THESIS

DETERMINANTS OF DEFORESTATION IN VIETNAM, 2008 – 2015

Submitted by

Thai Binh Tran

Department of Sociology

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Summer 2022

Master's Committee:

Advisor: Anthony Roberts

Kuoray Mao
Anita Pena

ABSTRACT

DETERMINANTS OF DEFORESTATION IN VIETNAM, 2008 – 2015

New methods including satellite data, geographic information systems (GIS), and remote sensing processing have discovered human expansion over forest areas referred as forest degradatio. This study acknowledges these findings but insists on using official data to address some drawbacks of previous studies. These drawbacks include (1) the focusing on limited areas, Central Highlands areas, instead of a national scale, (2) exclusion of resources trade from the analysis, (3) lacking consideration of the spatial and longitudinal autocorrelation, which is overlooked in panel analysis; and (4) the inconsistency of the relation between poverty and deforestation. This research investigated the effects of land-use change from agricultural expansion and timber extraction, resources trade, and community poverty on province-level forest coverage in Vietnam from 2008 to 2015 using panel and spatial autoregressive modelling. After accounting for resources trade, effect of agricultural expansion as well as forest extraction disappear. In addition, panel analysis suggests no covariate along poverty rate affects forest coverage while the spatial analysis suggests literacy rate and agricultural land are also have significant effects.

TABLE OF CONTENTS

ABSTRACT ii
Chapter 1 – Introduction 1
Chapter 2 – Literature Review 6
2.1. Deforestation and Forest Degradation in Vietnam 6
2.2. Theoretical Framework 8
2.2.1. Agricultural Production & Deforestation 8
2.2.2. Timber Extraction & Deforestation 9
2.2.3. Resources Trade & Deforestation 10
2.2.4. Socioeconomic & Demographic Drivers of Deforestation 11
2.2.5. A Theoretical Model of Deforestation 13
Chapter 3 – Data & Methodology 16
3.1. Data Sources 16
3.2. Province Sample 17
3.3. Variables 17
3.3.1. Dependent Variables 17
3.3.2. Focal Variables 18
3.3.3. Control Variables 19
3.4. Data & Model Diagnostics 20
3.4.1. Multivariate Diagnostics 20
3.4.2. Panel Data Diagnostics 21
3.4.3. Spatial Data Diagnostics 23
3.5. Regression Estimators 23
3.6. Model Assessment 26
Chapter 4 – Results 28
4.1. Introduction 28
4.2. Theoretical Framework & Hypothesis 29
4.3. Analytical Strategy 31
4.4. Correlation Analysis 33
4.5. Multivariate Regressions 33
4.6. Spatial Regression 34
4.7. Discussion & Conclusion 35
Chapter 5 – Discussions 38
Bibliography 43

Chapter 1: Introduction

Forest coverage in Vietnam has experienced a period of decline before 1990s (Meyfoidt and Lambin 2008; Meyfroidt et al. 2013). Starting in early the 1990s, provinces across the country implemented series new policies to protect forest area because of new environmental requirements as well as better understanding of impact of deforestation (Meyfroidt and Lambin, 2008). Therefore, the cover forest area stopped decreasing and enter a national reforestation period in early 1990s (Meyfroidt and Lambin, 2008). This accomplishment is due to the realization of consequences of deforestation such as reducing production capacity (Vu et al. 2017), weakening ground foundation leading to landside (Kumar and Bhagavanulu 2008), assessing to clean water (Mupalanga and Naito 2019) and marginalizing populations (Meyfroit et al. 2013).

Though these positive changes lead forest in Vietnam into a reforestation period (Meyfroidt and Lambin, 2013), studies show that such activities do not necessarily decline but rather getting more sophisticated in form of forest degradation (JICA, 2012; Meyfroidt et al. 2013; Quy et al., 2018). Deforestation is defined as (Hosonuma et al. 2012) as removal of trees and conversion of land for different purposes, one of which is significantly economic. Conversely, forest degradation is referred as 'a reduction in the capacity of a forest to produce ecosystem services' caused by 'unsustainable logging, agricultural, invasive species, fire, fuelwood, gathering and livestock grazing' (Thompson et al. 2013), which is harder to either observe or identify. These findings are enabled with the use of original techniques such as remote sensing processing and geographic information systems (Meyfroidt et al. 2013) rather than the use of official data.

This study acknowledges these findings but insists on using official data to address some drawbacks of previous studies. These drawbacks include (1) the focusing on limited areas, Central Highlands areas (Meyfroidt et al. 2013), instead of a national scale, (2) exclusion of resources trade from the analysis, (3) lacking consideration of the spatial and longitudinal autocorrelation, which is overlooked in panel analysis (Quy et al. 2018); and the inconsistency of the relation between poverty and deforestation (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003; Miyamoto 2020; Quy et al 2018). Addressing these problems contributes to the relation between deforestation with agricultural production as well as demographic factors which is significant for drafting development policy following by implementation on groups of populations, especially the poor and marginal communities.

The study addresses the following problems through examining the state of deforestation/forest degradation during the period from 2008 to 2015, creating a theoretical framework for explaining province-level variation in forest coverage, accounting for unobserved heterogeneity with panel data, examining the effects of resources trade on forest coverage and key determinants of forest coverage and testing for spatial dependence between provinces.

Previous studies evaluated deforestation in Vietnam with different economic measurements such as annual crops and marginalization of ethnic communities (Meyfroit et al. 2013), agricultural production with demographic control (Khuc et al. 2018), one of which is the effect of poverty. Additionally, empirical studies have produced inconsistent findings on the association between poverty and deforestation in developing countries Miyamoto 2020; Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003). Several studies point out the deforestation state is only temporary and would decrease given that the poverty rate decrease (Miyamoto 2020; Quy et al 2018). However, it is also inconsistent with which argue that deforestation only go down if there are a comprehensive improvement in different aspects such as living standards, literacy, open to business, etc (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003). The inconclusive comes from whether a reduction in poverty associates with a reduction in deforestation or it requires more factors to improve forest coverage. These factors include living standards, education, accessibility and open to business. The study addresses this problem by examining the effects of a set of demographic variables, including poverty, population, literacy rates and provincial government quality. Accordingly, this study expands the explanatory power of previous models of deforestation by including these covariates.

Previous study has discussed the role of external market but have not empirically studied how agricultural and forest exports affect deforestation in Vietnam. According to the sociological literature on resources trade, the expansion of the international market is a form of externalizing environmental costs to less developed countries from those developed (Jorgenson 2006; Hornborg 2001; Rice 2007). This leads to environmental degradation including, deforestation and forest degradation. Therefore, leaving out the external market factor would potentially lead to omitted variables bias in previous studies because resources exports is likely to correlated with deforestation/forest degradation and covariates such as agricultural production and timber extraction. This study addresses this problem by including two measures of resources trade: agricultural and forestry exports.

Considerations of the geographical proximity of provinces have not been sufficiently addressed in previous studies. Given the propensity for spatial dependence among processes in provinces, trading or economics activities have tendency to cluster in certain areas. For example, the forest coverage is different between provinces which perhaps creates distinct level of deforestation whether a province is close to an economic center if such relation exists. Without considering this spatial dependency, the statistical estimator might be improper leading to misleading hypotheses testing. As a result, this study conducts diagnosis following by a spatial regression for deforestation in Vietnam.

This thesis examines the association between deforestation with different agricultural crops using the data from General Statistics Office of Vietnam and Provincial Competitive Index. After balancing the data through listwise deletion, the data contains 432 observations ranging from 2008 to 2015 and consists of 54 provinces. The study begins with an examination and comparison of regression estimators for modeling deforestation. Based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hausman test, one-way fixed effect is the preferred model even though there is a trade-off between efficiency and consistency. The Moran's I test examines the spatial dependence in the data. The result does not suggest spatial autocorrelation in the dataset, but spatial regression is still included as a robustness check.

Results suggests that there is a negative correlation between poverty level and deforestation in which a one-point increase in poverty rate reduces forest coverage by 0.037 points. This is consistent with past studies that improving in poverty alone associates with increasing in forest coverage (Miyamoto 2020; Quy et al. 2018). Interestingly, even though the spatial diagnosis implied no spatial correlation in the data, given the results, the spatial analysis suggests that literacy rate and agricultural land have significant association with deforestation. This partially indicates that an increasing in forest coverage needs other factors rather than lower poverty rate (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003), though other demographic factors are not significant.

The thesis is organized into 4 additional chapters. In Chapter 2, I explain the state of deforestation and forest degradation in the case of Vietnam. Also, the chapter provides context within socioeconomic factors. The study builds a theoretical framework in which deforestation/forest degradation is affected by its relationship with agricultural expansion, timber extraction and external market. The framework extents to other demographic including population, literacy rate, labor force with training,

In Chapter 3, I review the process of collecting data, defines variables used to operationalize deforestation; agricultural production; economic & living standard controls; and export measurements; and diagnosis evaluations. Also, the chapter discusses different estimators for panel data, following by its assumptions as well as limitations. Statistical models include ordinary least square, random effect, fixed effect one way & two ways, Praise-Winston regression, and spatial regression.

In Chapter 4, I discuss results from a series of panel regression models as well as spatial models to discuss the effects of different dependent variables on forest coverage. In general, poverty rate is negatively correlated with forest coverage representing deforestation/forest degradation. The spatial model provides a similar result. Literacy rate has a positive correlation with forest coverage in both fixed effect model and spatial regression. Interestingly, agriculture land has a negative association with forest coverage only in spatial analysis model.

In Chapter 5, I discuss the implications of the results for the empirical literature on deforestation as well as the broader impacts of this research on potential policy solutions for addressing deforestation in Vietnam.

Chapter 2: Deforestation in Vietnam

This chapter discusses the theoretical framework for the study. It starts with defining deforestation/forest degradation following by identifying the trend of forest coverage in Vietnam. Next, the chapter discusses the role of agricultural production affecting deforestation through low techniques of cultivating and increasing demand for more goods. Timber extraction from forest area is due to the needs for multiple products. Resources trade captures the international demand for agricultural products because it is a mean for more developed countries to externalize their environmental costs. This leads to forms of environmental degradation, one of which is deforestation/forest degradation. Lastly, the chapter discusses different socioeconomic drivers and how each of them relates to deforestation, including population, poverty, and literacy rate. The framework is an explanatory model of deforestation.

2.1. Deforestation and Forest Degradation in Vietnam

Overall, Vietnam experiences high level of deforestation before 1990s (Meyfoidt and Lambin 2008; Meyfroidt et al. 2013). Even though the cover forest area stopped decreasing and enter a national reforestation period (Meyfroidt and Lambin, 2008), several studies have found that deforestation continues to exist but in more sophisticated ways, which are classified as forest degradation (JICA 2012).

Hosonuma et al. (2012) defines deforestation as ‘the (complete) removal of trees and the conversion from forest into other land uses such as agricultural, mining etc’. Forest degradation is referred as ‘a reduction in the capacity of a forest to produce ecosystem services’ caused by ‘unsustainable logging, agricultural, invasive species, fire, fuelwood, gathering and livestock grazing’ (Thompson et al. 2013). Degraded forest is harder to identify ‘because perceptions of forest degradation vary depending on the cause.' (Thompson et al. 2013). Types of forest degradation can be represented using five criteria 'including: productivity, biodiversity, unusual disturbances, protective functions, and carbon storage' (Thompson et al. 2013).

Different measures have been developed recently to re-evaluate deforestation and forest degradation such as geographic information system (GIS), which uses a database containing geographic data combining with software techniques to analyze as well as visualize the data (Quy et al., 2018); and remote sensing processing, which tracks the physical characteristics of an area by measuring its reflected and emitted radiation at a distance, and spatial patterns (Meyfroidt et al., 2013). GIS suggests a 1.77 and 0.65 million ha of lost and degraded forest, respectively, between 2000 and 2010 (Quy et al., 2008). Remote sensing shows negative net deforestation in Dak Lak and Dak Nong (two out of five provinces in Central Highlands, Vietnam) from 2000 – 2010 (Meyfroidt et al., 2013). Meyfroidt 2013 focuses on two provinces Dak Lak and Nak Nong in Central Highlands which is limited in scale.

Though this study acknowledges the use of original methods, it insists on the use of official data to evaluate deforestation because the paper focuses on addressing drawbacks of previous research. The descriptive analysis will at least answer the question of whether the forest coverage is still officially in the reforestation period.

[Supplemental Files: Figure 1]

Figure 1 shows the change of forest coverage in every province from 2008 to 2015. Most provinces have the area of forest staying relatively the same. However, 8 provinces experienced a change with absolute value larger than 45 thousand hectares in any year from 2008 – 2015. Of those eight provinces, Nghe An, Quang Nam, Lang Son, Lam Dong, Dak Lak are in top 10 provinces in terms of forest coverage in 2015 while Bac Kan and Dien Bien rank relatively high. The ranking is mostly identical with the data in 2008. In general, the change of forest coverage remains the same if not increases significantly in some provinces. The only exception is Dien Bien province

[Supplemental Files: Figure 2]

Figure 2 shows the area of forest coverage in each province in 2015 in thousand hectares. Nghe An is the province with the highest area of forest coverage with roughly 1,000,000 hectares. The lowest province in terms of forest coverage is Bac Ninh with nearly 0 thousand hectares. On average, forest coverage across provinces is around 219.88 thousand hectares in 2015. Compared to forest coverage in province in 2008, there are a slight increase in forest coverage in general.

The descriptive analysis suggests that overall, the area of forest coverage is still increasing from 2008 to 2015. This is consistent with previous studies on deforestation (Meyfroidt et al. 2013; Meyfroidt and Lambin 2008). Therefore, the result might imply that official data does not represent the true reality of forest conditions while validate the use of other methods rather than official data.

2.2. Theoretical Framework

This section reviews all elements of the deforestation/forest degradation framework. The framework explains the role of each factor and how they influence forest coverage. These factors include agricultural production, timber extraction, resources trade, and multiple socioeconomic drivers. Each sections provides an understanding of the association between the factor and deforestation/forest degradation in the case of Vietnam.

2.2.1. Agricultural Production & Deforestation

Vietnam economy was a socialist economy before the 1980s with collective farms, land and factories. In agricultural sector, the output of collective farms remains low threatening the food security of the population. With the new economic policy- the 'Doi Moi' from 1986, there was a transition from collective to de-centralized/ liberalized agriculture, specifically in land and capital (Kerkvliet and Porter 1995; Meyfroidt and Lambin 2008; Meyfroidt et al. 2013).

The creation of a more private and competitive economy creates a need for resources, raw materials and land for growing products which previously own only by the state. In addition, a large population were not guaranteed occupations as previously by the state leading to a need of tremendous number of new jobs in this new economy. Government addressed this problem by initiating different economic policies. The 'facilitating regions' and 'New Economic Zones'-NE Zones creates many waves of immigrants to forest areas, especially the Northeast and Central Highlands (Pfaff and Walker 2010; Hardy 2000, 2003). The shift increases both perennial outputs especially in coffee (Ha and Shively 2008), rubber (Ziegler et al 2009, Li and Fox 2012) and annual outputs securing food production.

This achievement comes at a cost for forest (Meyfroidt and Lambin 2008), local communities (De Koninck 2000, Baulch et al 2010). This study investigates the current forest conditions from 2008 – 2015 focusing on the use of annual crops.

2.2.2 Timber Extraction & Deforestation

Timber extraction aims at certain types of trees due to higher market values (Meyfroidt and Lambin 2009) or for general purposes such as fuelwood, charcoal production, uncontrolled fire and livestock grazing leading to forest degradation (Hosonuma et al 2012). In addition, timber logging provides valuable materials for products such as: furniture, medicine, beverage, etc. Exports of wood products from Vietnam have grown rapidly (Meyfroidt and Lambin 2008, 2009).

Also, the demand keeps increasing while the limitation on exploitation is fixed due to the needs to protect forest. As a result, illegal market for timber products exists in multiple provinces. Forest rangers are often underfunded to cover large area which fail to prevent illegal logging. These damages are often either underreport or difficult to evaluate upon discovered. Illegal timber logging becomes a perpetual problem without any reasonable solution (Meyfroidt and Lambin 2008, 2009; Meyfroidt et al 2013).

Timber extraction has perpetuated for multiple reasons (Meyfroidt and Lambin 2009; Hosonuma et al 2012). It might help explaining deforestation has reduced, but forest degradation increased at the same time (Meyfroidt and Lambin 2009). However, it is necessary to emphasize that timber extraction in this framework does not capture the total size of this factor which is the illegal part of it. This study accounts for timber logging in terms of volume of timber extraction, which is only the legal part of timber logging.

2.2.3 Resource Trade & Deforestation

Following the ecologically unequal exchange perspective (Jorgenson 2006; Hornborg 2001; Rice 2007), the stratified global economy allows developed countries to externalize their environmental cost which accelerate environmental degradation in developing countries. The regulation on environmental damage is generally stricter in developed countries compared to its developing counterpart. Therefore, it is not economical to produce certain agricultural products at home. At the same time, developing countries rely heavily on exporting and low wage for worker as advantage in the market. However, this exchange does not include the cost of intensive productions on environment aspects such as: land, air, or water. The environmental impact is often difficult to assess while taking a long time to present themselves. Therefore, by involving in the trade, developed countries can externalize the environmental cost to less-developing countries in form of payments for products while avoiding these potential problems.

Vietnam is a developing country which relies significantly on exporting agricultural products. Such expansion of the international market has been the main driving force for increasing production of agricultural products in Vietnam last decades (De Koninck 1999; Meyfroit 2013). In term of deforestation, economic developing and trade openness positively associated with deforestation (Jorgenson 2010) and environmental degradation at least in non-OECD countries (Jorgenson et al. 2010,). Similar to the argument with environmental degradation above, less develop countries does not have many options for economic development or the environmental cost does not outweigh better living standard for its populations. Therefore, economic expansion into forest area is a feasible activity because it provides multiple raw materials for trading, lands for either living or production purposes. However, contradiction to previous study, agricultural products export does not necessarily contribute to economic growth with inconsistent results in different products (Faridi 2012; Nguyen 2016).

The need for timber extraction and agricultural production are crucial for development in Vietnam because they provide raw materials as well as food security for the population. However, at the same time, partial the demand comes internationally which makes it difficult to evaluate each market separately. It might be safe to assume that such needs from external market amplify the expansion of both forest extraction and agricultural production.

2.2.4 Socioeconomic & Demographic Drivers of Deforestation

Agricultural production is tied with deforestation, particularly in poor and ethnic minority communities (Meyfroidt et al. 2013; Meyfroidt and Lambin 2008, 2011; DeFries et al. 2010). Specifically, marginalized community needs lands for agricultural subsistence and therefore clear forested areas for cultivation. However, without the use of fertilizer or outdated agricultural techniques, the soil become exhausted after a short period of time. It means that they can no longer exploit the land anymore. Therefore, the community has to move to a new area as well as cut down more tree for next production cycle.

However, whether reducing poverty led to decrease of deforestation is a controversial topic. On one hand, reducing poverty has been found to reduce deforestation in the long term (Miyamoto 2020; Quy et al 2018). Conversely, small economic improvement does not necessarily reduce deforestation (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003). This study addresses the inconsistency by including other demographic control such as literacy rate and labor force with training as association with poverty rate.

Poverty is an important variable, but its effects on deforestation are inconsistent across studies (Miyamoto 2020; Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003). On one hand, reducing in poverty is associated with deforestation regardless other factors (Miyamoto 2020). Conversely, literacy rate, access to economics center or open to business and living standard are equally important to reduce deforestation (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003). In Vietnam, impoverished population often locates in province with higher forest coverage. Also, Meyfroit et al. (2013) also suggested that deforestation in Central Highlands is partially driven by poor community forced away from their original lands. They often practice their economic activities, however with expansion of new business, they can either lack the skill to work in the industry or ability to communicate effectively. Regardless, these population have to move to margin forest area and start making a living again. Therefore, one viable option is extracting resources from forest.

Population growth has been discussed having positive association with deforestation (Jorgenson 2010; DeFries et al 2010). An increase in population provides additional workforces and works finely in an agricultural economic because it requires both resources as well as labor intensive for production. The population of Vietnam has been increased for some decades, given better standards of livings. Also, the bigger the population, the higher needs for land use for multiple purposes such as: housing, production, consumption, etc.

Literacy rate is another control variable along with poverty. Due to insufficient evidence of reducing poverty leading to lower deforestation, other factors such as education, business openness, etc matter (Zwane 2007). Literacy, life expectancy and standard of living increases deforestation at first, but declines afterward (Rodrigues et al. 2017, Murtazashvili et. al 2019). This is understandable because improving quality of livings includes many aspects while income is only one of it. When income is low, population might rely on extraction for economic gaining. However, as all aspects of their life improves, extraction, specifically deforestation, becomes economic inferior to other options. Such differences can arrive from better incomes, safe work conditions, or realization of potential damage by deforestation. Therefore, it takes time and large population for such conditions to gradually improve.

2.2.5 A Theoretical Model of Deforestation

Based on the above discussion, I developed a theoretical framework which identifies forest coverage as a function of cereal crop, annual crops, timber extraction, agricultural land, provincial competitive index, literacy rate, labor force participation, net migration, poverty rate, agricultural and forestry exports. The framework is summarized with the following function:

$$(1)\ fc_{it} = f_{ac_{i,t}, tb_{i,t}, ev_t, po_{i,t}, lr_{i,t}, ci_{i,t}, lf_{i,t}, nm_{i,t}, al_{i,t}}, \text{ in which}$$

fc stands for forest coverage in province i and year t, ac is for annual crops including crops for sugarcane, peanut and soybean, and cereal crops, tb is for timber extraction, ev is for export variables including agricultural and forestry export, po stands for poverty rate, lr is for literacy rate, ci is competitive provincial index, lf is labor force with training, nm is net migration and al is agricultural land. Subscript i stands for provinces while t stands for years. Export variable does not have subscript i because it is only available at the national level.

With multiple factors leading increasing in demand such as population or profit from exporting, it should be reasonable to assume changes in land use from growing these crops lead to potential deforestation and forest degradation. Especially, given that Vietnam is a developing country, it means that at the beginning it relies on agriculture which requires large number of tracts as well as labor, and moves away from it later. Therefore, the study predicts that annual crops have a negative correlation with forest coverage.

Previous studies have mentioned the impact of foreign market on environmental degradation in developing countries (Meyfroidt and Lambin 2008,2009; Meyfroidt et al 2013; Ziegler et al 2009; Li and Fox 2012). However, including measurements for this factor has not been included in deforestation studies in Vietnam or in form of direct expansion from internal factor. The more the needs for raw materials, the higher incentives for growing agricultural and forestry products. This paper expects a negative association between agricultural export and forest coverage.

Similar to agricultural export, forestry export attempts to evaluate the external demand using forestry export as percentage of export in total. However, the effect might be moderated because a significant proportion of timber logging in Vietnam is illegal which the data does not necessarily represent. Regardless, it is arguable that if the market demand is strong enough, it is expected that forestry export is negatively correlated with forest coverage. Also, it is also expected that timber extraction will also have a negative association with forest coverage.

The association between poverty and deforestation is inconsistent. Some suggests that reduction in poverty leads to reduction in deforestation. Moderately, others suggest that a small decrease in poverty rate does not necessarily correlate with a similar response in deforestation, but such change must come from improvements from multiple factors including literacy, quality of living, openness for business. Regardless, given a lower rate of poverty, population has access to other options having higher economic payoff or earns consequences of deforestation leading them away from the actions. In both cases, this study expects a negative correlation between poverty rate and forest coverage.

Chapter 3: Data & Methodology

This chapter discusses the sample, data, and variables used for the analysis. Next, a series of different diagnoses are conducted to find the best regression for modeling in the study.

3.1 Data Sources

The data for the study was collected from the General Statistic Office of Vietnam (GSO) (gso.gov.vn), which is a specialized bureau in Vietnam for providing official economic and social statistics. The administrative data is collected by government agencies surveying data local populations. At the end of the year, all statistics are reported back to the GSO to compile into the annual statistical yearbook based on estimates at the national – level or provincial – level. The publication describes their measurements and calculations for each variable in details. Each province also has their statistical yearbook reported in provincial level or subdivision level. As for collecting data, the part after 2010 can be extracted directly from the website organized according to variables. Previous observations are collected by extracting and combining data point from their annual statistical yearbook, which can be found in the archive page of the GSO. Omitted variables from the analysis such as provincial gross domestic product or perennial crop annual production cannot be found on both their website and their annual yearbook.

The study is also based on data from the Provincial Competitive Index (CPI), which is a measurement created by Vietnam Chamber of Commerce and Industry (VCCI) and US Agency for International Development (USAID). This data measures ‘economic governance quality of provincial authorities in creating a favorable business environment for development of the private sector.’ CPI’s measurements aspects including 1) entry costs, 2) land access and security of tenure, 3) transparency and access to information, 4) time costs and regulatory compliance, 5) informal charges, 6) policy bias, 7) proactivity of provincial Leadership, 8) business support services, 9) labor and training and 10) legal institutions. All these indicators are measured with a10-point scale. The PCI main index is the weighted combination of ten subindexes on a 100-point scale ranking governance quality of each province.

3.2 Province Sample

Originally, the dataset contains 64 provinces. Based on the availability of data, it is necessary to perform listwise deletion for complete case analysis. the overall sample of the study is a panel sample consisting of 432 observations (ranging from 2008 – 2015, which is the equivalence of 8 years, and a total of 9 variables). There are 52 provinces in the sample [1]. Given the observations equal the multiplication of the number of provinces and year range total ($432 = 8 * 54$), the panel sample was balanced with 84% of the provinces containing full observations for the period of this study. The panel sample needed to be balanced for the spatial analysis.

3.3. Variables

3.3.1. Dependent Variable

Forest Coverage

I operationalize deforestation with the total forest coverage in a province. Forest coverage is measured as the percentage of a province covered by forest. Data on forest coverage was

[1] 52 provinces used in this analysis are Ba Ria – Vung Tau, Bac Giang, Bac Kan, Bac Lieu, Ben Tre, Binh Dinh, Binh Phuoc, Ca Mau, Cao Bang, Da Nang, Dak Lak, Dak Nong, Dien Bien, Dong Nai, Dong Thap, Ha Giang, Ha Noi, Ha Tinh, Hai Duong, Hai Phong, Hau Giang, Ho Chi Minh City, Hoa Binh, Khanh Hoa, Kien Giang, Lai Chau, Lam Dong, Lang Son, Lao Cai, Long An, Nam Dinh, Nghe An, Ninh Binh, Ninh Thuan, Phu Tho, Phu Yen, Quang Binh, Quang Nam, Quang Ngai, Quang Ninh, Quang Tri, Soc Trang, Son La, Tay Ninh, Thai Binh, Thai Nguyen, Thanh Hoa, Tien Giang, Tra Vinh, Tuyen Quang, Vinh Phuc, Yen Bai

obtained from the GSO (gso.gov.vn). The average forest coverage at the province-level was 4.15% and the average difference in forest coverage between provinces was 4.69%.

3.3.2. Focal Variables

Agricultural Production

Agricultural productions represent annual crops including two variables crop and cereal. Crop is the sum of sugarcane, peanut, and soybean (measured in thousand tons). The average of annual crop production at the province-level was in the sample is 9.159 thousand tons and standard deviation is 11.388 thousand tons. Cereal accounts for volume of cereal per capita (ton/person). The variable represents food factor in section 2.6. Cereal crop has the average in the sample of 507.15 thousand tons and the spread measured by standard deviation of 425.548 thousand tons. These variable measures agricultural expansion. Though, the study originally attempts to include both annual crops (peanut, soybean and sugarcane) as well as perennial crops (cashew, coffee, pepper) as previous study (Meyfroit et al. 2013). However, data for perennial crops after 2005 cannot be located and therefore will not be included.

Agricultural land measures the percentage of agricultural land relative to the area of the province. It represents agriculture factor in section 2.6. Agricultural land has the average of .0445% and the standard deviation of .0438%

Timber Extraction

Timber extraction used to operationalize timber logging measures in volume of timber extract in km^3. The mean and standard deviation for timber extraction is 91.26 km^3 and 166.238 km^3, respectively. Timber extraction only measures the size of legal extraction while leaving out the illegal part.

Agricultural & Forest Exports

Agricultural export and forestry export are the relative of national agricultural and forest product as percentage of total export at the national level by each year. Both variables theoretically estimate the demand of the international market for agricultural and forestry products. In this sample, agricultural export has the mean as 15.23% of the GDP and the difference across the year as 6.487%, respectively. Forestry export has the average of 1.324% of the total export value and the difference of .595% across the years.

Poverty rate

The poverty rate is measured as the percent of household in a given province lived under $15/month. The average value 13.87% of the population and the difference of 10.276% across all provinces from 2008 - 2015.

3.3.3 Control Variables

PCI (Province Competitive Index)

Province Competitive Index estimates how strong a province in terms of openness for business in scale 0 – 100 (CPI website). The data is collected from the office of PCI. For this study, PCI represents governance quality and/or governance competency. PCI has the average in the sample of 58.82 over 100 and the standard deviation of 6.115 over 100

Literacy rate

The literacy rate is measured as percent of population who can read and write. The variable is available at the provincial level. Percent of literacy population has the mean and average difference of 92.64% and 7.087%.

Percent of labor force having training

The variable is defined as the percent of the labor force competed high school education. It is at the provincial level. The variable has the average of 16.07% of the labor force and difference of 7.017%, and

Net migration rate

The variable is the difference between percent of population moving in a province and percent of population moving out a province. It is available for each province with the average of -1.334% and the difference of 8.013%.

Table 1 shows the variable definitions and summary statistics for each variable in the analysis

[Supplemental Files: Table 1]

3.4 Data & Model Diagnostics

This section begins with multiple diagnostics to determine the most appropriate estimator for the regression modelling. However, I begin by examining two important features of multivariate data: collinearity and heteroskedasticity. The most two basic tests are the variance inflation factor (VIF) for multi-collinearity and Breusch-Pagan test for Heteroskedasticity, which are two important assumptions of ordinary least square. Then, the paper follows with the Wooldridge test autocorrelation for determining the extant of temporal dependence in the data. Next, the paper argues for the use of random effect and fixed effect model resulting in the use of Hausman test for better estimator. Finally, the Moran's I spatial autocorrelation test is conducted to test for spatial correlation.

3.4.1 Multivariate Diagnostics

Multicollinearity is attributable to an extremely high correlation between the independent variables in the model. If two or more variables are high correlated, this causes the standard errors to inflate which increases the likelihood of Type II error when testing hypotheses on the effects of the covariates on the dependent variable. Accordingly, I estimate the variance inflation factor (VIF) for each of the independent variables in the model. VIF is a measurement of the degree to which standard errors are inflated in multiple regression model. The VIF measures the degree to which standard errors are over-estimated. For example, a 2.5 VIF indicates the SE is 150% inflated. The threshold for VIF is 10 with some more conservative threshold such as 5 or 2.5. For this analysis, the average VIF among the independent variables is 2.03 while the variable with highest VIF is poverty with 4.04. This value is considered an indicator of multicollinearity with a 2.5 threshold but should be reasonable.

A major assumption in linear regression is the residual variance is constant across the levels of the independent variables (homoskedasticity). However, it is often not the case, because the residual variance increases or decreases across the levels of the independent variables which indicates – heteroskedasticity. This study performs the Breusch-Pagan test for heteroskedasticity. The null hypothesis in the test is that there is homoskedasticity (constant residual variance) while its alternative hypothesis is heteroskedasticity violation (varying residual variance). The Chi-square for the test was 492.95 ($p<.001$). This suggests the residual variance is heteroskedastic. Therefore, to address this problem, I use robust-clustered standard errors when testing the coefficients while directly modeling temporal and spatial autocorrelation.

3.4.2 Panel Data Diagnostics

Panel data possesses potential problem with temporal autocorrelation (serial correlation). Temporal autocorrelation is the correlation between values of different time series within panels.

For ordinary least squares regression, autocorrelation violates the minimum variance property of such analysis causing the estimate no longer is reliable. In this study, I utilize the Wooldridge test for autocorrelation to determine whether the data shows significant temporal autocorrelation. The null hypothesis in the test is that there is no autocorrelation. The F-Ratio of the test was 559.21 ($p<.001$). This suggest the residuals were correlated with each other within the panels. Therefore, to address this problem, I use robust-clustered standard errors and directly model temporal autocorrelation with random- and fixed-effects.

Both models are used to control for unobserved heterogeneity given individual effect for both unobserved time-invariant heterogeneity created by years and unobserved unit-invariant heterogeneity caused by provinces. This creates problem in which residual is correlated with the independent variables in the model which biases the estimation of the coefficients (Halaby 2004). The Hausman test determines whether the coefficients of two estimators are similar where (one estimator is consistent Fixed Effect Model-FEM) and the other estimator is efficient (Random Effect Model-REM). More specifically, it is used to test the 'strict exogeneity' assumption of the REM. The random intercept is uncorrelated with the independent variables in the model. Violating this assumption renders the REM inconsistent. If the assumption of the random effect is correct, then the random effect is preferred because it is a more efficient estimator than the FEM. However, if the assumption is violated, the FEM is preferred because it is a consistent estimator[2].

[2] Unbiasedness, consistency and efficiency are three characteristics determining quality of a statistical model. Unbiasedness refers to how correct on average of the estimator compared to the population parameter. A more efficient estimator has a smaller variance. Consistency refers to the estimator getting closer and closer toward the population parameter as the sampling size increases. However, it is often that these three characteristics cannot be obtained at the same time and there must be a trade-off between each quality.

The null hypothesis is that the REM is preferred while the alternative hypothesis suggests that the FEM is the better fit. The Chi-Square of the Hausman test was 29.22 ($p<.001$). This suggests the strict exogeneity assumption of REMs was violated and the REM are inconsistent. Therefore, I prefer the FEM.

The result of the Hausman test suggests that fixed effect model is a better estimator because the province-level random intercept is likely correlated with the independent variables in the model. For the one-way fixed effect, the individual effect is included as a vector while in the two ways fixed effect, the individual effect is included as a dummy matrix, in which the column of the dummy matrix is the number of individual observations minus one.

3.4.3 Spatial Data Diagnostics

Given that spatial factors have not been addressing in previous studies, assessing spatial autocorrelation is an important part of this study. Spatial autocorrelation is the cluster of observations by space causing spatial dependence (LeSage 2008). To decide whether such analysis is necessary, the Moran's I test is conducted for the data of each distinct year from 2008 – 2015. Moran's I measures spatial autocorrelation based on feature locations and features value. The test evaluates whether the pattern of a spatial dataset is clustered, dispersed, or random. If the p value is insignificant, the null hypothesis is rejected then the spatial distribution is potentially due to randomness. In case of the p value is significant, then it is necessary to consider the z score. If the z score is positive, the spatial pattern is likely to be clustered while it is potentially dispersed given a negative z score. As for this analysis, the Moran's I p value is larger than the threshold of 0.10 indicating that the spatial pattern is quite likely due to randomness.

3.5 Regression Estimators

The data for this analysis is a panel data, which is also referred to as longitudinal or cross-sectional time-series data. It includes multiple variables of a set of observations across time and the same units. Normally, ordinary least squared (OLS) is the simplest approach for estimating a multivariate regression model. It has the following form

$$(1) y_{i,t} = \alpha + X_{i,t}\beta + u_{i,t}$$

In Equation 1, $y_{i,t}$ is the dependent variable vector which represents the forest coverage for province i in year t. $X_{i,t}$ is the vectors including focal variables and control variables. $u_{i,t}$ is the idiosyncratic error for province i in year t. β is the coefficient vector of independent variables while α is the intercept.

OLS is built on 6 basic assumptions. 1) linear in parameter, 2) error are independent and identically distributed, 3) the conditional mean is zero, 4) No multi-collinearity (perfect collinearity), 5) no heteroskedasticity, and 6) no autocorrelation. If all conditions are satisfied, then OLS is the best linear unbiased estimators or B.L.U.E. Otherwise, any violation leads to OLS no longer B.L.U.E and therefore, there exists better linear estimators for the data. The reason is by satisfying the (2) assumptions of independent and identically distributed, it is already indicative that the observations cannot have autocorrelation, which is assumption #6. However, the data for this analysis is in panel form with individual dimension and time dimension. In this case, the individual dimension is the number of provinces while the time aspect is the number of years in the dataset.

Both models are used to control for unobserved heterogeneity given individual effect for both unobserved time-invariant heterogeneity created by years and unobserved unit-invariant heterogeneity caused by provinces. This creates problem in which residual is correlated with the independent variables in the model which biases the estimation of the coefficients (also a violation of the third assumption of OLS). Also, the nature of panel data makes the error terms no longer is dependent and identically distributed, which is a violation of the second assumptions above. This problem can be addressed by introducing the either the residual terms vector in random effect model or including the fixed intercepts for individual effect in fixed effect model. The random effect has the form

$$(2)\ y_{i,t} = \alpha + X_{i,t}\beta + U_i + u_{i,t}$$

In Equation 2, the difference with equation (1) is the term U_i, which is a panel residual term for random effect with addressing the unobserved unit-invariant.

One-Way Fixed effect has the form

$$(3)\ y_{i,t} = X_{i,t}\beta + \alpha_i + u_{i,t}$$

α_i is a vector of panel specific intercepts addressing the unobserved time-invariant heterogeneity.

Two-Way Fixed effect has the form

$$(4)\ y_{i,t} = X_{i,t}\beta + \alpha_i + \ a_t + \ u_{i,t}$$

The term α_i is a vector of panel – specific intercepts addressing the unobserved time-invariant while the term α_t is a vector of year-specific intercepts addressing unobserved unit-invariant heterogeneity.

Prais-Winston (PW) regression addresses the problem of autocorrelation, which is assumptions #6 of OLS estimate. Normally, the data is processed with first difference but results in losing the first observation of each panel. Prais -Winston model solves this problem by weighting the first difference by a modified parameter which increases the efficiency of the estimator. The Prais-Winston regression has the form:

$$(5a)\ y_{it} - \hat{r}y_{it-1} = (X_{it} - \hat{r}X_{it-1})\beta + u_{it}$$

$$(5b)\ \hat{r} = \sqrt{1 - \rho^2}$$

In equation 5, the function is a first difference regression. The equation goes through a rho transformation by subtracting the first lag vector of forest coverage on the left-hand side and the first lag matrix of the independent variables on the right-hand side. The term u is the idiosyncratic error for province i and year t. The term $\hat{r}$ in function 5b is the rho hat autocorrelation parameter as a squared value of the difference between 1 and rho square.

Spatial regression accounts for spatial dependency in the regression model by including a weighted spatial matrix in the independent part (LeSage 2008). The weighted spatial matrix is a representation of the spatial structure in the dataset by quantifying the adjacent observations.

$$(7)\ y_{it} = \rho W y_{it} + \beta X_{it} + a_i + \delta W u_{it}$$

In this equation, there is an addition of the term ρW, which provides spatial weight for the forest coverage in province i and year t. The term $W$is the spatial matrix. A similar term $\delta W$is used to similar purposes for the error terms u. All other terms y, X and α remains the same purposes as the fixed effect one way model.

3.6 Model Assessment

I assess regression estimators based on examining how these models fit the data. Specifically, I use the Akaike Information Criterion (AIC) and Bayesian Information Criterion

(BIC). Both statistics are measures of fitness as alternative to R Squared or Adjusted R Squared. BIC takes account of the number of observations in the model while AIC does not.

AIC has the form

$$(8) AIC = 2k - 2\ln(\hat{L})$$

In this equation, k is the number of estimators in the model and $\hat{L}$ is the maximized value of the likelihood function of the model. AIC is the difference between twice value of k and the natural log of $\hat{L}$

BIC has the form

$$(9) BIC = kln(n) - 2\ln(\hat{L})$$

In this equation, n is the number of observations. BIC is the difference between the product of k and natural log of n and twice value of natural log of $\hat{L}$

[Supplemental Files: Table 2]

Both measurements are used to evaluate relative quality of statistical model. Specifically, model having lowest relative measurements is better. Also, a 10-point difference in AIC and BIC usually indicates a better fit. Based on AIC and BIC statistics, one way fixed effect has the lowest BIC values while two ways fixed effect has the lowest AIC values. However, BIC accounts for the number of observations in the statistical model while AIC does not. This indicates fixed effect one way is the best estimator.

Chapter 4: Results

4.1 Introduction

This study addresses drawbacks in previous research that use official data. These drawbacks include (1) the focusing on limited areas instead of a national scale (Meyfroidt et al 2013), (2) exclusion of resources trade from the analysis, (3) lacking consideration of the spatial and longitudinal autocorrelation, which is overlooked in panel analysis (Quy et al. 2018); and the inconsistency of the relation between poverty and deforestation (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003; Miyamoto 2020; Quy et al 2018).

The data is at province-level providing a more comprehensive picture of deforestation in Vietnam. To tackle the second problem, the inclusion of agricultural and forestry trade helps improve the explanatory power of the theoretical framework. Next, the study tests for spatial and temporal dependence. Lastly, this study uses fixed effect model (FEM) and spatial regression to address previous drawbacks accordingly. The fitness tests suggest fixed effect one way is the best estimators

$$(1)\ fc_{i,t} = \alpha_i + \beta_1 ac_{i,t} + \beta_2 tb_{i,t} + \beta_3 ev_t + \beta_4 po_{i,t} + \beta_5 lr_{i,t} + \beta_6 ci_{i,t} + \beta_7 lf_{i,t} + \beta_8 nm_{i,t} + \beta_9 al_{i,t} + u_{i,t}$$

fc stands for forest coverage in province i and year t, ac is for annual crops including crops for sugarcane, peanut and soybean, and cereal crops, tb is for timber extraction, ev is for export variables including agricultural and forestry export, po stands for poverty rate, lr is for literacy rate, ci is competitive provincial index, lf is labor force with training, nm is net migration and al is agricultural land. Subscript i stands for provinces while t stands for years. Export variable does not have subscript i because it is only available at the national level. α_i is a vector of panel – specific intercepts addressing the unobserved time-invariant $u_{i,t}$ is the idiosyncratic error for province i in year t. β_n are the coefficients of the variables.

4.2 Theoretical Framework & Hypotheses

As mentioned in the theoretical framework, the basic of this analysis is built on the association between deforestation and agricultural expansion, timber logging, resource trading and multiple subfactors of socioeconomic drivers.

Agricultural expansion satisfies the resources for a private economy while also creating more jobs with better living standards for a part of the population. However, the forest along with local population are often left out of the picture. In previous studies, both annual and perennial crops have positive association with deforestation and forest degradation (Meyfroit et al. 2013). Therefore, this study expects such same result with annual crop. The first hypothesis is that there should be a negative correlation between annual crop and forest coverage.

$$H_1: Annual\ crop\ is\ negatively\ associated\ with\ forest\ coverage$$

The second factor of the framework is timber extraction. Timber extraction has been consistently rising both legally and illegally. Though the illegal part is often underreported if not non-exist in the official data which affects the ability to capture the true scale of timber logging, the high demand for forestry products would still manifest through official data. The studies therefore still anticipate a negative correlation between forest coverage and timber extraction

$$H_2: Timber\ extraction\ is\ negatively\ associated\ with\ forest\ coverage$$

This study introduces exports variable as a way to account for the effects of the world economy on forest coverage in Vietnam. The expansion of the international market is another factor accelerate deforestation/ forest degradation. According to the unequal ecological exchange perspective, developed countries externalize environmental cost to less-developed counterparts by driving up the needs for certain products. More developed countries can externalize their environment costs to developing countries through payments for products. This creates environmental degradation in exporting countries, specifically deforestation in this case. It has been consistent that elements of economic development for developing countries such as agricultural expansion and exportation are positively correlated with deforestation. Developing countries have a large proportion of GDP earned from exporting extracted raw materials, particularly forestry and agricultural products. Therefore, by including them, this study predicts a negative association between forest coverage and agricultural export as well as forestry exports.

$$H_3: Agricultural\ export\ is\ negatively\ associated\ with\ forest\ coverage$$

$$H_4: Forestry\ export\ is\ negatively\ correlated\ with\ forest\ coverage$$

The need for timber extraction and agricultural production are crucial for development in Vietnam because they provide raw materials as well as food security for the population. However, at the same time, partial the demand comes internationally which makes it difficult to evaluate impact of each market separately. What is the magnitude of each market and their effect on the other? This study interests in the relationship between two markets. The hypothesis expects the resources trade amplifies the negative association between agricultural production, forest extraction with forest coverage.

H_5: Forestry export amplify the negative association between timber extraction and forest coverage

H_6: Agricultural exports amplify the negative association between agricultural production and forest coverage.

Poor communities need land for agricultural production. However, poor cultivating techniques often leads the soil worsen quickly, therefore they need to move to another area and log trees again for new soil. It has been controversial on the effect of poverty on deforestation, though poverty has a negative association with forest coverage. Whether poverty is the sole factor affecting deforestation or other covariates share association as well. Some studies suggests that improvement in poverty rate has a positive impact on deforestation while others disagree such result. They show that a slight change in only poverty does not have any effect or negative effect on deforestation. Such improvement must complement with other factors such as living standards, training in labor, etc. Therefore, this study tests for the association between poverty rate and forest coverage. The expectation is that poverty will have a positive correlation with forest coverage.

$$H_7: Poverty\ is\ negatively\ correlated\ with\ forest\ coverage$$

4.3 Analytical Strategy

I test the hypotheses described above in two stages. In the first stage, I estimate Pearson correlations to measure the association between forest coverage and the covariates. This is a preliminary test of the association. However, correlations measure zero-order associations which are susceptible to spuriousness/confounding.

Therefore, in the second stage, I estimate multivariate regression models of forest coverage.

Both fixed effect and random effect are used to control for unobserved heterogeneity given individual effect for both unobserved unit-invariant heterogeneity created by years and unobserved time-invariant heterogeneity caused by provinces. This creates problems of endogeneity in which residuals are correlated with the independent variables in the model which biases the estimation of the coefficients (Halaby 2004). More specifically, it is used to test the 'strict exogeneity' assumption of the REM. The assumption is the random intercept is uncorrelated with the independent variables in the model. Violating this assumption renders the REM inconsistent. If the assumption of the random effect is correct, then the random effect is preferred because it is a more efficient estimator than the FEM. However, if the assumption is violated, the FEM is preferred because it is a consistent estimator. A more efficient estimator has a smaller spread range. Consistency refers to the estimator getting closer and closer toward the population parameter as the sampling size increases.

The Hausman test in Chapter 3 suggests the strict exogeneity assumption of REMs are violated in the data. Additionally, the AIC and BIC of different estimators shows the one-way fixed effects models fit the data best. Recall the fixed effect one way model has the function:

$$(2)\ fc_{i,t} = \alpha_i + \beta_1 ac_{i,t} + \beta_2 tb_{i,t} + \beta_3 ev_t + \beta_4 po_{i,t} + \beta_5 lr_{i,t} + \beta_6 ci_{i,t} + \beta_7 lf_{i,t} + \beta_8 nm_{i,t} + \beta_9 al_{i,t} + u_{i,t}$$

In Equation 9, fc stands for forest coverage in province i and year t, ac is for annual crops including crops for sugarcane, peanut and soybean, and cereal crops, tb is for timber extraction, ev is for export variables including agricultural and forestry export, po stands for poverty rate, lr is for literacy rate, ci is competitive provincial index, lf is labor force with training, nm is net migration and al is agricultural land. Subscript i stands for provinces while t stands for years. Export variable does not have subscript i because it is only available at the national level. α_i is a vector of panel – specific intercepts addressing the unobserved time-invariant. $u_{i,t}$ is the idiosyncratic error for province i in year t

As a robust check, I re-estimate the FEMs using a spatial autoregressive estimator. The spatial model includes a spatial matrix accounted for the adjacent correlation between provinces. The spatial fixed effect model has the function:

$$(3)\ fc_{i,t} = \rho W y_{it} + \alpha_i + \beta_1 ac_{i,t} + \beta_2 tb_{i,t} + \beta_3 ev_t + \beta_4 po_{i,t} + \beta_5 lr_{i,t} + \beta_6 ci_{i,t} + \beta_7 lf_{i,t} + \beta_8 nm_{i,t} + \beta_9 al_{i,t} + \delta W u_{i,t}$$

In this equation, there is an addition of the term ρW, which provides spatial weight for the forest coverage in province i and year t. The term $W$is the spatial matrix. A similar term $\delta W$is used to similar purposes for the error terms u. All other terms y, X and α remains the same purposes as the fixed effect one way model.

4.4 Correlation Analysis

[Supplemental Files: Table 3]

Forest coverage has correlation of .095, -.236 and .182 with annual crop, cereal crop and agricultural land. Forest coverage almost has no relation with annual crop and agricultural land while has a weak negative relation with cereal crop (.087). Timber extraction has no relation to forest coverage which are similar to both export variable-agricultural and forestry (.022 and .0188, respectively). Poverty rate has a moderate positive (.525) relation with forest coverage.

4.5 Multivariate Regressions

[Supplemental Files: Table 4]

In this section, the study evaluate hypothesis using multivariate regression models which measures the partial effects of the covariates on forest coverage. Table 4 show the FEMs of forest coverage.

There are three models in the table. The first model excludes interaction effect between covariates to directly test the Hypothesis 1,2,3,4,7, which examines the association between forest coverage and agricultural expansion, forest extraction, resources trade and poverty rate

The second model includes interaction effect between timber extraction and forestry export to test Hypothesis 5-forestry export amplifies the association between timber extraction and forest coverage. The third model includes the interaction terms between agricultural land and agricultural export to test Hypothesis 6, in which agricultural export amplifies the effects of agricultural expansion on forest coverage.

The first hypothesis H_1considers the association between annual crop and forest coverage. Model 1 in Tables 4 show indicators of agricultural production are not significantly associated with forest coverage ($p>.10$). Therefore, I do not find any supporting evidence for the first hypothesis.

The second hypothesis H_2 expects a negative association between forest coverage and timber extraction. However, model 1 of both types of fixed effect does not result in any significant association ($p>.10$).

As for hypothesis H_3, H_4, H_5 and hypothesis H_6, Model 1 shows that exports are not significantly associated with forest coverage. In Model 2 and Model 3, the interaction terms between timber extraction and forestry exports, and agricultural land and agricultural export are also insignificant. These results, therefore, do not confirm both hypotheses (p > .10).

The only supported hypothesis is H_7 on the association between poverty and forest coverage. In model 1, poverty is negatively associated with forest coverage (p < .10). For fixed effect one way model, one point increasing in the poverty rate leads to a decrease of .037 percent in forest coverage, holding all else constant. In 2015, the province with the highest poverty rate is Lai Chau with 32% and a forest coverage of 4.591%. If poverty reduced to the mean national level of 14.57%, forest coverage of Lai Chau would have increased to 5.234%.

4.6 Spatial Regression

[Supplemental Files: Table 5]

The Moran I test suggests that there is no spatial correlation between provinces. However, as a robustness check, I estimate models that account for spatial autocorrelation in forest coverage and the residual. Table 5 shows estimate from spatial autoregressive models with province level fixed effects. Models 1-3 confirms that poverty is negatively associated with forest coverage. A one-point increasing in percent of poverty leads to a decreasing of .037% in forest coverage, holding all else constant which is similar to normal fixed effect models in table 3. Surprisingly, literacy rate and agricultural lands are also significant (p <.001). A one point increasing in literacy rate leads to a .06 percent increasing in forest coverage in model 1, holding everything constant. In other models, the effect is relatively the same except model 3 with .061% increment in forest coverage. In 2015, the province with the lowest literacy rate is Lai Chau with

59.2% and forest coverage is 4.591%. If the literacy rate raised the national mean of 92.36%, forest coverage in Lai Chau reaches 6.581%. With agricultural land, as for model 1, a one-point increasing in agricultural land associates with .081% reducing in forest coverage, holding all else constant. The effect changes slightly with -.086 in model 2 and -.076 in model 3. In 2015, Da Nang has the lowest agricultural land relative to total land area of .029% and forest coverage rate of .246%. If agricultural land raised to the national mean of 4.365%, Da Nang's forest coverage would have been -.105%, which does not make sense in reality because it is a negative number.

4.7 Discussion & Conclusion

This study addresses four problems of previous research: (1) focusing on particular regions rather than national conditions of forest (Meyfroidt et al 2013); (2) exclusion of resources exchange (Quy et al 2018); (3) missing consideration of longitudinal and spatial autocorrelation in panel analysis (Quy et al 2018); and (4) the inconsistency on whether poverty is the only factor associated with deforestation (Miyamoto 2020). This study addresses the first problem by creating a province-level model to assess deforestation from 2008 – 2015. The descriptive analysis suggests that forest is still in reforestation state which are consistent with previous studies using official data (Meyfroidt and Lambin 2013, Quy 2018).

This paper includes resources trade following the literature on unequal ecological exchange (Jorgenson 2006; Hornborg 2001; Rice 2007) to improve the explanatory power of deforestation with panel data (Quy et al 2018). Developing countries experience environmental degradation through exporting raw material as comparative advantages (Jorgenson 2006; Hornborg 2001; Rice 2007). In the case of Vietnam, exports of agricultural and forestry products grow rapidly (De Koninck 1999; Meyfoidt and Lambin 2009). Agricultural expansion and forest extraction has a positive association with deforestation as for both annual and perennial crops

(Meyfroidt and Lambin 2008, 2009, 2011; Meyfroidt et al. 2013; De Koninck 1999; Ziegler et al 2010; Defries et al 2010). Although both types of exploitation are created by domestic forces, this paper argues that the international market either have direct association or amplification effect on agricultural expansion as well as forest extraction. The results provide evidence for neither of the arguments. In addition, the effects between agricultural expansion and forest extraction with forest coverage are insignificant which contradict with all previous studies. (Meyfroidt and Lambin 2008, 2009, 2011; Meyfroidt et al. 2013; De Koninck 1999; Ziegler et al 2010; Defries et al 2010)

This study tests for temporal with the Wooldridge tests and spatial autocorrelation with the Moran's I test. The result suggests an autocorrelation with time though there is no evidence for spatial dependence. Further fitness tests with AIC and BIC statistics also suggest that panel data analysis provides better results than ordinary least square model, which is an improvement with previous study (Quy et al. 2018).

Lastly, this paper tests the association between deforestation and the poverty rate. Given the inconsistency of whether other factors along poverty rate is associated with forest coverage, this study includes a set of covariates including population, literacy rate, net migration, labor force with training. The fixed effect one way show evidence for none of these control variables, which is consistent with (Miyamoto 2020; Quy et al 2018). As for poverty rate, a one point increasing in poverty rate leads to a decreasing of .037 percent in forest coverage which is weaker than .056 in previous study (Quy et al. 2018). In 2015, the province with the highest poverty rate is Lai Chau with 32% and a forest coverage of 4.591%. If poverty raised to the mean national level of 14.57%, forest coverage of Lai Chau would have reduced to 31.63%. The spatial fixed effect model suggests literacy rate and agricultural land are also have significant relationship with forest coverage along with poverty rate which partially consistent with social values (Ferraro and Simorangkir 2020) and land use (Geist and Lambin 2003).

Chapter 5: Conclusion

Forest coverage in Vietnam has experienced a period of decline before 1990s (Meyfoidt and Lambin 2008; Meyfroidt et al. 2013). Starting in early the 1990s, provinces across the country implemented series new policies to protect forest area because of new environmental requirements as well as better understanding of impact of deforestation (Meyfroidt and Lambin, 2008). Though these positive changes lead forest in Vietnam into a reforestation period (Meyfroidt and Lambin, 2013), studies show that such activities do not necessarily decline but rather getting more sophisticated in form of forest degradation (JICA, 2012; Meyfroidt et al. 2013; Quy et al., 2018). Deforestation is defined as (Hosonuma et al. 2012) as removal of trees and conversion of land for different purposes, one of which is significantly economic. Conversely, forest degradation is referred as 'a reduction in the capacity of a forest to produce ecosystem services' caused by 'unsustainable logging, agricultural, invasive species, fire, fuelwood, gathering and livestock grazing' (Thompson et al. 2013), which is harder to either observe or identify. These findings are enabled with the use of original techniques such as remote sensing processing and geographic information systems (Meyfroidt et al. 2013) rather than the use of official data.

This study acknowledges these findings but insists on using official data to address some drawbacks of previous studies. These drawbacks include (1) the focusing on limited areas, Central Highlands areas (Meyfroidt et al. 2013), instead of a national scale, (2) exclusion of resources trade from the analysis, (3) lacking consideration of the spatial and longitudinal autocorrelation, which is overlooked in panel analysis (Quy et al. 2018); and (4) the inconsistency of the relation between poverty and deforestation (Ferraro and Simorangkir 2020; Zwane 2007; Geist and Lambin 2003; Miyamoto 2020; Quy et al 2018). Addressing these problems contributes to the relation between deforestation with agricultural production as well as demographic factors which is significant for drafting development policy following by implementation on groups of populations, especially the poor and marginal communities.

This paper first develops a model of province-level deforestation from 2008 – 2015 to assess the national condition of forest. Second, agricultural exports and forestry exports are included into a theoretical framework where changes in forest coverage is a linear function of annual crop production, timber extraction, and poverty rate. The additions provide more explanatory power because it accounts for the effects of resources trade missing in previous studies. Next, this study addresses problem of panel data which is obtained from the General Statistic Office of Vietnam and the Provincial Competitive Index. The data spreads across time and provinces making it a panel data which causing violation of the assumptions of independent & identical distribution and zero conditional mean if we use traditional estimator such as OLS. Instead, the paper tackles unobserved endogeneity by using random effect and fixed effect model. In addition, this paper also addresses the problem of spatial dependence which was not considered previously. Lastly, to deal with the inconsistency of poverty with forest coverage, the study includes a set of control variables including population, literacy rate, net migration, labor force with training.

The theorical framework is combination of agricultural expansion, timber extraction, resources trade, and poverty on deforestation/forest degradation. Agricultural expansion is the domestic force driven by economic policies. By moving from a collective to a more private economy, Vietnam has an immense need for raw materials, jobs creations as well as guaranteeing food security. Such policies give rise to inflows of immigrants to original forest area. Timber extraction aims at certain types of trees due to higher market values (Meyfroidt and Lambin 2009) or for general purposes such as fuelwood, charcoal production, uncontrolled fire and livestock grazing leading to forest degradation (Hosonuma et al 2012). In addition, timber logging provides valuable materials for products such as: furniture, medicine, beverage, etc. Resources trade follows the equal ecological exchange (Jorgenson 2006; Hornorg 2001; Rice 2007) in which the developed countries externalize their environmental costs to developing countries through importing products. Less developed countries rely on raw materials export as same as cheap labor for comparative advantage. Exports of wood products from Vietnam have grown rapidly (Meyfroidt and Lambin 2008, 2009) which justified the inclusion of resources trade in the framework. In addition, Agricultural production is tied with deforestation, particularly in poor and ethnic minority communities (Meyfroidt et al. 2013; Meyfroidt and Lambin 2008, 2011; DeFries et al. 2010). Specifically, marginalized community needs lands for agricultural subsistence and therefore clear forested areas for cultivation. However, without the use of fertilizer or outdated agricultural techniques, the soil become exhausted after a short period of time. It means that they can no longer exploit the land anymore. The community then has to move to a new area as well as cut down more tree for next production cycle. Therefore, poverty is a proxy used to account for deforestation of impoverished communities. These four parts are the main factors of this paper’s arguments for deforestation/forest degradation in Vietnam.

This study tests for a total of 7 hypotheses addressing the theoretical framework above. There are three models. The first model excludes interaction effect between covariates to directly test the Hypothesis 1,2,3,4,7, which examines the association between forest coverage and agricultural expansion, forest extraction, resources trade and poverty rate. The second model includes interaction effect between timber extraction and forestry export to test Hypothesis 5- forestry export amplifies the association between timber extraction and forest coverage. The third model includes the interaction terms between agricultural land and agricultural export to test Hypothesis 6, in which agricultural export amplifies the effects of agricultural expansion on forest coverage. There are no evidence supporting for the association between agricultural expansion, timber extraction and forest coverage. The same results are consistent with resources trade and forest coverage. As for model 2 and 3, there is also no evidence that resources trade amplifies the effects of either agricultural expansion or timber extraction. Only Hypothesis 7 measures the association between the poverty rate and forest coverage is negatively associated. The effect is similar in the spatial analysis but with the additional evidence of literacy rate and agricultural land.

A large limitation of this study is the use of official dataset given the contradiction in previous findings and official number of forest conditions. The descriptive analysis also shows that forest coverage keeps increasing over the period from 2008 to 2015 nationally. The official data might not be a good indicator to track sophisticated extraction methods as forest degradation. This leads to the next limitations of this study which is the complete lacking qualitative data. An analysis at large scale leaves out the unique conditions of every district, city or province. It also lacks insights of the people experiencing these changes including farmers, workers, business owners, people of ethnic and impoverished communities, or policy makers. Both weaknesses can be addressed with different data types. Previous research has already applied original techniques such as remote sensing processing or geographic information systems

(GIS) to track forest changing over time. In addition, the use of qualitative methods can provide perspectives on how agricultural expansion affects deforestation in different areas.

The following limitations are due to the nature of secondary data. First, a limitation of this study is the constrains the number of year due to data availability. Panel data for spatial regression requires absolute balance. Although some missing observations can be reconstructed using interpolation and extrapolation with limited uses, a more completed dataset ranging from a longer period might create a better estimator. Missing data for perennial crops as well as illegal timber logging causes potential omitted variables bias. Also, the agricultural and forestry export variables is at national level instead of the provincial level. These problems are harder to address because of the GSO's inconsistency in collecting certain data. A possible solution is to perform study with a dataset from independent organizations, which might not exist because of the scale needed to collect such data. Also, a more appropriate approach might be the district level data of the GSO which is available only in purchasing. However, such data will subject to the same problem mentioned above.

Bibliography

Baulch, B., Hoa, T.M.N., Phuong, T.T.P., Hung, T.P., (2010). Ethnic Minority Poverty in Vietnam. Working Paper 169, Chronic Poverty Research Centre, Manchester, UK.

Cramb, R.A., Colfer, C.J.P., Dressler, W., Laungaramsri, P., Quang, T.L., Mulyoutami, E., Peluso, N.L., Wadley, R.L., 2009. Swidden transformations and rural livelihoods in Southeast Asia. *Human Ecology* 37, 323–346.

Cruz, Wilfrido, Donald Grebner, and William F. Hyde. (2013). The Board of Regents of the University of Wisconsin System Environmental Motivations for Migration: Population and Deforestation in the Philippines Pressure, Poverty Gregory S. Amacher, Wilfrido Cruz, Donald Grebner, and William F. Hyde. 74(1):92–101.

DeFries, R.S., Rudel, T., Uriarte, M., Hansen, M., (2010). Deforestation driven by urban population and agricultural trade in the twenty-first century. *Nature Geosciences* 3, 178–181.

De Koninck, R., (1999). Deforestation in Vietnam. *International Development Research Center (IDRC)*, Ottawa.

Drukker, David M. (2003). Testing for Serial Correlation in Linear Panel-Data Models. *The Stata Journal: Promoting Communications on Statistics and Stata* 3(2):168–77. doi: 10.1177/1536867x0300300206.

Faridi, M. Z. (2012). Contribution of agricultural exports to economic growth in Pakistan. *Pakistan Journal of Commerce and Social Sciences (PJCSS)*, *6*(1), 133–146. http://hdl.handle.net/10419/188047

Ferraro, P. J., & Simorangkir, R. (2020). Conditional cash transfers to alleviate poverty also reduced deforestation in Indonesia. *Science Advances*, *6*(24), eaaz1298. https://doi.org/10.1126/sciadv.aaz1298

General Statistics Office. 1990-2019. 'Statistical Summary Book of Vietnam.'

Halaby, Charles N. (2004). Panel Models in Sociological Research: Theory into Practice. *Annual Review of Sociology* 30(Villareal 2002):507–44. doi: 10.1146/annurev.soc.30.012703.110629.

Ha, D.T., Shively, G., (2008). Coffee boom, coffee bust and smallholder response in Vietnam's Central Highlands. *Review of Development Economics* 12 (2) 312–326.

Hardy, A., (2000). Strategies of migration to upland areas in contemporary Vietnam. *Asia Pacific Viewpoint* 41 (1) 23–34.

Hardy, A., (2003). Red Hills: Migrants and the State in the Highlands of Vietnam. *University of Hawai'i Press*, Honolulu.

Hornborg, A. (2006) Footprints in the Cotton Fields: The Industrial Revolution as Time- Space Appropriation and Environmental Load Displacement. *Ecological Economics* 59: 74–81.

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., Angelsen, A. and Romijin, E. (2012) An assessment of deforestation and forest degradation drivers in developing countries. *Environ. Res. Lett*. 7, 040009

Hosonuma, Noriko, Martin Herold, Veronique De Sy, Ruth S. De Fries, Maria Brockhaus, Louis Verchot, Arild Angelsen, and Erika Romijn. 2012. An Assessment of Deforestation and Forest Degradation Drivers in Developing Countries. *Environmental Research Letters* 7(4). doi: 10.1088/1748-9326/7/4/044009.

JICA, (2012). Potential Forests and Land Related to Climate Change and Forest in Vietnam. *Vietnam National University of Forestry*, Hanoi, Vietnam (219 pp)

Jorgenson, A.K. (2006) Unequal Ecological Exchange and Environmental Degradation: A Theoretical Proposition and Cross-National Study of Deforestation, 1990–2000. *Rural Sociology* 71(4): 685–712.

Jorgenson, A.K., Dick, C. and Austin, K. (2010) The Vertical Flow of Primary Sector Exports and Deforestation in Less-Developed Countries: A Test of Ecologically Unequal Exchange Theory. *Society and Natural Resources* 23(9): 888 – 897.

Lambin, E.F., Meyfroidt, P., (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3465–3472.

Lambin, E. F., & Geist, H. J. (2003). Regional Differences in Tropical Deforestation. *Environment*, *45*(6), 22–36. https://doi.org/10.1080/00139157.2003.10544695

LeSage, J. P. (2008). An introduction to spatial econometrics. In *Revue d'Economie Industrielle* (Vol. 123, Issue 3, pp. 19–44). *Editions Techniques et Economiques*. https://doi.org/10.4000/rei.3887

Li, Z., Fox, J.M., (2012). Mapping rubber tree growth in mainland Southeast Asia using time-series MODIS 250 m NDVI and statistical data. *Applied Geography* 32 (2) 420–432.

Meyfroidt, P., Lambin, E.F., (2008). The causes of the reforestation in Vietnam. *Land Use Policy* 25 (2) 182–197.

Meyfroidt, Patrick, and Eric F. Lambin. (2009). Forest Transition in Vietnam and Displacement of Deforestation Abroad. *Proceedings of the National Academy of Sciences of the United States of America* 106(38):16139–44. doi: 10.1073/pnas.0904942106.

Meyfroidt, P., Vu, T. P. and Hoang, V. A. (2013) Trajectories of deforestation, coffee expansion and displacement of shifting cultivation in the Central Highlands of Vietnam. 23 1187 – 1198.

Miyamoto, Motoe. (2020). Poverty Reduction Saves Forests Sustainably: Lessons for Deforestation Policies. *World Development* 127:104746. doi: 10.1016/j.worlddev.2019.104746.

Murtazashvili, Ilia, Jennifer Murtazashvili, and Raufhon Salahodjaev. (2019). Trust and Deforestation: A Cross-Country Comparison. *Forest Policy and Economics* 101:111–19. doi: https://doi.org/10.1016/j.forpol.2019.02.001.

Nguyen, Thanh Hai. (2016). Impact of Export on Economic Growth in Vietnam: Empirical Research and Recommendations. *International Business and Management* 13(3):45–52. doi: 10.3968/9040.

Pfaff, A., & Walker, R. (2010). Regional interdependence and forest “transitions”: Substitute deforestation limits the relevance of local reversals. *Land Use Policy*, *27*(2), 119–129. https://doi.org/10.1016/j.landusepol.2009.07.010

Rambo, A. Terry. (1996). “Vietnam's Rural Transformation. Edited by Benedict J. Tria Kerkvliet and Doug J. Porter. Boulder, Colo.: Westview Press, 1995. Xiv, 215 Pp. $54.95.” *The Journal of Asian Studies* 55(3):784–86.

Rice, J. (2007) Ecological Unequal Exchange: Consumption, Equity, and Unsustainable Structural Relationships within the Global Economy. *International Journal of Comparative Sociology* 48(1): 43–72.

Rudel, T.K., DeFries, R., Asner, G.P., Laurance, W.F., (2009). Changing drivers of deforestation and new opportunities for conservation. *Conservation Biology* 23, 1396–1405.

Thompson, I.D., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymell, V., Sabogal, C., (2013). An operational framework for defining and monitoring forest degradation. *Ecol. Soc*. 18, 20.

Tsurumi, Tetsuya, and Shunsuke Managi. (2014). The Effect of Trade Openness on Deforestation: Empirical Analysis for 142 Countries. *Environmental Economics and Policy Studies* 16(4):305–24. doi: 10.1007/s10018-012-0051-5.

Quy, V. K., Tran, B. Q., Meyfroidt, P. and Paschke, M. W. (2018) Drivers of deforestation and forest degradation in Vietnam: An exploratory analysis at the national level. *Forest Policy and Economic*s 90 128 – 141.

Vogelmann, James E., Phung Van Khoa, Do Xuan Lan, Jacob Shermeyer, Hua Shi, Michael C. Wimberly, Hoang Tat Duong, and Le Van Huong. (2017). Assessment of Forest

Degradation in Vietnam Using Landsat Time Series Data. *Forests* 8(7). doi: 10.3390/f8070238.

Vu, Tam Bang, Eric Iksoon Im, Karla Hayashi, and Ryan Torio. (2017). Cyclones, Deforestation, and Production of Food Crops in Vietnam. *Economics of Disasters and Climate Change* 1(3):245–62. doi: 10.1007/s41885-017-0010-5.

Ziegler, A.D., Bruun, T.B., Guardiola-Claramonte, M., Giambelluca, T.W., Lawrence, D., Nguyen, T.L., (2009). Environmental consequences of the demise in Swidden in Montane Mainland Southeast Asia: hydrology and geomorphology. *Human Ecology* 37, 361–373.

Zwane, Alix Peterson. 2007. "Does Poverty Constrain Deforestation? Econometric Evidence from Peru." *Journal of Development Economics* 84(1):330–49. doi: 10.1016/j.jdeveco.2005.11.007.