to indicate the changes of wages over time. On the other hand, the development of urban centers suggests the growth of urban population, which was believed to be positively correlated with deforestation (Angelsen 1999; DeFries et al. 2010). In addition, classic models of urbanization suggest that as the deforestation frontier closes, the redistribution of population from rural areas to cities would be intensified, alleviating the pressure for agricultural land (Walker 1993; Simmons et al. 2002). All of these motivate the paper to investigate the urban development.

The contemporary urban settlements and urban activities could be “visually” observed by the artificial illumination of buildings, transportation infrastructure, and other components of the built environment (Mellander et al. 2015). One way to measure the economic activities, growth in population, and urbanization is using the amount of night-time lights that can be observed from outer space, which are statellite-based time series data available through DMSP-OLS (Defense Meteorological Satellite Program - Operational Linescan System) night-time lights rasters for years 1992 to 2013 (Amaral et al. 2006; Huang et al. 2014). Utilizing night-time lights as proxy for urban development overcomes the difficulties in collecting socio-economic and demographic data at the subnational level over the long term (Mellander et al. 2015), and provides a reasonable way to compare urban development over regions and periods (Henderson et al. 2012). According to the data of night-time lights in the Ouro Preto do Oeste region from 1992 to 2010, the three cities, Jaru, Ouro Preto do Oeste, and Ji-Paraná, had highest lights intensity and biggest extent of lights over the study period. The comparison of lights among urban centers reflects differences in population growth, wages growth, and stages of urbanization. In 1992, besides the three cities, only Mirante da Serra and Urupá are lit areas. In 1993, Teixeirópolis and Nova União started to be lit up with lower intensity; light intensity in Mirante da Serra is increasing significantly. Vale do Paraíso was in dark until the end of 1993. In the later years till 2010, all municipalities were experiencing increasing light intensity but with different growth rates. Besides the three cities, the northest Vale do Paraíso and Rondominas[17]and southest Mirante da Serra and Urupá grew fast and had higher light intensity. The lowest light intensity was in Teixeirópolis.

Given the regional differences in urban development, we estimate the effect of distance to the nearest urban center on deforestation for each municipality separately, in order to further explain the heterogeneities underlying the general evolving relationship between market access and deforestation. While the distance effects in Ouro Preto do Oeste does not appear clear change patterns, the distance effects in Mirante da Serra and Urupá demonstrate the similar trend that we found for the whole region: deforestation is negatively related to distance in pre-2000 years and positively related to distance after 2000. From the night-time lights maps, Ouro Preto do Oeste is the most mature urban center in the region since 1992, then its surrounding municipalities began to grow as well, especially Mirante da Serra and Urupá. Referring new economic geography theory

[17] Rondominas is a district within the municipality of Ouro Preto do Oeste, which is included in the calculation of the nearest distance to urban centers.

on urban development (Krugman 1999), when centrifugal (militate against concentration) force was dominant in a developed city (i.e. Ouro Preto do Oeste), centripetal (concentration) force may be dominant in its surrounding developing towns (i.e. Mirante da Serra and Urupá). The evolvement of distance effect on deforestation was alone with the growth of urban of Mirante da Serra and Urupá. The significant positive effect of distance in after-2000 years may be due to dominant effect of urban wages and/or shorter population weighted distance in Mirante da Serra and Urupá.

7. Conclusion

Market access represented by distance to urban center has multiple dimensions with different effects on deforestation. The mechanisms by which distance affects deforestation include farm gate price of agricultural output, the wage of off-farm labor, and the shadow cost of illegal deforestation (policy enforcement). The relationship between distance to urban center and deforestation is not constant unchanged, and is determined by the dominant mechanisms in particular periods and regions. With a 24-year panel on about 8,800 farm properties in the Ouro Preto do Oeste region of Rondônia, the empirical results show that after controlling for forest stock and other drivers of deforestation, the effect of distance to urban center on both deforestation probability and deforestation level has changed over time. Specifically, the average effect of distance to the nearest urban center was significantly negatively correlated with deforestation in the 2004-and-pre-years, but positively correlated with deforestation in after 2004 years. Supported by structural break tests, 2004 is the most significant break point.

An investigation in the development of milk (the dominant agricultural output) markets and transportation cost suggests that the effect of distance to urban center on farm gate price of milk became less dominant along with the increasing competition of milk plants over time. In addition, the increasing monitoring and control of deforestation after 2004 advanced the dominance of the policy enforcement mechanism. The diminishing role of transportation cost in farm gate price combined with the increasing influence of environmental enforcement forced the effect of distance on deforestation from negative to positive. Furthermore, there were regional heterogeneities underlying the general evolving relationship between market access and deforestation, which were consistent with regional differences in urban development.

To confirm that forest law enforcement was the key factor in this structural shift in the relationship between urban access and deforestation, our future research plans include distinguishing cities by the types of administrative offices present and by the level of economic activity, based on night-time lights data. We expect this to provide further insight on the evolving relationship between rural and urban areas on this aging tropical forest frontier.

.

REFERENCES

Aguiar, Ana Paula Dutra, Gilberto Câmara, and Maria Isabel Sobral Escada. "Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity." Ecological modelling 209, no. 2 (2007): 169-188.

Allison, Paul D. *Fixed effects regression models*. Vol. 160. London: Sage, 2009.

Amaral, Silvana, Antonio MV Monteiro, Gilberto Câmara, and J. A. Quintanilha. "DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon." *International Journal of Remote Sensing* 27, no. 05 (2006): 855-870.

Angelsen, Arild. "Agricultural expansion and deforestation: modelling the impact of population, market forces and property rights." *Journal of development economics* 58, no. 1 (1999): 185-218.

Angelsen, Arild. "Forest cover change in space and time: combining the von Thünen and forest transition theories." *World Bank policy research working paper* 4117 (2007).

Assunção, Juliano, Clarissa Gandour, and Rudi Rocha. "Deforestation slowdown in the Brazilian Amazon: prices or policies?." Environment and Development Economics 20, no. 06 (2015): 697-722.

Assunção, Juliano, Clarissa Gandour, and Romero Rocha. "DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement." (2014).

Azevedo, Andrea A., Raoni Rajão, Marcelo A. Costa, Marcelo CC Stabile, Marcia N. Macedo, Tiago NP dos Reis, Ane Alencar, Britaldo S. Soares-Filho, and Rayane Pacheco. "Limits of Brazil's Forest Code as a means to end illegal deforestation." *Proceedings of the National Academy of Sciences* 114, no. 29 (2017): 7653-7658.

Bell, Andrew, and Kelvyn Jones. "Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data." *Political Science Research and Methods* 3, no. 01 (2015): 133-153.

Börner, Jan, Krisztina Kis-Katos, Jorge Hargrave, and Konstantin König. "Post-crackdown effectiveness of field-based forest law enforcement in the Brazilian Amazon." *PLoS One* 10, no. 4 (2015): e0121544.

Brancalion, Pedro HS, Letícia C. Garcia, Rafael Loyola, Ricardo R. Rodrigues, Valério D. Pillar, and Thomas M. Lewinsohn. "A critical analysis of the Native Vegetation Protection Law of Brazil (2012): updates and ongoing initiatives." *Natureza & Conservação*14 (2016): 1-15.

Burbidge, John B., Lonnie Magee, and A. Leslie Robb. "Alternative transformations to handle extreme values of the dependent variable." *Journal of the American Statistical Association* 83, no. 401 (1988): 123-127.

Caviglia-Harris, Jill L., and Daniel W. Harris. "Integrating survey and remote sensing data to analyze land use at a fine scale: insights from agricultural households in the Brazilian Amazon." *International regional science review*31, no. 2 (2008): 115-137.

Caviglia-Harris, Jill L., Erin O. Sills, Luke Jones, Shubhayu Saha, Daniel Harris, Suzanne McArdle, Dar Roberts, Marcos Pedlowski, and Rebecca Powell. "Modeling land use and land cover change in an Amazonian frontier settlement: Strategies for addressing population change and panel attrition." *Journal of Land Use Science* 4, no. 4 (2009): 275-307.

Caviglia-Harris, Jill, and Daniel Harris. "The impact of settlement design on tropical deforestation rates and resulting land cover patterns." *Agricultural and Resource Economics Review* 40, no. 3 (2011): 451.

Caviglia-Harris, Jill, Dar Roberts, and Erin Sills. Longitudinal Survey Data of Households in Ouro Preto do Oeste, Rondonia, Brazil, 1996-2009. ICPSR34905-v1. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2014-07-15. http://doi.org/10.3886/ICPSR34905.v1

Chomitz, Kenneth M., and David A. Gray. "Roads, land use, and deforestation: a spatial model applied to Belize." *The World Bank Economic Review* 10, no. 3 (1996): 487-512.

Chow, Gregory C. "Tests of equality between sets of coefficients in two linear regressions." *Econometrica: Journal of the Econometric Society* (1960): 591-605.

Cragg, John G. "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica: Journal of the Econometric Society* (1971): 829-844.

Dalla-Nora, Eloi Lennon, Ana Paula Dutra de Aguiar, David Montenegro Lapola, and Geert Woltjer. "Why have land use change models for the Amazon failed to capture the amount of deforestation over the last decade?." *Land Use Policy* 39 (2014): 403-411.

DeFries, Ruth S., Thomas Rudel, Maria Uriarte, and Matthew Hansen. "Deforestation driven by urban population growth and agricultural trade in the twenty-first century." *Nature Geoscience* 3, no. 3 (2010): 178-181.

de Souza Soler, Luciana, and Peter H. Verburg. "Combining remote sensing and household level data for regional scale analysis of land cover change in the Brazilian Amazon." *Regional Environmental Change* 10, no. 4 (2010): 371-386.

Drummond, José, and ANA FLÁVIA BARROS-PLATIAU. "Brazilian environmental laws and policies, 1934–2002: a critical overview." *Law & Policy* 28, no. 1 (2006): 83-108.

Duan, Naihua, Willard G. Manning, Carl N. Morris, and Joseph P. Newhouse. "Choosing between the sample-selection model and the multi-part model." *Journal of Business & Economic Statistics* 2, no. 3 (1984): 283-289.

Engel, Christoph, and Peter G. Moffatt. "dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression." *Stata Journal* 14, no. 4 (2014): 778-797.

Faminow, Merle D. Cattle, deforestation and development in the Amazon: An economic, agronomic and environmental perspective. No. 333.7409811 F198. New York: Cab International, 1998.

Faleiros, Gustavo. “Looking Down on Deforestation: Brazil Sharpens Its Eyes in the Sky to Snag Illegal Rainforest Loggers.” Scientificamerican.com. https://www.scientificamerican.com/article/brazil-satellites-catch-illegal-rainforest-loggers/ (accessed April 12, 2011)

Fearnside, Philip M. "Deforestation in Brazilian Amazonia: history, rates, and consequences." *Conservation biology* 19, no. 3 (2005): 680-688.

Friedline, Terri, Rainier D. Masa, and Gina AN Chowa. "Transforming wealth: Using the inverse hyperbolic sine (IHS) and splines to predict youth’s math achievement." *Social science research* 49 (2015): 264-287.

Fujita, Masahisa. *Urban economic theory: land use and city size*. Cambridge university press, 1989.

Geoghegan, Jacqueline, Laura Schneider, and Colin Vance. "Temporal dynamics and spatial scales: Modeling deforestation in the southern Yucatán peninsular region." *GeoJournal* 61, no. 4 (2004): 353-363.

Glaeser, E.L. and Mare, D.C., 2001. Cities and skills. *Journal of labor economics*, *19*(2), pp.316-342.

Google Earth. Historical imagery, 1984-2009. Ouro Preto do Oeste - State of Rondônia, Brazil 10°36'03.00" S 62°18'34.50" W. November 13, 2017.

Greene, William. "Fixed effects and bias due to the incidental parameters problem in the Tobit model." *Econometric reviews* 23, no. 2 (2004): 125-147.

Hansen, Matthew C., Peter V. Potapov, Rebecca Moore, Matt Hancher, SAa Turubanova, Alexandra Tyukavina, David Thau et al. "High-resolution global maps of 21st-century forest cover change." *science* 342, no. 6160 (2013): 850-853.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. "Measuring economic growth from outer space." *The American Economic Review* 102, no. 2 (2012): 994-1028.

Hsu, Ai-Chi, and Shu-Chin Liu. "The hurdle models choice between truncated normal and lognormal." *Applied Economics* 40, no. 2 (2008): 201-207.

Huang, Qingxu, Xi Yang, Bin Gao, Yang Yang, and Yuanyuan Zhao. "Application of DMSP/OLS nighttime light images: A meta-analysis and a systematic literature review." *Remote Sensing* 6, no. 8 (2014): 6844-6866.

INPE, "Monitoramento da Floresta AmazôNICA Brasileira Por Satélite Projeto Prodes." INPE online. http://www.obt.inpe.br/prodes.

IBGE, *2010 Annual Report*, IBGE online

IBGE, *2016 Municipal Livestock Research*, IBGE online

IPAM. "Evolução na política para o controle do desmatamento na Amazônia brasileira: o PPCDAM." *Clima e Floresta* (2009).

Jones, Donald W., Virginia H. Dale, John J. Beauchamp, Marcos A. Pedlowski, and Robert V. O'Neill. "Farming in Rondonia." *Resource and Energy Economics*17, no. 2 (1995): 155-188.

Key, Nigel, Elisabeth Sadoulet, and Alain De Janvry. "Transactions costs and agricultural household supply response." *American journal of agricultural economics* 82, no. 2 (2000): 245-259.

Krugman, Paul. "The role of geography in development." *International regional science review* 22, no. 2 (1999): 142-161.

Leite, Laercio L., and Peter A. Furley. "*Land development in the Brazilian Amazon with particular reference to Rondônia and the Ouro Prêto colonisation project"* in J. Hemming (ed.), Change in the Amazon Basin II, Manchester University Press., 1985.

Lewis-Beck, Michael S. "Interrupted time series." *New tools for social Scientists* (1986): 209-240.

Mann, Michael L., Robert K. Kaufmann, Dana Bauer, Sucharita Gopal, Maria Del Carmen Vera-Diaz, Daniel Nepstad, Frank Merry, Jennifer Kallay, and Gregory S. Amacher. "The economics of cropland conversion in Amazonia: The importance of agricultural rent." *Ecological Economics* 69, no. 7 (2010): 1503-1509.

Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson. "Night-time light data: A good proxy measure for economic activity?." *PloS one* 10, no. 10 (2015): e0139779.

Min, Yongyi, and Alan Agresti. "Modeling nonnegative data with clumping at zero: a survey." *Journal of the Iranian Statistical Society* 1, no. 1 (2002): 7-33.

Mundlak, Yair. "On the pooling of time series and cross section data."*Econometrica* 46, no. 1 (1978): 69-85.

North, Douglass C. "Transaction costs through time." *Economics Working Paper Archive at WUSTL* (1994).

Olsen, Maren K., and Joseph L. Schafer. "A two-part random-effects model for semicontinuous longitudinal data." *Journal of the American Statistical Association* 96, no. 454 (2001): 730-745.

Padoch, Christine, Eduardo Brondizio, Sandra Costa, Miguel Pinedo-Vasquez, Robin Sears, and Andrea Siqueira. "Urban forest and rural cities: multi-sited households, consumption patterns, and forest resources in Amazonia." *Ecology and Society* 13, no. 2 (2008).

Partridge, Mark D., Dan S. Rickman, Kamar Ali, and M. Rose Olfert. "Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy." *Journal of International Economics* 78, no. 1 (2009): 126-140.

Pedlowski, M. A., and V. H. Dale. *Land-use practices in Ouro Preto do Oeste, Rondonia, Brazil*. No. ORNL/TM-12062. Oak Ridge National Lab., TN (United States), 1992.

Pence, Karen M. "The role of wealth transformations: An application to estimating the effect of tax incentives on saving." *The BE Journal of Economic Analysis & Policy* 5, no. 1 (2006).

Pendleton, Linwood H., and E. Lance Howe. "Market integration, development, and smallholder forest clearance." *Land Economics* 78, no. 1 (2002): 1-19.

Pfaff, A., G. Amacher, and E. Sills. "Realistic REDD: understanding and improving domestic policy impacts on forest loss." *Review of Environmental Economics and Policy* (2013).

Popkin, Gabriel. "Satellite alerts track deforestation in real time." *Nature* 530, no. 7591 (2016): 392.

RADAMBRASIL, Mapa exploratório de solos, Projeto Radambrasil: Programa de Integração Nacional, 1:1,000,000, Minist. das Minas e Energ., Dep. Nac. de Prod., Rio de Janeiro, 1978.

Roberts, D. A., I. Numata, K. Holmes, G. Batista, T. Krug, A. Monteiro, B. Powell, and O. A. Chadwick. "Large area mapping of land-cover change in Rondônia using multitemporal spectral mixture analysis and decision tree classifiers." *Journal of Geophysical Research: Atmospheres* 107, no. D20 (2002).

Sadoulet, Elisabeth, and Alain De Janvry. *Quantitative development policy analysis*. No. 338.9011 S3. Baltimore: Johns Hopkins University Press, 1995.

Saha, Shubhayu. *Spatial externalities: Evidence on social interaction, market expansion and mineral extraction from Brazil and India*. North Carolina State University, 2008.

Schneider, Robert. *Land Abandonment, property rights, and agricultural sustainability in the Amazon*. World Bank, Latin America Technical Department, Environment Division, 1993.

Shone, Bryan M., and Jill L. Caviglia-Harris. "Quantifying and comparing the value of non-timber forest products in the Amazon." *Ecological Economics*58, no. 2 (2006): 249-267.

Sills, Erin, and Jill L. Caviglia-Harris. "Evolution of the Amazonian frontier: Land values in Rondônia, Brazil." *Land Use Policy* 26, no. 1 (2009): 55-67.

Sills, Erin. "Economics of the Evolution of the Amazon Frontier." Chapter 24 in S. Kant and J. Alavalapati, eds. *Handbook of Forest Economics*. Florence, GB: Routledge, 2014.

Simmons, Cynthia S., Stephen Perz, Marcos A. Pedlowski, and Luiz Guilherme Teixeira Silva. "The changing dynamics of land conflict in the Brazilian Amazon: The rural-urban complex and its environmental implications." *Urban Ecosystems* 6, no. 1 (2002): 99-121.

Singh Inderjit, Lyn Squire, and John Strauss. *Agricultural household models: extensions, applications, and policy*. Johns Hopkins University Press, 1986.

Soares-Filho, Britaldo, Raoni Rajão, Marcia Macedo, Arnaldo Carneiro, William Costa, Michael Coe, Hermann Rodrigues, and Ane Alencar. "Cracking Brazil's forest code." Science 344, no. 6182 (2014): 363-364.

Soares-Filho, Britaldo, Raoni Rajão, Frank Merry, Hermann Rodrigues, Juliana Davis, Letícia Lima, Marcia Macedo, Michael Coe, Arnaldo Carneiro, and Leonardo Santiago. "Brazil's market for trading forest certificates." *PLoS One* 11, no. 4 (2016): e0152311.

StataCorp. *Stata 14 Base Reference Manual*. College Station. TX: Stata Press, 2015.

Von Thünen, Johann Heinrich, and Peter Geoffrey Hall. *Isolated state: an English edition of Der isolierte Staat*. Pergamon, 1966.

Walker, Robert. "Deforestation and economic development." *Canadian Journal of Regional Science* 16, no. 3 (1993): 481-497.

Wooldridge, Jeffrey M. Econometric analysis of cross section and panel data. MIT press, 2010.

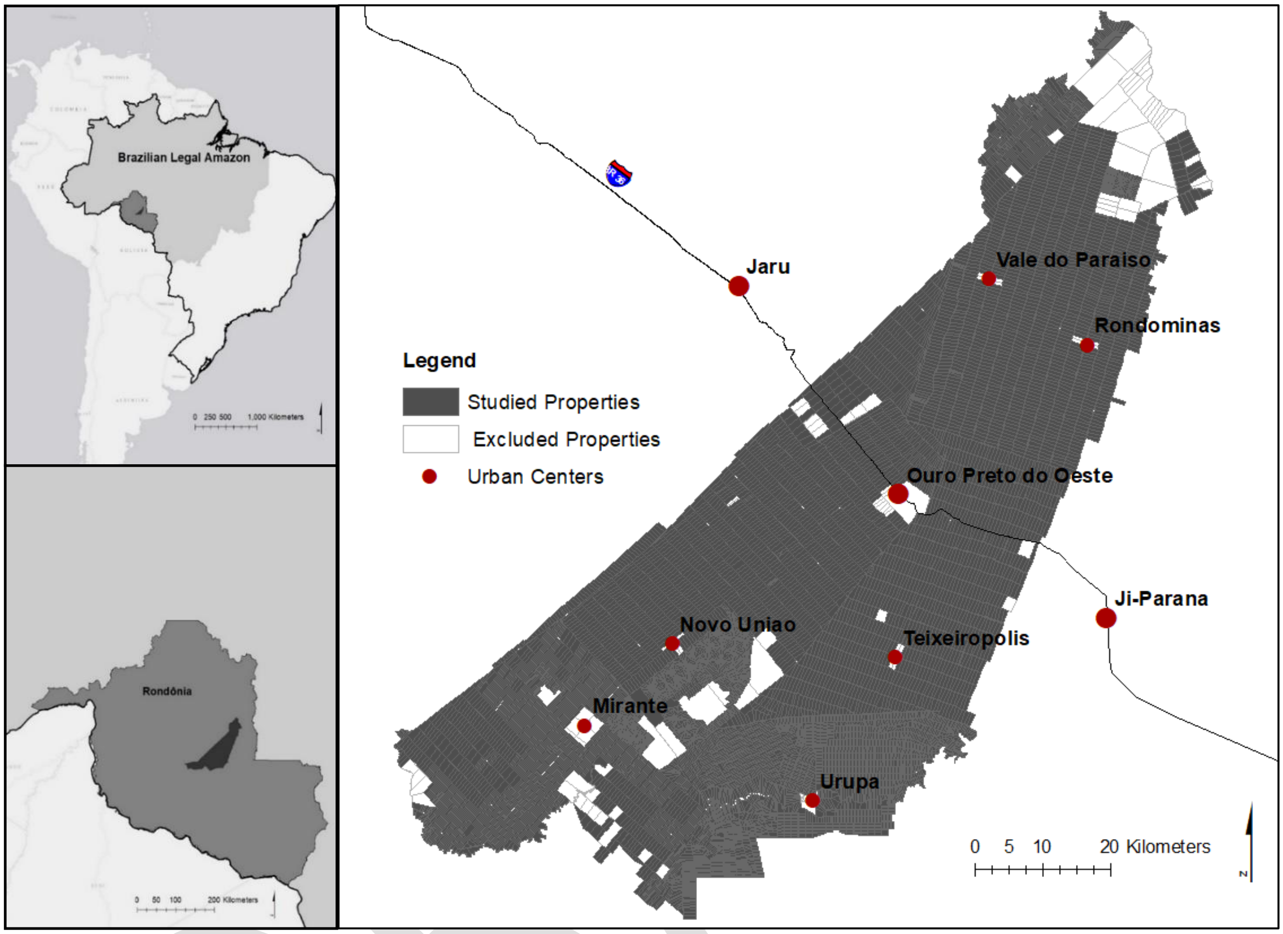

Figure 1. Studied Properties (n=8793) in the Ouro Preto do Oeste Region

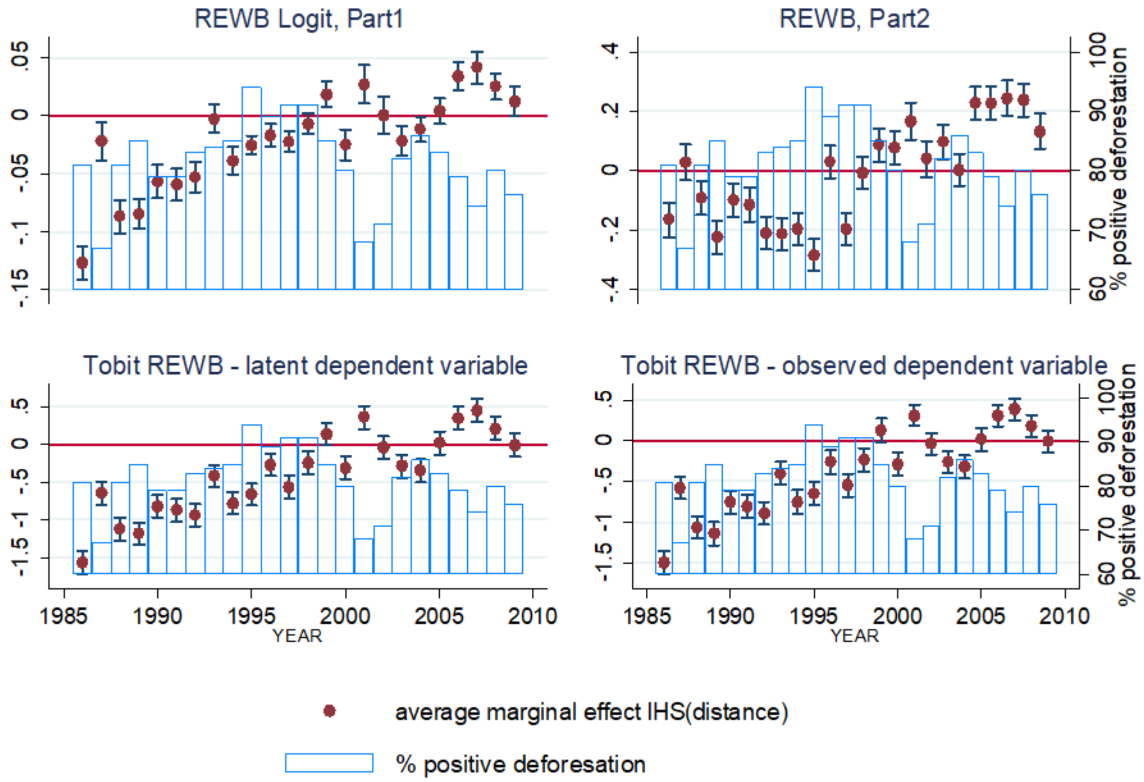

Figure 2. Average Marginal Effect of IHS(distance) by different years

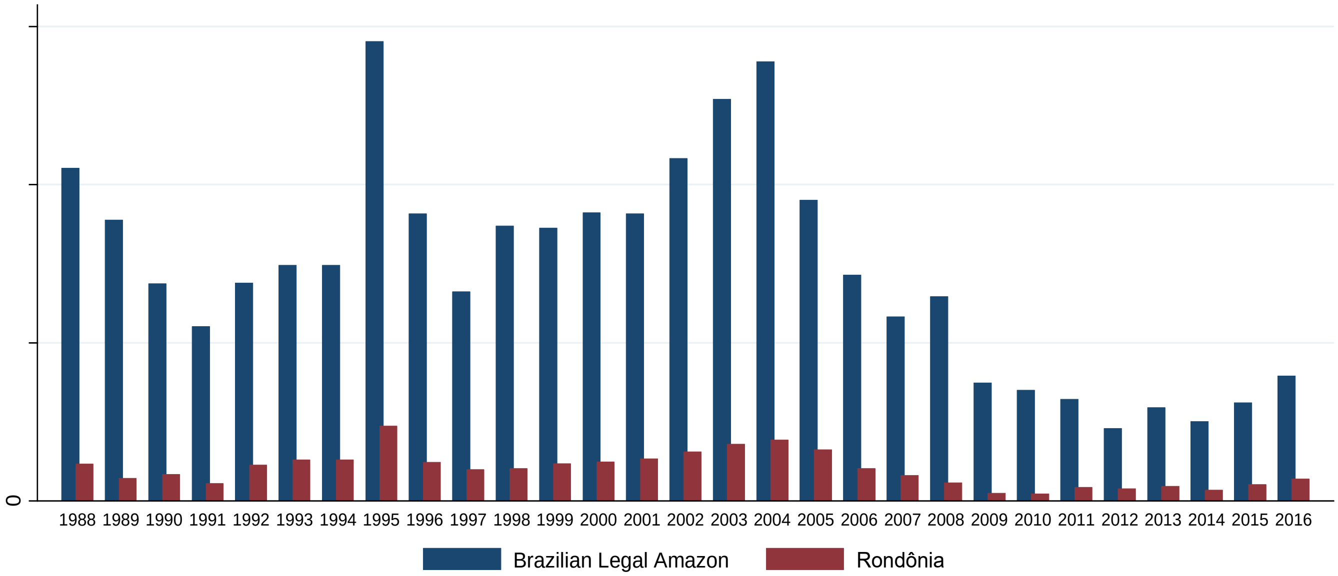

Figure 3. Annual Rate of Deforestation in Rondônia and Brazilian Legal Amazon

Data Source: INPE (2017)

Table 1. Variable Definitions and Descriptive Statistics

Variable	Definition	Mean	Std. Dev.	Min	Max	Observations
Dependent variable						
Deforestation	Deforestation areas: annual loss of primary forest cover on the lot, in hectares.	1.582	3.231	0.000	153.403	N = 211032
Market access						
milk price	Annual milk price at state level, in 2009 Brazilian Reais per liter.	0.327	0.170	0.000	0.641	N = 211032
distance	Road distance to the nearest urban center (Jaru, Ji-Paraná, Mirante da Serra, Nova União,Ouro Preto do Oeste, Rondominas, Teixeirópolis, Urupá, Vale do Paraíso), in meters.	16989.980	8270.798	806.000	50681.000	N = 211032
Characteristics of Land						
property age	Age of the property are years since first primary forest cleared. When the properties have forest cleared before 1985 (the first year of the study period), the ages were based on the official settlement records from INCRA.	16.904	11.572	0.000	39.000	N = 211032
slope	Weighted average of lot slope, calculated using the percentage of the lot in each 5% inclination class, from 0% to 50%.	4.079	2.048	2.500	22.284	N = 211032
soil quality	Percentage of the lot's area classified as good for agriculture.	0.012	0.104	0.000	1.000	N = 211032
water access	Dummy variable, whether the lot has relatively easy access to water sources (main rivers and/or bodies of water).	0.352	0.478	0	1	N = 211032
pie shape	Dummy variable, whether the lot has a pie shape.	0.075	0.264	0	1	N = 211032
Other control variables						
forest cover $_{(t-1)}$	Primary forest cover (hectares) in the last year on the lot.	27.104	25.636	0.000	239.355	N = 211032
lot size	Size of the property measured in hectares.	61.784	37.781	1.288	239.993	N = 211032

Table 2. Average Marginal Effects (AME) for the Estimated Models

Variable	Two-part Model							Tobit Model	
	Part 1				Part 2				
	(1) Probit WBRE	(2) Logit WBRE	(3) Logit RE	(4) Logit FE	(5) WBRE	(6) RE	(7) FE	(8) Tobit WBRE- latent dependent variable	(9) Tobit WBRE- observed dependent variable
Within-effects									
IHS (forest cover $_{(t-1)}$) within-effect	0.108***	0.106***	0.091***	0.001***	0.910***	0.667***	0.945***	1.992***	1.840***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.006)	(0.005)	(0.006)	(0.014)	(0.012)
IHS (property age) within-effect	0.111***	0.110***	0.087***	0.001***	0.450***	0.284***	0.433***	1.951***	1.801***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.007)	(0.005)	(0.007)	(0.016)	(0.015)
IHS (milk price)	0.002*	0.002**	0.005***	0.000**	-0.024***	-0.062***	-0.003	-2.838*	-2.620*
	(0.001)	(0.001)	(0.001)	(0.000)	(0.005)	(0.004)	(0.005)	(1.768)	(1.632)
Between-effects									
IHS (forest cover $_{(t-1)}$) between-effect	0.045***	0.046***			0.218***			-3.159	-2.917
	(0.002)	(0.002)			(0.008)			(3.257)	(3.007)
IHS (property age) between-effect	0.040***	0.038***			0.040***			0.999***	0.922***
	(0.002)	(0.002)			(0.007)			(0.026)	(0.024)
IHS (distance)	-0.021***	-0.022***	-0.014***	0.001***	-0.016*	-0.020**	0.205***	-0.399***	-0.381***
	(0.002)	(0.002)	(0.002)	(0.000)	(0.009)	(0.009)	(0.032)	(0.035)	(0.032)
IHS (slope)	-0.003	-0.003	-0.016***					-0.601***	-0.555***
	(0.003)	(0.004)	(0.004)					(0.053)	(0.048)
IHS (lot size)	0.033***	0.031***	-0.040***		-0.144***	-0.302***		-0.877***	-0.810***
	(0.003)	(0.003)	(0.003)		(0.014)	(0.015)		(0.042)	(0.039)
soil quality	0.054***	0.055***	0.015		0.422***	-0.185***		0.507***	0.468***
	(0.013)	(0.014)	(0.015)		(0.014)	(0.013)		(0.179)	(0.166)
water access (dummy)	0.005*	0.005*	-0.002		0.136***	0.036		-0.061*	-0.057*

Variable	Two-part Model							Tobit Model	
	Part 1				Part 2				
	(1) Probit WBRE	(2) Logit WBRE	(3) Logit RE	(4) Logit FE	(5) WBRE	(6) RE	(7) FE	(8) Tobit WBRE- latent dependent variable	(9) Tobit WBRE- observed dependent variable
	(0.003)	(0.003)	(0.003)		(0.047)	(0.049)		(0.039)	(0.036)
pie shape (dummy)	-0.083***	-0.084***	-0.061***		-0.013	-0.056***		-1.646***	-1.520***
	(0.005)	(0.005)	(0.006)		(0.010)	(0.011)		(0.091)	(0.084)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	206,136	206,136	206,136	167,602	171,774	171,774	171,774	211,032	211,032
Number of groups (properties)	8,792	8,792	8,792	7,150	8,790	8,790	8,790	8,793	8,793

Note: These are average marginal effects (the average of the individual marginal effects) from the estimated models. For WBRE adjusted tobit model (8), there are two types of marginal effects of interest in the paper: a) marginal effect on the latent variable, ${D_{it}}^*$; b) marginal effect on the observed (censored) dependent variable, D_{it}.

The dependent variable in models of part 1 is 1, if the (IHS transformed) annual deforested areas larger than zero; The dependent variable in models of part 2 is IHS transformed annual deforested areas; The dependent variable in the tobit models is IHS transformed annual deforested areas, if the (IHS transformed) annual deforested areas larger than zero, and is zero if otherwise.

Standard errors in parenthesis; "***", "**", "*", indicate significance at the 1%, 5%, and 10% levels.

Table 3. Average Marginal Effect of IHS(distance) by Different Years

Year	Two-part Model			Tobit Model			
	Part 1		Part 2				
	(1) Probit WBRE	(2) Logit WBRE	(5) WBRE	(8) Tobit WBRE- latent dependent variable	(8) Tobit WBRE- observed dependent variable	$p(F01_{t-1} > 0)$	p(DF > 0)
1986	-0.122***	-0.127***	-0.165***	-1.565***	-1.491***	99.99%	80.60%
1987	-0.022***	-0.022***	0.028	-0.647***	-0.576***	99.97%	66.66%
1988	-0.082***	-0.087***	-0.092***	-1.120***	-1.062***	99.95%	81.49%
1989	-0.081***	-0.085***	-0.224***	-1.182***	-1.137***	99.92%	85.00%
1990	-0.055***	-0.057***	-0.098***	-0.821***	-0.753***	99.83%	79.01%
1991	-0.058***	-0.059***	-0.115***	-0.872***	-0.799***	99.83%	79.01%
1992	-0.050***	-0.053***	-0.210***	-0.937***	-0.892***	99.81%	82.77%
1993	-0.003	-0.003	-0.214***	-0.418***	-0.402***	99.81%	84.19%
1994	-0.039***	-0.039***	-0.196***	-0.780***	-0.749***	99.77%	85.45%
1995	-0.023***	-0.026***	-0.285***	-0.663***	-0.648***	99.64%	93.51%
1996	-0.016***	-0.017***	0.030	-0.275***	-0.264***	99.44%	89.48%
1997	-0.021***	-0.022***	-0.197***	-0.565***	-0.545***	99.37%	91.47%
1998	-0.008*	-0.007*	-0.008	-0.245***	-0.236***	99.07%	91.38%
1999	0.016***	0.018***	0.086***	0.135*	0.127*	98.70%	84.61%
2000	-0.025***	-0.025***	0.078***	-0.308***	-0.286***	98.37%	80.28%
2001	0.027***	0.027***	0.165***	0.357***	0.311***	97.79%	68.13%
2002	-0.0003	0.0003	0.039	-0.044	-0.039	97.46%	70.78%
2003	-0.021***	-0.022***	0.097***	-0.286***	-0.263***	97.10%	81.66%
2004	-0.012*	-0.012**	0.002	-0.345***	-0.320***	96.57%	86.30%
2005	0.004	0.004	0.227***	0.016	0.014	94.75%	83.13%
2006	0.033***	0.034***	0.227***	0.344***	0.304***	93.42%	79.19%
2007	0.042***	0.041***	0.243***	0.452***	0.387***	92.47%	73.69%
2008	0.023***	0.025***	0.237***	0.206***	0.180***	91.47%	79.97%
2009	0.009*	0.013**	0.132***	-0.012	-0.010	89.82%	75.78%

Note: $p(F01_{t-1} > 0)$ denotes percentage of properties having positive primary forest cover in the previous year; p(DF > 0) denotes percentage of properties having positive deforestation.
"***", "**", "*", indicate significance at the 1%, 5%, and 10% levels.

Table 4. Testing for Difference Between the Average Effect of Distance

	(1) Probit WBRE	(2) Logit WBRE	(5) WBRE	(8) Tobit WBRE- latent dependent variable	(8) Tobit WBRE- observed dependent variable
A. Break year: 2004					
χ^2 statistic	298.830	323.670	390.560	449.830	490.900
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years \leq 2004$	-0.031	-0.032	-0.067	-0.557	-0.528
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years > 2004$	0.022	0.023	0.213	0.201	0.175
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
B. Break year: 1998					
χ^2 statistic	401.680	432.210	551.830	806.850	839.070
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years \leq 1998$	-0.045	-0.046	-0.134	-0.776	-0.735
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years > 1998$	0.009	0.009	0.139	0.047	0.037
	(0.002)	(0.001)	(0.000)	(0.220)	(0.280)
C. Break year: 1998 and 2004					
χ^2 statistic	462.520	496.380	618.050	845.240	889.380
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years \leq 1998$	-0.045	-0.046	-0.134	-0.776	-0.735
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of AME between $years > 1998\ and \leq 2004$	-0.003	-0.002	0.078	-0.082	-0.078
	(0.460)	(0.535)	(0.000)	(0.055)	(0.045)
Mean of AME between $years > 2004$	0.022	0.023	0.213	0.201	0.175
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The null hypothesis to be tested is: $\text{Mean}(\text{AME}(\text{IHS}(\text{distance}))_{years \leq breakpoint}) = \text{Mean}(\text{AME}(\text{IHS}(\text{distance}))_{year > breakpoint})$. Wald test is performed, which reports its significance levels using a chi-squared distribution. p-values are in parenthesis. AME denotes average marginal effect.

Table 5. Testing for Known Structural Break Point(s) Using an Interrupted Model

	(1) Probit WBRE	(2) Logit WBRE	(5) WBRE	(8) Tobit WBRE- latent dependent variable	(8) Tobit WBRE- observed dependent variable
A. Break year: 2004					
χ^2 statistic	828.050	719.020	828.050	1162.880	
	(0.000)	(0.000)	(0.000)	(0.000)	
AME of IHS(distance), if $year \leq 2004$	-0.028	-0.030	-0.068	-0.556	-0.520
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AME of IHS(distance), if $year > 2004$	0.025	0.025	0.196	0.179	0.156
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
B. Break year: 1998					
χ^2 statistic	1663.470	1548.420	642.130	3968.180	
	(0.000)	(0.000)	(0.000)	(0.000)	
AME of IHS(distance), if $year \leq 1998$	-0.041	-0.043	-0.120	-0.756	-0.715
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AME of IHS(distance), if $year > 1998$	0.007	0.008	0.123	0.031	0.028
	(0.017)	(0.008)	(0.000)	(0.412)	(0.416)
C. Break year: 1998 and 2004					
χ^2 statistic	2689.300	2465.750	1365.650	4011.600	
	(0.000)	(0.000)	(0.000)	0.000	
AME of IHS(distance), if $year \leq 1998$	-0.041	-0.043	-0.120	-0.758	-0.717
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AME of IHS(distance), if $1998 < year \leq 2004$	-0.003	-0.002	0.064	-0.095	-0.086

	(1) Probit WBRE	(2) Logit WBRE	(5) WBRE	(8) Tobit WBRE- latent dependent variable	(8) Tobit WBRE- observed dependent variable
	(0.385)	(0.493)	(0.000)	(0.028)	(0.028)
AME of IHS(distance), if $year > 2004$	0.021	0.022	0.190	0.188	0.164
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Using the interrupted models (3 scenarios of breakpoint(s)), Wald test is performed to help determine if there are abrupt changes in both the intercept and the slope of the regression line. The Wald test is computed as chi-squared statistics. p-values are in parenthesis. AME denotes average marginal effect.