University of California
Santa Barbara

# Responses to Institutional Constraints

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

Doctor of Philosophy
in
Economics

by

Ryan Michael Abman

Committee in charge:

Professor Robert Deacon, Chair
Professor Douglas Steigerwald
Professor Paulina Oliva

September 2015

The Dissertation of Ryan Michael Abman is approved.

---

Professor Douglas Steigerwald

---

Professor Paulina Oliva

---

Professor Robert Deacon, Committee Chair

July 2015

Responses to Institutional Constraints

## Acknowledgements

First and foremost, I thank my wonderful wife, Zamira Abman, for her unconditional support and grace through the successes and setbacks that come with graduate studies. Her encouragement and positive attitude know no bounds. I am also incredibly thankful to my incredible and supportive family. My parents, Steven and Carolyn Abman, have been a source of both inspiration and strength.

I will forever remain grateful for the guidance provided by my three committee members. Robert Deacon offered tremendous feedback, vision and perspective on my ideas and always guided me towards the interesting underlying economic questions. Douglas Steigerwald pushed me to improve each draft of every paper and was always available for advice every step of the way. Paulina Oliva offered incredible technical expertise and has forever shaped how I approach empirical work. Numerous other faculty members at UCSB helped shape my work including Javier Birchinall, Chris Costello, Gary Libecap, Kyle Meng and Dick Startz.

My work was also shaped by talented and hardworking individuals in North Hall. I extend a sincere 'thank you' to Zach Bethune, Vedant Koppera, Stefanie Fisher, Tom Zimmerfaust, Chris Severen, Jenna Sterns and Corey White both their intellectual contributions as well as their friendship during my time at UCSB. I would also like to thank the members of the Environmental Economics lunch seminars and the Econometrics Research Group for their feedback on drafts and presentations of this work.

# Curriculum Vitæ
Ryan Michael Abman

## Education

Ph.D., Economics, University of California Santa Barbara, 2015

M.A. Economics, University of California Santa Barbara, 2010

B.A. International Affairs and Economics, University of Colorado, Boulder, 2008

## Research Fields

Environmental and Natural Resource Economics, Econometrics, Political Economy

## Committee Members

Robert Deacon (Chair), Douglas Steigerwald, Paulina Oliva

## Honors and Fellowships

Dean's Central Fellowship, UC Santa Barbara, AY 2013-2014

Travel Grant, UCSB Graduate Division, 2014

Graduate Fellow, Property and the Environment Research Center, 2013

Graduate Dean's Advancement Fellowship, UC Santa Barbara, 2013

Outstanding Teaching Assistant Award, 2013

## Teaching Assistance

Econometric Theory (Ph.D.)

Econometrics (Undergraduate)

Introduction to Microeconomics

Introduction to Macroeconomics

Intermediate Macroeconomic Theory

Introduction to Probability and Statistics

## Conferences and Seminar Presentations

**2015**: CIDE (Mexico City and Aguascalientes), San Diego State University, University of Queensland, and Utah State University

**2014**: CU Environmental and Resource Economics Workshop (Vail, CO), Summer School in Political Economy (GSE, Barcelona), Land Use and Ecosystem Services Workshop (UCSB)

**2013**: Western Economic Association International Conference (Seattle), AERE Annual Meetings (Banff, Canada), PERC Graduate Student Workshop (Bozeman, MT), Occasional Workshop in Environmental and Resource Economics (UCSB)

**Abstract**


Responses to Institutional Constraints

by

Ryan Michael Abman

Institutions, as mechanisms of social order, often constrain the behavior of individuals within a society. Political institutions constrain the behavior of politicians, financial institutions constrain the behavior of businesses and payment processors and social institutions often constrain the behavior of individuals. These institutions often play an important role in constraining activities that may be seen as illicit or unwanted and careful analysis of these constraints can allow researchers to learn more about activities that are often hidden or go unreported.

This dissertation explores the role of institutional constraints on unwanted behavior by studying deforestation in Brazil and Malawi as well as underground activity in fraudulent software sales. These cases share the commonality that they are influenced by institutional constraints. Politicians in Brazil are constrained by reelection incentives, perpetrators of fraudulent antivirus software are constrained by payment processors and the cultural practice of ethnic favoritism in public good provision leads to particular ethnic groups in Malawi receiving much more fertilizer subsidies than others.

The first chapter examines deforestation in Brazil. Local political authority (formal or informal) over natural resources may create rents for politicians. The political decision to use or allocate resources involves balancing private rents with reelection prospects. I examine the case of deforestation in Brazil and a presidential decree granting the federal government the authority to punish counties that failed to limit total deforestation within their borders. This collective punishment aimed to generate pressure on local politicians

to slow deforestation. Using binding term limits as a source of variation in reelection eligibility, I find eligibility has no effect on deforestation prior to the decree. After the decree, reelection eligible mayors reduced annual deforestation 10% more than mayors ineligible for reelection. These findings are consistent with the equilibrium outcome of a lobbying model. Policies such as sanctions, which target the electorate in order to influence political behavior, may be less effective when politicians are not accountable to voters.

The second chapter examines Fake antivirus (AV) programs which have been utilized to defraud millions of computer users into paying as much as one hundred dollars for a phony software license. As a result, fake AV software has evolved into one of the most lucrative criminal operations on the Internet. In this chapter, we examine the operations of three large-scale fake AV businesses, lasting from three months to more than two years. More precisely, we present the results of our analysis on a trove of data obtained from several backend servers that the cybercriminals used to drive their scam operations. Our investigations reveal that these three fake AV businesses had earned a combined revenue of more than $130 million dollars. A particular focus of our analysis is on the financial and economic aspects of the scam, which involves legitimate credit card networks as well as more dubious payment processors. In particular, we present an economic model that demonstrates that fake AV companies are actively monitoring the refunds (chargebacks) that customers demand from their credit card providers. When the number of chargebacks increases in a short interval, the fake AV companies react to customer complaints by granting more refunds. This lowers the rate of chargebacks and ensures that a fake AV company can stay in business for a longer period of time. However, this behavior also leads to unusual patterns in chargebacks, which can potentially be leveraged by vigilant payment processors and credit card companies to identify and ban fraudulent firms. This chapter is joint work with Brett Stone-Gross, Richard Kremmerer, Christopher Kruegel, Douglas Steigerwald, and Giovanni Vigna and was published as Stone-Gross et al. (2013).

The final chapter returns to deforestation and studies it in the context of agriculture in Malawi. The effect of development policies on the environment is often ambiguous ex ante. Programs designed to improve agricultural productivity may increase deforestation by raising the marginal productivity of agricultural land, thus increasing the demand for land clearing. However, in a setting of subsistence farming on unproductive land, increasing agricultural productivity may reduce the need to shift cultivation to maintain the desired yields. This chapter examines the impact of agricultural subsidies on deforestation in Malawi by leveraging ethnic favoritism in government resource allocation. By exploiting a change in the ethnicity of the Malawi president following the 2004 election, we show that coethnic districts received more fertilizer subsidies and experienced significant declines in deforestation compared to districts with other predominant ethnicities. This paper studies a case in which poverty alleviation programs have beneficial environmental impacts demonstrating that, in certain contexts, input subsidies may provide a 'win-win' scenario. This chapter is joint work with Conor Carney.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Reelection Incentives, Blacklisting and Deforestation in Brazil

## 1.1 Introduction

The loss of tropical forests around the world has received increasing attention from researchers, particularly in the context of global climate change. The carbon dioxide released from deforestation is estimated to have comprised nearly 20% of all global greenhouse gas emissions from 2000-2007 (Pan et al., 2011). Deforestation has been most prevalent in tropical forests which contain a disproportionate amount of the world's biodiversity. These external consequences of tropical deforestation have motivated research aimed to better understand the economic, social and political drivers of deforestation. This paper studies the political drivers of deforestation, particularly the tradeoffs faced by local politicians between private rents and reelection prospects.

Local political authority over forest resources may create rents for politicians. Political decisions over forest resource use must consider two factors, private rents and reelection incentives. If these two incentives are aligned, for example, if deforestation is driven by politically popular local economic activity, reelection incentives may have no influence over resource use. However, if increased resource use is at odds with the demands of the electorate, politicians face a trade off in their resource decisions. Those politicians without reelection incentives will choose to deforest more if this action entails more private rents, whereas reelection eligible politicians may choose to deforest less in hopes of winning reelection.

This paper examines a Brazilian presidential decree that allowed the federal government to single-out (blacklist) counties that failed to limit deforestation and impose collective penalties on all landowners. Landowners in blacklisted counties lost eligibility for subsidized credit and faced restrictions on selling beef as well as expensive land titling requirements. This decree created a collective demand to control deforestation even though private demand for deforestation still remained. I argue that this decree exogenously changed the demands of the electorate on local politicians in the Amazon with regard to controlling deforestation and provides a unique opportunity to study how local politicians balance private rent with reelection incentives in their control over natural resources.

The introduction of this decree offers the opportunity to study the incentives faced by local politicians who have de facto influence over deforestation yet risk being blacklisted if they allow too much forest clearing to occur. To guide the empirical analysis, I present a lobbying model whereby a local politician allocates permission to deforest among a large number of landowners in exchange for political contributions. While each individual landowner would privately like to deforest, landowners fear the politician may allow too much deforestation and they would bear the costs of the collective punishment. Landowners observe the total allocation of deforestation by the politician and decide whether to vote for the incumbent or a challenger. The equilibrium of the model suggests that the reelection incentive will lead the politician to choose less deforestation than he would without the possibility of reelection.

Using county-level satellite data on deforestation, I estimate a difference-in-difference model on a panel of 639 counties in the Brazilian Amazon. I utilize binding term limits to compare deforestation in counties with a mayor eligible for reelection to counties with a mayor who is term-limited and thus ineligible for reelection. Prior to the introduction of the policy, I find no difference between deforestation in counties with a mayor eligible for reelection and counties with a term-limited mayor. However, after the introduction of the policy, counties with a reelection eligible mayor reduce annual deforestation by 10% more than counties with a term limited mayor. I show evidence that suggests voters punish incumbent mayors for blacklisting by offering 10 - 11% less vote share and reelecting them at a 16% lower rate in the following election than incumbent mayors in non-blacklisted counties.

Understanding how political motivations influence use of natural resources is empirically difficult for two primary reasons. First, the preferences of the voters over how a resource is used may depend on the politician's ability to extract rent from the resource. Politicians may choose to tie social program spending to resource revenue so as to effectively align the incentives of the electorate with the private incentives of the politician. Second, the degree to which the politician is accountable to the voters may also depend on the availability of resource rents. For example, rents in the form of political campaign contributions may be used to discourage political competition allowing the incumbent politician to be less accountable to the electorate and pursue their own agenda.[1]

This paper overcomes both of these obstacles. First, the policy that allowed the federal government to collectively punish an entire jurisdiction was passed by presidential decree and, as such, required no support from local politicians. The policy exogenously changed the demands of the electorate regarding how politicians controlled deforestation.

[1]See for example Mahdavi (2015); Brollo et al. (2013).

Second, local mayors in Brazil may serve no more than two consecutive terms. This electoral rule offers well-defined variation in political accountability as second term mayors may not run for reelection. This paper provides an example in which, in the absence of electoral accountability, local political authority over natural resources leads to patterns of resource use that are at odds with the demands of the electorate.

This paper contributes to a growing literature on the political economy of deforestation as well as the role of electoral accountability and political behavior in natural resource use. The early literature on the political economy of deforestation documents the importance of a government's ability and willingness to enforce property rights in explaining country-level deforestation (Deacon, 1994; Bohn and Deacon, 2000). Follow-up work moved from cross country analysis to within country variation in governance to study deforestation(Ferreira and Vincent, 2010; Wendland et al., 2014). Other papers have been more explicit in modeling the interaction between landowners and politicians. Based on Grossman and Helpman (1994), these papers argue that a significant amount of deforestation can be explained through the channel of corruption, whereby local economic interests lobby or bribe the government to allow more forest clearing activity (Barbier et al., 2005) or to receive agricultural subsidies (Bulte et al., 2007) and put more land into extensive agriculture. This paper contributes to this work by explicitly examining the role of electoral accountability as well as utilizing within-country data to mitigate concerns that arise from unobserved heterogeneity in national institutions with respect to forest management.

Recent work has explicitly examined local political incentives and deforestation within a given country. Burgess et al. (2012) examine the effect of a proliferation of local jurisdictional units on deforestation in Indonesia finding that this led to a significant increase in deforestation. Cisneros et al. (2013) examine corruption among local politicians and deforestation in the Brazilian Amazon. They find that mayors caught stealing public funds and engaging in other corrupt behavior allow more deforestation. Morjaria (2014) finds the transition to democracy in Kenya led to significantly higher deforestation in politically contested districts as the president was able to allocate permits in exchange for votes. While Burgess et al. (2012) and Cisneros et al. (2013) study the private rents aspect of political decisions over deforestation and Morjaria (2014) studies the reelection aspect of political decisions over deforestation, this paper examines both aspects, particularly the trade off between the two incentives.

The paper proceeds as follows. Section 2 discusses the blacklisting policy as well as local politics and political influence over forest resources in Brazil. Section 3 develops a political agency model of deforestation under the threat of blacklisting. Section 4 presents the data and empirical methodology and Section 5 discusses the empirical findings. The paper concludes with some final remarks.

## 1.2 Regulation, deforestation and local politics in Brazil

Brazil has overseen an unprecedented decline in deforestation over the past decade. From its peak of approximately 28 thousand square kilometers deforested in 2004, annual deforestation fell to a mere 4,571 square kilometers in 2012.(INPE, 2013) This dramatic decline has been attributed to a combination of falling commodity prices and new policing strategies by the federal government. (Assunção et al., 2012; Hargrave and Kis-Katos, 2013; Nepstad et al., 2014) Despite strict laws governing deforestation on private land, the enforcement of such laws was largely ineffective until 2004. The federal environmental agency adopted the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm), a command and control policy aimed at curbing deforestation. This plan dramatically increased the area of forest resources set aside for protection and introduced real time satellite monitoring of deforestation (which began in 2004) to detect illegal deforestation activity. The improved monitoring led lead to more raids, arrests, confiscation of cattle and farm equipment as well as fines. Upon detecting deforestation via satellite, the Federal enforcement agency in charge of deforestation monitoring (IBAMA) would send out local inspectors to the site. This greatly increased the capacity for enforcement, led to a dramatic increase in fines for deforestation. Despite an extremely low fine payment rate, research has found that these inspections were important in decreasing the rate of deforestation. (Assunção et al., 2012, 2013; Hargrave and Kis-Katos, 2013)[2]

While the real-time satellite monitoring system greatly improved the federal government's ability to target monitoring efforts, enforcement required catching illegal deforesters in the act of clearing land or transporting timber. Agents from the federal government were often met with local protests, threatened by armed *grileiros* or arrived to find timber mills already abandoned as locals had received advanced notice of their inspection. The federal government wanted a mechanism to control deforestation that did not entirely rely on the ability of inspectors to find actors engaging in deforestation.

### 1.2.1 The Blacklisting Policy

In 2008 the command and control legislation was strengthened by Presidential decree 6,321 which created the legal basis for blacklisting counties with intense deforestation activity. There are three principle criteria for singling out particular municipalities; total deforested area, total deforested area over the past three years, and an increase in the deforestation rate in at least three of the past five years. A blacklisted county is subject to

[2] Notable economics literature on deforestation in the Brazilian Amazon has examined the economic causes of deforestation (Pfaff, 1999; Pfaff et al., 2007) as well as the effects of conservation policies such as protected area establishment (Barber et al., 2012; Pfaff et al., 2013; Nolte et al., 2013).

more stringent monitoring and environmental law enforcement as well as potential land title revisions by the National Institute for Colonization and Agrarian Reform (INCRA). Furthermore, to qualify for subsidized credit, landowners in blacklisted counties must prove that their land has been registered through a georeferenced land registry and that they have no outstanding fines from IBAMA. This increases the cost for all landowners within the municipality of titling land and complying with federal laws. Furthermore, some international soy and beef suppliers will not purchase from municipalities under priority status thus closing market options for goods produced within the municipality. In 2009, federal prosecutors began to take legal action against beef suppliers who had purchased cattle raised in blacklisted counties.

An article in the Economist Magazine on September 14, 2013 reporting on one blacklisted municipality, Paragominas, states,

> "Being blacklisted did not just bring public humiliation to the citizens of Paragominas, it also hit their wallets. Businesses in municipalities on the list were not eligible for cheap credit from state owned banks. ... The federal public prosecutor in Para, Daniel Avelino, followed the supply chain back from the supermarkets through the beef companies to the ranchers to find out which animals had been produced on illegally deforested land, and threatened the super markets with prosecution."

Such measures were designed to make local politicians accountable for high rates of deforestation and previous research has argued that blacklisting can be politically costly.(Assunção et al., 2013; Hargrave and Kis-Katos, 2013) Blacklisting imposed costs on everyone in the county, not just those engaging in deforestation. By threatening a public bad for excessive deforestation, the federal government wanted to force local mayors to solve the collective action problem of limiting deforestation.

Upon passing the policy, 36 municipalities were immediately blacklisted in 2008. Seven were added in March of 2009 and seven more were added in March of 2010. Figure 1.1 provides a map of these counties. Empirical results reported later indicate that deforestation in blacklisted municipalities dropped by 25% under blacklisted status.[3] Removal from the blacklist required that the county formally register 80% of its territory (with the exception of approved protected areas and indigenous lands) with the INCRA as well as maintain annual deforestation below a limit set by the Ministry of the Environment. As of 2013, only two counties had left the blacklist.

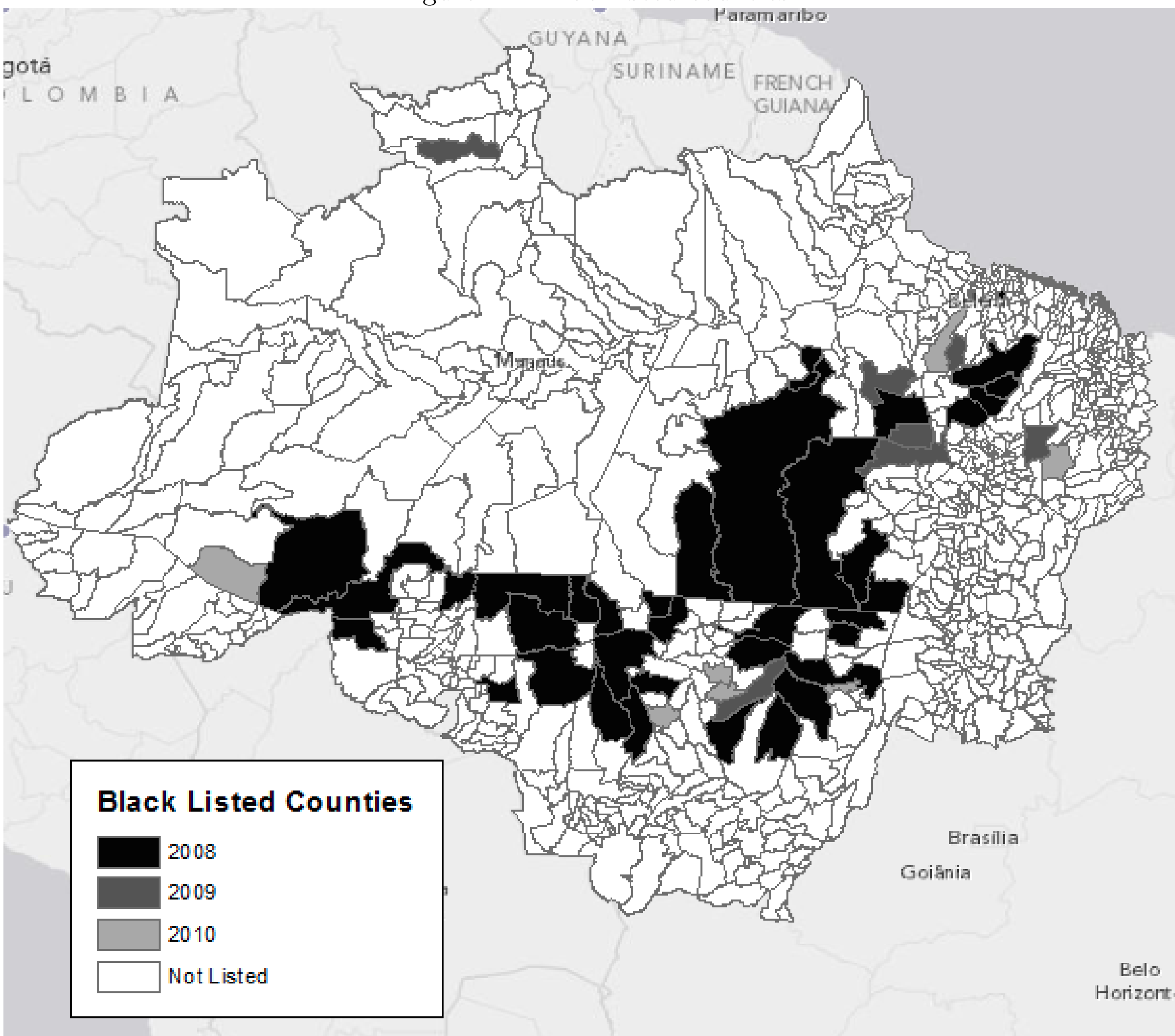
For the purposes of this study, there are two important characteristics of blacklisting which deserve attention. First, as discussed above, blacklisting shifts the cost of deforestation from the individual agent engaging in deforestation (landowner, sawmill operator, rancher, etc.) to the entire county. Second, the blacklisting policy was passed by presidential decree, which did not require the consent or votes of local politicians. For this reason, the blacklisting decree can be seen as exogenous from the perspective of voters and local politicians. I argue that the introduction of the policy exogenously changed the demands of the voters on local politicians with respect to controlling deforestation. Voters who would have previously been indifferent to the amount of deforestation in their county risk bearing costs associated with the collective punishment from blacklisting if local deforestation is not controlled.

### 1.2.2 Mayoral Control of Forest Resources

While formal authority over land use largely lies with the Federal government, local politicians exert substantial influence over deforestation. In 2008, the Brazilian Environmental minister, Carlos Minc, argued the observed increase in deforestation was linked to local elections, claiming that mayors were ignoring illegal logging activity in order to

[3] See section 5 for details

Figure 1.1: Blacklisted counties




win votes (Balakrishnan, 2008). Cisneros et al. (2013) explicitly examine local corruption and deforestation among mayors in the Amazon from 2002-2009. Using data from randomized corruption audits focusing on management of public funds, the authors find that average corruption is correlated with higher deforestation. Furthermore, they find mayors increase deforestation in response to an audit finding corrupt behavior. They attribute this finding to mayors shifting corrupt behavior from public funds toward unmonitored corrupt behavior, enabling illegal deforestation.

Mayors have informal channels by which to influence deforestation. Some mayors have engaged in fraud to legalize illegal timber by falsifying the necessary paperwork.[4] Landowners are only allowed to extract 15 cubic meters of timber per hectare in a given year and all timber extraction from protected areas is illegal. Falsified documents may hide the source of illegally extracted timber.

Mayors may work with local squatters or large landowners to coordinate squatter settlements in protected areas or on privately held land. Landowners with squatters on their land are entitled to compensation from the federal government which is often above the market price for land. Mayors may also tolerate or even facilitate the illicit selling of untitled land which effectively removes forest cover restrictions faced by the titled land. Mayors have also been involved in diverting equipment intended for public use and small family farms to large private properties, reducing the cost to the landowner of clearing their land.[5] Mayors have also fought to block the establishment of protected areas in their jurisdictions. While the creation of protected areas is often done at the Federal or State level, they require consultation of local mayors before their establishment.

### 1.2.3 Local Politics in Brazil

Local mayors play an important role the provision of public services such as health and education and in settling local disputes. Mayors are elected by popular vote in each county and serve terms of four years. Local elections are held on separate four-year cycles from state and federal elections and take place in October of the election year with the winner taking office in January of the following year. Since 1997, mayors are allowed to

[4] For example, the mayor of Nova Mamore was accused by the Public Ministry of the State of Rondonia of falsifying documentation pertaining to timber that had been extracted illegally between 2007 and 2008 to benefit a timber company with which he had connections.(Rondoniagora, 2014)

[5] An example of one such case occurred in the county of Davinopolis, whereby the mayor was able to procure machinery dedicated to small, family farms through the PRONAF program and use it to clear land for a large landowner who would not have qualified for the assistance otherwise.

serve at most two consecutive terms and may only run for office again after a one-term hiatus.[6]

Few candidates eligible for a second term actually serve one. Table 1.1 summarizes election results for the 760 counties in the legal Amazon. 70% of incumbents ran for a second term yet less than 40% of eligible candidates were reelected from 2000 to 2012.[7] The absence of a clear incumbency advantage implies that incumbent politicians have the incentive to work hard and cater to the interests of the voters in their first term in order to win reelection. This reelection incentive has been shown to motivate reelection eligible politicians to steal fewer resources (27% less) than term limited mayors (Ferraz and Finan, 2011) and to allocate conditional cash transfers more effectively than term limited mayors (De Janvry et al., 2012). Reelection eligible mayors who do not perform well are punished by voters (Ferraz and Finan, 2008; De Janvry et al., 2012).

The constitution of Brazil makes voting mandatory for all citizens ages 18 to 65. The introduction of electronic voting in Brazil created a large de facto enfranchisement of less educated voters and allowed for near-perfect records of voting to be measured even in remote locations (Fujiwara, 2010). Municipal elections occur every four years and do so on a separate cycle from state and federal elections. The chief executives of the three levels of government (federal, state, municipal) are chosen by majority vote with municipalities under 200,000 in population electing the candidate with a plurality and those with greater populations carrying out a run-off election to ensure a majority vote. Elections are held in October and the winning candidate takes office at the beginning of January of the following year.

[6] Prior to 1997, mayors could not serve a second consecutive term.

[7] This accords with Ferraz and Finan (2011) who study all Brazilian counties.

## 1.3 A Political Agency Model of Brazilian Deforestation

To understand the role of reelection incentives and blacklisting in the Brazilian Amazon, I consider a context in which a county is populated by a large number of landowners who produce an identical agricultural good with deforested land. An individual landowner is only able to clear additional land with permission from the mayor. The mayor, in turn, receives campaign contributions or bribes from the landowners and makes allocation decisions accordingly. A collective penalty is imposed on all landowners if the aggregate amount of deforestation exceeds a known threshold. This feature captures the blacklisting policy. Landowners (who are also voters) rely on the mayor to limit the total amount of deforestation allocated in order to avoid the collective punishment. Landowners are able to constrain the behavior of the mayor in the first term via the reelection mechanism, but are unable to do so in the mayor's final term.

I model this using a two-period political agency model based on Besley (2006) where payoffs are determined by a menu auction game based on Grossman and Helpman (1994). I assume that a mayor can be one of two types - corrupt or benevolent - based on whether the mayor's preferences include the welfare of the landowners. I assume the mayor type is private information held by the mayor. In each period, landowners simultaneously and competitively submit a contribution schedule to the mayor, which maps any given amount of permissible deforestation into a contribution. Upon receiving the schedules, the mayor allocates permission to deforest among landowners and collects the appropriate payments. If the total amount of deforestation exceeds the pre-determined blacklisting threshold, all landowners incur a penalty which is independent of the amount of deforestation the individual landowner was allocated in that period. The preferences of a benevolent mayor include the welfare of the landowners. Consequently, a benevolent mayor will limit the amount of deforestation so as to avoid the collective punishment. A corrupt mayor, however, prefers to allow more deforestation in order to collect more contributions and has no direct preferences regarding the punishment of the landowners.

After the first period, landowners vote to decide whether to reelect the incumbent mayor or elect a challenger whose type is unknown. The reelection decision allows voters to cast out a mayor who allocates too much land for clearing in the first period, knowing that the mayor's type must be corrupt and that the same mayor will continue to allocate too much land in the second period, inducing another collective punishment. The reelection decision benefits landowners in two ways; it disciplines the actions of corrupt mayors in the first period, and allows voters to improve the probability of having a benevolent mayor in office for the second period.

In what follows, I describe the players of the game, provide the timing of the game and describe the equilibrium.

### 1.3.1 Landowners

Consider a county with $N$ identical landowners who use their land to produce an identical agricultural product. They produce this product using newly cleared land in that period, $h_{i,t}$ and sell the good at market price, $P_t$ which they all take as given. The production function is increasing and concave in its argument and each landowner faces a shock to total factor productivity in each period, $\alpha_{i,t}$ which is drawn independently and identically from a stationary distribution. The draw of total factor productivity is revealed to the landowner at the beginning of each period.

In this model, it is assumed that landowners may only clear land if they receive permission from the mayor. This assumption captures the fact that, while landowners themselves can clear land without explicit permission, mayoral support can make the land clearing cheaper or, in the case of falsified documents, make illegal deforestation appear legal. The assumption that land clearing may only take place with the support of the government is also made in previous work such as Barbier et al. (2005) and Burgess et al. (2012).

In order to gain this support, each landowner submits a contribution schedule to the mayor which maps any given level of permitted deforestation into a lump-sum contribution. These payments are contingent upon the landowner receiving permission to deforest a particular amount of land. This paper focuses on truthful bidding strategies for reasons discussed below. Such schedules allow the mayor to observe the willingness to pay for deforestation for all landowners for all possible quantities of deforestation in a particular period. The willingness to pay for deforestation varies across landowners and across periods by the realization of $\alpha_{i,t}$, an independent and identically drawn total factor productivity shock. While all landowners are *ex ante* identical, after the realization of the shock, landowners with high draws will be willing to pay more for deforestation than those with low draws. The payment schedule for landowner $i$ in period $t$ is denoted $s_{i,t}(h_{i,t})$. The contribution schedule is assumed to be non-negative for all values of $h_{i,t}$.[8] The other potential cost faced by the landowner is the private cost incurred from blacklisting. Blacklisting occurs if the total sum of deforestation in the county exceeds a particular threshold $\overline{H}$ and costs each landowner the same amount, $\Delta$.

The utility function of landowner $i$ is linear in profits and is as follows:

$$U_{landowner_{i,t}} = P_t \alpha_{i,t} q(h_{i,t}) - s_{i,t}(h_{i,t}) - \Delta \mathbb{1}\{\sum_{j=1}^{N} h_{j,t} > \overline{H}\} \tag{1.1}$$

where $\mathbb{1}\{\sum_{j=1}^{N} h_{j,t} > \overline{H}\}$ is an indicator function equal to 1 if the sum of allocated deforestation exceeds the threshold, and zero otherwise. Note the collective penalty from

[8]This condition simply states that landowners cannot charge the mayor for permission to deforest, which accords with intuition.

blacklisting affects all landowners identically, regardless of their level of deforestation in that period.

Landowners in the model have choices over two dimensions. First, in each period, after the productivity parameter is revealed, they submit contribution schedules to the mayor in order to compete for permission to deforest. Second, at the end of the first period, landowners have the choice to vote in favor of the incumbent mayor or vote for a challenger. The equilibrium strategies for both choices are discussed below.

### 1.3.2 Mayors

In this model, I assume there are two types of mayors, corrupt and benevolent. The type is denoted by $\theta \in \{0, 1\}$ with a $\theta = 0$ corresponding to the corrupt mayor and $\theta = 1$ corresponding to the benevolent mayor. The mayor type is privately revealed to the mayor, but the distribution of types is common knowledge, with $\Pr(\theta = 1) = \pi$. The forgone utility from serving as mayor is normalized to zero, so utility gained from actions in office are net of potential outside income or other sources of utility. Formally the utility function of the mayor in period $t$ is as follows:

$$U_{mayor_t} = \theta \sum_{i=1}^{N} u_{i,t}(h_{i,t}) + \sum_{i=1}^{N} s_{i,t}(h_{i,t}) \tag{1.2}$$

$$S.T. \sum_{i=1}^{N} h_{i,t} \leq H^{MAX}$$

The job of the mayor is allocate deforestation among the landowners. This decision is two-fold, first the mayor must choose how much total deforestation to allow, and then the mayor must choose how to allocate the total amount of deforestation among the landowners within the county. In each period, mayors observe the submitted contribution schedules of the landowners and decide how much to allocate to each landowner. The mayor is constrained to the total amount of deforestation he may allocate in any given period to $H^{MAX}$.[9]

### 1.3.3 Equilibrium of Menu Auction Game

In each period, landowners bid for permission to clear land while the mayor chooses the total amount of land to allocate and, conditional on that level of total land, how to allocate land among the landowners. The menu auction in this model is a simple application of Bernheim and Whinston (1986) and I draw from their results to establish the equilibrium of the game. For this section, I drop the time subscript, $t$, and focus simply on the strategies in a given period.

Let $s_i^o(h_i) \in S_i(h_i)$ denote the contribution schedule submitted by the landowner and $\boldsymbol{h}^o$ denote the $1 \times N$ vector of allocations made by the mayor. For any given level of $H$ chosen by the mayor, $(\{s_i^o\}_{i \in N}, \boldsymbol{h}^o)$ is a subgame-perfect Nash equilibrium if and only if:

1. $s_i^o(h_i)$ is feasible for all $i \in N$
2. $\boldsymbol{h}^o$ maximizes $\theta \sum_{i=1}^N u_i(h_i) + \sum_{i=1}^N s_i^o(h_i)$ such that $\sum_{i=1}^N h_i \leq H$
3. $\boldsymbol{h}^o$ maximizes $u_i(h_i) + \theta \sum_{i=1}^N u_i(h_i) + \sum_{i=1}^N s_i^o(h_i)$ for every $i \in N$ such that $\sum_{i=1}^N h_i \leq H$.

While these conditions have been extensively discussed in the political economy literature, I briefly describe their intuition.[10] Condition (1) states that all contribution offers must be feasible. This condition proves trivial with the nonnegative assumption imposed above as well as the fact that offering to pay more than the marginal benefit

[9]This constraint restricts the mayor's influence over deforestation to be finite. From the standpoint of the model, what is needed is that the resource the mayor allocates is, in some way, scarce. The main findings of the model hold if, rather than imposing a limit on the total amount of allocation, the mayor faces a marginal cost of allocating deforestation.

[10]See Bernheim and Whinston (1986) as well as Grossman and Helpman (1994).

for deforestation is not optimal for the landowner. Condition (2) states that the allocation of deforestation permission must maximize the utility of the mayor. Condition (3) states that the contribution schedule and allocation of deforestation must maximize the joint utility of the the mayor and any landowner. This condition must hold for a Nash equilibrium to exist. If it did not hold, landowners could alter their strategy and capture more surplus.

Among the ways landowners could choose $s_i(h_i)$, I focus on truthful contribution schedules, whereby the landowner offers his full net-willingness-to-pay for any given level of deforestation. Because permission to deforest is rivalrous, each landowner is competing against the offers of other landowners for permission to clear land. Consider a simple case of two landowners/bidders $(A, B)$ for one unit of deforestation. If $\alpha_A > \alpha_B$, landowner $A$ will offer the politician $P\alpha_B[q(1)-q(0)]+\epsilon$, the full-willingness-to-pay of landowner $B$ plus some arbitrarily small amount, $\epsilon$, to ensure the politician strictly prefers to offer the unit of deforestation to $A$. $B$ will offer his full-willingness-to-pay for that unit of deforestation, $P\alpha_B[q(1) - q(0)]$, and will be indifferent between winning the auction and losing. Now consider a third landowner/bidder $(C)$ with $\alpha_A > \alpha_C > \alpha_B$. In this very same auction for one unit of deforestation, $B$ and $C$ offer their full-willingness-to-pay. While $A$ will still win, he must offer a greater amount to the mayor, $P\alpha_C[q(1) - q(0)] + \epsilon$. As the number of landowners becomes large, the difference between $\alpha_A$ and $MAX\{\alpha_B, \alpha_C, \ldots, \alpha_N\}$ will become arbitrarily small and the optimal bid for $A$ will approach his full-willingness-to-pay for deforestation.[11] Truthful bidding will ensure that deforestation is allocated efficiently among landowners (landowners who value deforestation the most are the ones who receive permission to do so) but truthful bidding will also allow the mayor to capture

[11] Truthful bidding is used in Grossman and Helpman (1994),Barbier et al. (2005),and Bulte et al. (2007) and others. Theorems (1)-(3) of Bernheim and Whinston (1986) prove that, even in menu auctions with small numbers of bidders, truthful bidding strategies 1) always exist in a bidder's best-response correspondence, 2) lead to an efficient allocation choice, and 3) are coalition-proof in the presence of nonbinding communication among bidders.

the full surplus from the auction.[12]

The truthful bidding strategy by the landowner, denoted $s_i^*(h_i)$, implies that in a given period, landowner $i$ will offer the following contribution schedule for any level of deforestation permitted by the mayor.

$$s_i^*(h_i) = P\alpha_i[q(h_i) - q(0)] \tag{1.3}$$

The truthful bidding strategy implies that the landowner does not account for the likelihood of blacklisting in the contribution schedule. The landowner effectively takes the mayor's choice of $H$ as given (although it is not observed) for the menu auction and competes for permission to deforest in exactly the same way if $H > \overline{H}$ or $H < \overline{H}$. This stems from the fact that the penalty incurred by the landowner is additively separable from the amount of private deforestation undertaken and from the fact that the individual landowner cannot, unilaterally, alter the total amount of deforestation chosen by the politician by reducing his contribution schedule.

Upon receiving the contribution schedules, the mayor must then decide how much land to allocate and to whom this permission should be granted. I begin by addressing the question of how a mayor allocates the permission to clear land among the landowners and then address the decision on the total amount of land to clear. By Theorem 2 of Bernheim and Whinston (1986), the Truthful Nash Equilibrium will yield an efficient action choice regardless of the type of mayor. However, the difference in preferences of the two mayor types may potentially lead to different, efficient allocations. This motivates the following proposition.

**Proposition 1:** *Given a set of truthful contribution schedules* ($\{s_i^*\}_{i=1}^N$) *and a fixed*

[12]This case is analogous to the example of all voters as special interest groups investigated by Grossman and Helpman (1994). "When all voters are active in the process of being influence, the rivalry among competing interests is most intense, and the government captures all of the surplus from political relationships" (page 846).

*level of total deforestation $H$, the optimal allocation of deforestation permission, $\boldsymbol{h}^*$, does not depend on mayor type, $\theta$.* For proof, see appendix A.0.1. The intuition for the proposition is simple. Regardless of the type of mayor, mayors prefer to allocate some total amount of deforestation to the landowners who value it the most because, under truthful bidding, this will maximize the total contributions received.

Proposition 1 reduces of the mayor's choice to the total level of deforestation permitted. The corrupt mayor only has preferences for maximizing contributions with no regard to the costs of the blacklisting on the landowners in the county. Let $S^*(H)$ be the maximized total contributions from the efficient allocation of $H$ total deforestation. Truthful bidding implies that landowners bid their full willingness to pay for deforestation regardless of the level of $H$ chosen by the mayor. Therefore, $S^*(H)$ is strictly increasing in $H$ and the corrupt mayor prefers the highest level of $H$ possible.

The preferences of the benevolent mayor lead to a different outcome. While increasing $H$ beyond $\overline{H}$ increases contributions, it also causes the electorate to incur the collective punishment which enters into the preferences of the benevolent mayor. As long as $S^*(H^{MAX}) - S^*(\overline{H}) < N\Delta$, the benevolent mayor will choose to limit total deforestation to $\overline{H}$. It is this difference that provides the basis for the signaling game.

### 1.3.4 Political Agency and Electoral Accountability Equilibrium

The menu auction provides the per-period payoffs of the game, but the notion of political agency component lies in the signaling model discussed here. The timing of the game is as follows:

1. Nature reveals to the incumbent mayor his type, $\theta$.

2. Landowners observe first period total factor productivity realizations ($\alpha_{i,1}$) and

submit their contribution schedules $(s_{i,1}(h_{i,1}))$ accordingly.

3. Mayors observe the contribution schedules and choose the total amount of deforestation permission to allocate, $H_1$.[13] Landowners pay mayors and clear land. First period payoffs are realized.

4. Landowners decide whether to vote for the incumbent mayor or for a challenger whose type is drawn at random from the distribution of types. The mayor, an incumbent or challenger, is elected.

5. Landowners observe second period total factor productivity realizations $(\alpha_{i,2})$ and submit their contribution schedules $(s_{i,2}(h_{i,2}))$ accordingly.

6. Mayors observe the contribution schedules and allocate permission to clear land $(H_2)$. Landowners pay mayors and clear land. Second period payoffs are realized and the game ends.

The equilibrium concept for this game is a perfect Bayesian Nash equilibrium, whereby the mayor behaves optimally in each stage of the game given the reelection rule of the voters and the voters update their beliefs regarding the mayor's type using Bayes' rule upon observing first-period actions. The game is solved by backwards induction. The second period actions by the mayor depend only on type. Benevolent mayors will choose $H_2 = \overline{H}$ to avoid blacklisting so long as $S^*(H^{MAX}) - S^*(\overline{H}) < N\Delta$. Corrupt politicians will choose $H_2 = H^{MAX}$ because the maximized sum of contributions is strictly increasing in $H$.

Under the truthful bidding strategy, the voters all prefer to have a benevolent mayor in office for period 2 as the collective punishment would make all landowners worse off.

[13] By Proposition 1, the amount allocated to the individual landowners will simply be the contribution maximizing allocation for the chosen total allocated land, $H_1$.

The beliefs about the mayor's type will be informed from the payoffs received in period one. A benevolent mayor sets $H_1 = \overline{H}$. The corrupt mayor faces a trade-off. He can set $H_1 = H^{MAX}$ and collect $S_1^*(H^{MAX})$ in contributions, but, in doing so he reveals his type as corrupt to the voters and will be voted out of office, receiving zero payoffs in period 2. The corrupt mayor may also choose $H_1 = \overline{H}$, forgo the additional $S_1^*(H^{MAX}) - S_1^*(\overline{H})$ in contributions, but improve his chances of being reelected and collecting $S_2^*(H^{MAX})$ in contributions in period 2. Allow $\lambda$ to be the probability a corrupt mayor will choose $H_1 = \overline{H}$.

Upon facing no penalty in the first period, the landowners beliefs about the probability the incumbent mayor is benevolent are:

$$\Pr(\theta = 1|No\ penalty) = \frac{\pi}{\pi + (1-\pi)\lambda} \geq \pi \tag{1.4}$$

Equation (1.4) implies that, having incurred no penalty in period one, the probability the mayor in office is benevolent is greater than the probability a mayor selected at random will be benevolent. Voters would thus prefer to reelect the incumbent rather than electing a challenger to serve as mayor in period 2. If voters incur a penalty in period 1, they know that the incumbent mayor is corrupt ($\Pr(\theta = 1|\Delta) = 0 \leq \pi$) and would prefer to elect the challenger. Thus, the optimal strategy of the voters is to vote for the incumbent when he sets $H_1 = \overline{H}$ and to vote for a challenger when the incumbent sets $H_1 > \overline{H}$.

Given the reelection rule set forth by the voters, the deforestation choice of the corrupt mayor for $H_1$ is now clear. If the set of contribution schedules offered to a corrupt first-term mayor is sufficiently high, the mayor will set $H_1 = \overline{H}$ and forgo the discounted expected second-period contributions, $\beta\mathbb{E}[S_2^*(H^{MAX})]$, where $\beta$ is the discount rate. However, if the first period returns to exceeding the threshold are smaller than the discounted expected second period contributions, the corrupt mayor will choose to limit first period deforestation to win reelection. Recall that $\lambda$ is probability a corrupt politician limits deforestation to the mandated threshold. It can formally be expressed as

$$\lambda = \Pr\left(S_1^*(H^{MAX}) - S_1^*(\overline{H}) < \beta\mathbb{E}\Big[S_2^*(H^{MAX})\Big]\right). \tag{1.5}$$

While the mayor is able to capture the surplus in the menu auction game, the electoral accountability mechanism benefits the voters in two ways. First, it incentivizes a corrupt mayor to take the action preferred by the voters with probability $\lambda$ in the first period. This is known as the disciplining effect of electoral accountability. Second, it increases the probability that voters have a benevolent politician in office for period 2 by allowing them to remove corrupt politicians with probability $(1-\lambda)$. This is referred to as the selection effect.

The implications of the equilibrium of the model on deforestation by period depend on the parameters of the model. The expected values of total deforestation in each period are as follows.

$$\mathbb{E}[H_1] = \overline{H}(\pi + (1-\pi)\lambda) + H^{MAX}(1-\lambda)(1-\pi) \tag{1.6}$$

$$\mathbb{E}[H_2] = \overline{H}(\pi + (1-\lambda)(1-\pi)\pi) + H^{MAX}((1-\pi)\lambda + (1-\lambda)(1-\pi)^2) \tag{1.7}$$

In the first period, both benevolent mayors as well as disciplined corrupt mayors will limit deforestation to the threshold level. Only undisciplined corrupt mayors will allow deforestation to exceed the threshold amount. In the second period, benevolent mayors (reelected incumbents or benevolent challengers replacing undisciplined corrupt mayors from the first period) limit deforestation to the threshold. All corrupt mayors (those that hid their type to get reelected and corrupt challengers replacing undisciplined corrupt mayors from the first period) will allow as much deforestation as possible. Under the blacklisting regime, expected deforestation in the first period will be lower than expected deforestation in the second period if a sufficiently large share of corrupt politicians are disciplined by the reelection incentive, or precisely $\pi < \frac{\lambda}{1-\lambda}$.

### 1.3.5 Preblacklist Equilibrium

In order to contrast the equilibrium described above with that of the equilibrium prior to the blacklisting regime, allow $\Delta$ to equal zero. This removes the collective punishment and changes the incentives for voters and politicians alike. Voters now prefer the mayor to set $H_1 = H_2 = H^{MAX}$ and both types of mayors do exactly that. Voters gain no information from first period actions, but are indifferent between the two types of mayors for the second period, so the decision of whether or not to reelect becomes orthogonal to mayor type. In contrast to the model with blacklisting there is no expected difference in deforestation between period one and period two. By allowing period one and period two to be analogous to the first term and the second term (term limit) of a mayor, the differences in the two variants of the model generate hypotheses to test in the data.

### 1.3.6 Overlapping regimes in 2009-2012

The application of the model to this particular empirical context requires one caveat. Because deforestation is only observed until 2012, only one electoral cycle is observed under the blacklisting regime. The blacklisting policy was enacted in an election year which implies the deforestation patterns observed may not reflect the full equilibrium of the model. The cohort eligible for reelection in 2008 served a first term under the pre blacklist regime, which implies there was no selection effect from electoral accountability. Without selection, the distribution of mayor types in the second period is identical to that of the first period. This implies that the expected deforestation for second term mayors in this regime is equal to

$$\mathbb{E}[H_2] = \overline{H}(\pi) + H^{MAX}(1 - \pi). \tag{1.8}$$

However, the cohort of first term mayors face the reelection incentives under the blacklisting regime, implying the expected level of total deforestation among these first term mayors is equal to that in (1.6). Because no selection occurs until the election in 2012, the difference in deforestation will be entirely driven by the disciplining effect. This implies that deforestation will necessarily be greater among the second-term cohort than the cohort eligible for reelection for 2009 to 2012 because some of the corrupt first term mayors will try to mask their type by deforesting less.[14]

## 1.4 Empirical Analysis

The model discussed above generates two empirically-testable hypotheses with regards to deforestation and one with regard to reelection. First, deforestation prior to 2008 should be the same in counties with reelection-eligible mayors as in counties with term-limited mayors. Second, after 2008, deforestation in counties with term-limited mayors should be higher than deforestation in counties with reelection eligible mayors. Using variation in binding term-limits, I employ a difference in differences strategy to test these two hypotheses. Lastly, incumbent mayors in blacklisted counties should be

[14] This can be seen by a simple comparison of equations (1.6) and (1.8).

voted out of office. To test this hypothesis, I estimate an incumbent mayor's vote share in 2012 as well as whether the mayor is reelected in 2012 as a function of the blacklist status of his/her county.

The data for this paper are comprised of set of annual observations from 2002 to 2012 on 760 Brazilian Municipalities in the legal amazon (see Figure 1). Data for this paper are publicly available and come from three different Brazilian agencies. The primary measure of deforestation comes from the PRODES project of the Brazilian Space Research Agency (INPE, 2013), which provides annual estimates of deforestation by county. These estimates are produced from analyzing images of forest cover from Landsat images. The fine scale of the data comes at the cost of infrequent observations. Each year, the best, cloud-free images are chosen from July to September (the time window with the least amount of cloud cover in the amazon rainforest) and are overlaid with a map of existing, intact forest cover. This map is then updated by areas where forest loss was detected, providing the new map of existing forest cover for comparison the following year. Hansen et al. (2008) identify potential short comings in PRODES deforestation estimates using MODIS satellite data. Once an area is deforested, it is permanently removed from the study area in PRODES which makes the data incapable of examining changes in secondary forest cover.[15] Second, the cerrado and cerradao regions of the legal amazon also experience forest loss, but are not studied by PRODES. While the policies studied in this paper apply to the amazon regions explicitly, PRODES data cannot be used to test whether the effects exist outside of their intended area. (Alternative measures of deforestation are considered for this purpose in section 6). The focus on amazon forest of PRODES data limits the initial sample to 639 counties.

Reelection eligibility, incumbent status and vote share are determined by matching

[15] Secondary forest cover is tree growth in once deforested area. For example, if pasture land is abandoned, and trees grow back inside the cleared area, this would be secondary forest cover. Any new clearing of this forest would not be identified by PRODES.

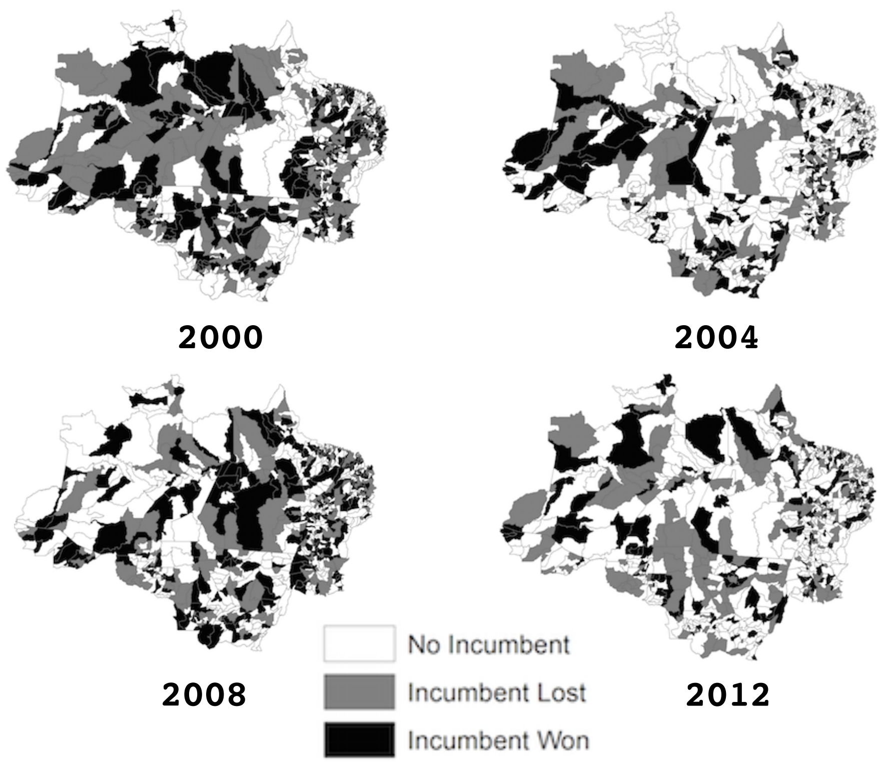
the names of previous winners to current candidates in each of the elections from 1996 (when all incumbent mayors were eligible for reelection) through 2012 using data from from the Brazilian Supreme Electoral Court (TSE). Table 1.1 provides the summary statistics for incumbent politicians and reelection to a second term and Figure 1.2 maps the counties in which incumbents win and where incumbents lose in the 2000, 2004, 2008 and 2012 elections.

Table 1.1: Election Summary

| Year | 2000 | 2004 | 2008 | 2012 |
|---|---|---|---|---|
| Incumbents Running | 505 | 328 | 423 | 362 |
| Incumbent Victories | 276 | 181 | 255 | 164 |
| Total Elections | 760 | 760 | 760 | 760 |

I include controls for economic factors that have been shown to drive deforestation. Controls for each municipality come from the Municipal Agricultural Survey by the Brazilian Institute of Geography and Statistics (IBGE) and include survey data on cattle, harvested acres dedicated to soybeans, total value of production from soybeans, official timber production quantity and value. From this data, I calculate nominal farm-gate prices for soybeans at the municipality level (following Hargrave and Kis-Katos (2013). I use commodity price data from the World Bank to capture time variation in beef prices. As this is the international price, this measure does not vary across municipalities, only across time. All prices are converted to 1998 Brazilian Reals. I interact both the price data with the corresponding lagged quantity data to allow changes in beef prices, for example, to have differential impacts on areas with more cattle from those with no cattle. Finally, to control for protected areas, I use data form the World Database of Protected Areas. I calculate the fraction of land inside any form of protected area for each municipality in each year. Table 1.2 summarizes the data used in this paper.

Figure 1.2: Elections in Amazon counties




### 1.4.1 Deforestation and reelection eligibility

I utilize two measures of deforestation for the empirical analysis. First, I use the natural log of forest loss as measured in square kilometers, which allows for intuitive interpretation of coefficient estimates. Second, for ease of cross county comparison, I construct a standardized measure of the incremental deforestation in a given county, $i$, in a given year, $t$. The normalized measure of deforestation, $N\_Defor_{i,t}$ is created as

Table 1.2: Summary Statistics

| **Variable** | **Mean** | **Std. Dev.** | **Min** | **Max** |
|---|---|---|---|---|
| Cattle | 103495.8 | 149900.1 | 0 | 2143760 |
| Cattle per km$^2$ | 34.42569 | 38.04579 | 0 | 246.1589 |
| Planted Soy Area (acres) | 8628.706 | 41781.47 | 0 | 608000 |
| Planted Soy Area (%) | 10.3572 | 22.78965 | 0 | 96.78 |
| Harvested Soy Area (acres) | 8621.128 | 41738.18 | 0 | 608000 |
| Harvested Soy Area (%) | 10.39455 | 22.86373 | 0 | 96.78 |
| Soy Price | 0.0588603 | 0.1116716 | 0 | 0.7302875 |
| Protected Area (%) | 18.20657 | 28.44221 | 0 | 100 |
| Share of Family Farming - Area | 0.4482019 | 0.2577647 | 0.0018942 | 1 |
| Share of Family Farming - Units | 0.8589439 | 0.1338808 | 0.0337079 | 1 |
| Share of Family Farming - Value | 0.5605426 | 0.3092991 | 0.0018625 | 1 |
| Number of clusters | 639 | | | |
| Obs per cluster | 11 | | | |

follows:[16]

$$N\_Defor_{i,t} = \frac{Defor_{i,t} - \overline{Defor_i}}{SD(Defor_i)} \tag{1.9}$$

In the above equation, $Defor_{i,t}$ corresponds to the amount of deforested area in that county and year as measured in square kilometers. $\overline{Defor_i}$ is the average annual deforested area in that county over the sample period and $SD(Defor_i)$ is the standard deviation of deforested area over the sample period.

The model suggests that the term limit status of the mayor should have no impact on deforestation prior to 2008, but term limited mayors will, on average, deforest more than their reelection eligible counterparts after 2008. Firgure 1.3 plots the average deforestation for all counties with mayors running for reelection in 2004. After 2004, mayors who win reelection are term limited and those who lose and are replaced by a mayor eligible for reelection in 2008. From 2001 to 2008, there is no systematic difference between the two groups despite the change in reelection incentives for the winning mayors.

[16]This is identical to the normalized measure of deforestation used by Assunção et al. (2012, 2013)

Figure 1.3: Deforestation from 2001-2008 by 2004 incumbent eligible counties

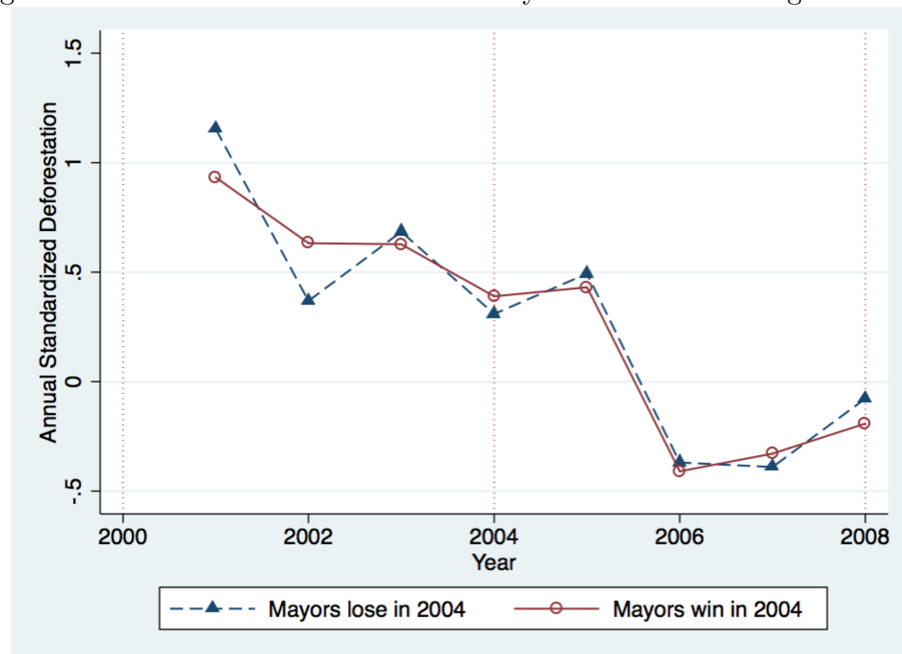



Figure 1.4: Deforestation from 2005-2012 by 2008 incumbent eligible counties

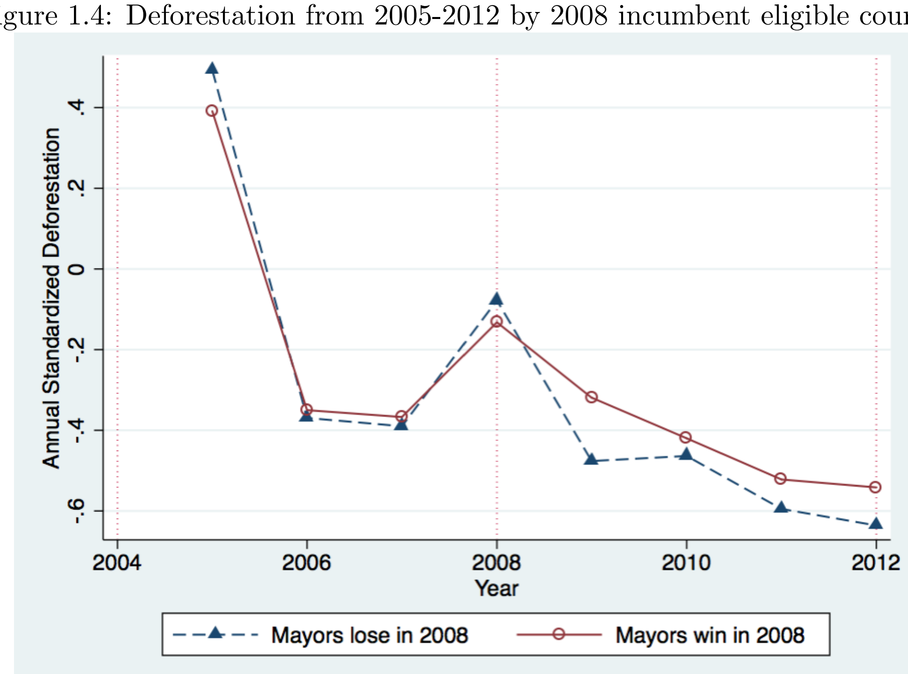



Figure 1.4 is constructed using the same design as Figure 1.3 but compares mayors who win reelection in 2008 with those who lose. Unlike the previous graph, the election

year (2008) also corresponds to the introduction of the blacklisting policy. As above, to the left of 2008, both groups are eligible for reelection and, from 2009 onward, the winning group is term limited while the losing group is replaced by mayors eligible for reelection in 2012. In contrast to Figure 1.3 after the election, the winning group experiences greater deforestation in each of the following years. These graphs motivate the formal analysis.

Utilizing the data described above, I examine whether deforestation is greater in counties with term limited mayors from counties with reelection-eligible mayors both before and after the introduction of the blacklisting policy via the following difference in differences model:

$$y_{i,t} = \beta_1 RE_{i,t} + \beta_2 RE_{i,t} \times Post_t + \gamma X_{i,t} + \alpha_i + \eta_{s,t} + \epsilon_{i,t} \tag{1.10}$$

In this model, $y_{i,t}$ is either of the two deforestation measures described above for municipality $i$ in year $t$ $RE_{i,t}$ is an indicator equal to 1 if a mayor is in his/her first term and thus eligible for reelection. $Post_t$ is an indicator equal to 1 if $t$ is greater than 2008. I include county-level fixed effects to account for time invariant characteristics within a county that may influence deforestation as well as state-by-year fixed effects to capture year-to-year unobserved factors that influence changes in average deforestation within a given state.

In equation (1.10), $\beta_1$ estimates the difference in deforestation between reelection eligible mayors and non reelection ineligible mayors prior to the blacklisting policy. The difference in deforestation between these two groups after the policy is the sum of $\beta_1$ and $\beta_2$. The predictions put forth by the model suggest that $\beta_1$ should equal zero and the sum of $\beta_1$ and $\beta_2$ should be positive.

For estimates of $\beta_1$ to be unbiased, I assume that, conditional on the controls, county fixed effects and state-by-year fixed effects, there are no unobserved factors that are correlated with the reelection status of the mayor in office that would influence deforestation. Time variation in reelection eligibility allows for separate identification of $\beta_1$ from the county fixed effect and the state-by-year fixed effect. This implies that any threat to identification of $\beta_1$ must differentially affect reelection eligible counties from term-limited counties within a given state and year through some channel other than the eligibility of the mayor to run again for office.

The identifying assumptions for an unbiased estimate of $\beta_2$ are slightly different than those for $\beta_1$. Unlike the traditional difference in differences model, the treatment and control status of counties changes with each election. The identifying assumption for $\beta_2$ is that nothing else changes the relationship of reelection eligibility to deforestation in 2008 other than the introduction of the blacklisting policy. One threat to identification would occur if the blacklisting policy itself influenced voters to reelect incumbent mayors who had demonstrated preferences for either high or low deforestation. The estimates of $\beta_2$ in this case would not reflect mayoral influence as a function of reelection status, but rather a systematic change in the preferences of voters. As voters are harmed by blacklisting, the introduction of the policy should alter their preferences towards politicians more willing and able to control deforestation, which would push the effect in the opposite direction of the theory. This effect would only attenuate the estimated differences in deforestation between reelection eligible mayors and term-limited mayors.

Table 1.3 compares the means of the control variables as well as time-invariant measures of landowner composition (small family farms versus large landowners) by reelection eligibility across the three election cycles. There are few differences significant at the 10 percent level, none of which appear in the 2009-2012 election cycle. Table 1.3 indicates that reelection eligible counties and term limited counties are largely similar on observable characteristics.

One shortcoming of the empirical analysis is that the deforestation data only covers one political cycle after the introduction of the blacklist policy. It could be the case that mayors now choose to delay deforestation to their second term. While the model suggests that deforestation should be lower under both first and second term mayors than it would have been without blacklisting, if the shifting of deforestation from first term to second term was sufficiently large average deforestation may not change. The empirics here can only examine the disciplining effect on deforestation for the 2009-2012 political cycle.

### 1.4.2 Voter punishment in response to blacklisting

To test whether voters punish politicians for blacklisting, I estimate an OLS regression on two measures of incumbent performance; incumbent vote share and an indicator for whether the incumbent was reelected. I estimate the relationship between blacklisting status and the incumbent performance while including controls for previous margin of victory for the incumbent, the number of total candidates in the election and including state by party fixed effects.

It should be stressed that the aim of these regressions is to provide suggestive evidence that the reelection mechanism may constrain mayoral behavior. If blacklisting has no effect on reelection prospects, any differences we observe in deforestation must be explained by a different channel. It should be cautioned, however, that these estimates should not be interpreted as casual, rather they are intended look for a potentially meaningful correlation in the data.

## 1.5 Results

The estimation results for the difference-in-difference model in equation (1.10) are listed in Table 1.4. Columns (1) and (2) present the preferred specification for the entire sample on the natural log of deforested area and the standardized measure of deforested area respectively. These columns impose the assumption that there were no differences in deforestation prior to the policy by reelection eligibility. Columns (3) and (4) are analogous to (1) and (2), but estimate the pre-2008 difference rather than assuming it to be zero.

The estimates indicate that reelection eligible mayors reduced annual deforestation 10% more than term limited mayors after the policy. This finding is statistically significant at the 5% level, using estimated standard errors clustered at the county-level. The results support the second hypothesis from the model, namely that reelection eligible mayors will be more responsive to the demand to control deforestation. The estimates of reelection eligibility on deforestation prior to the policy are not significantly different from zero, indicating no difference in deforestation from reelection incentives prior to the policy. This supports the first hypothesis of the model and also rules out political experience or ability as a confounding factor in the difference in deforestation under the policy.

Table 1.5 presents the results for the estimation of the effect of blacklisting on incumbent performance in 2012 by estimating incumbent vote share, defined as the number of votes in favor of the incumbent divided by the sum of the number of votes in favor of the incumbent and the number of votes in favor of the most popular, non-incumbent candidate. Column (1) begins with the single covariate, an indicator equal to one if the county was blacklisted in 2008-2011 and zero otherwise. I phase in controls for the previous margin of victory, a quadratic for number of candidates running in the election as well as state-by-party fixed effects in columns (2)-(4).

Being blacklisted is correlated with a 5.4-5.7 percentage point reduction in incumbent vote share. With the mean incumbent vote share in 2012 just above 50 percent, this represents an 11 percent reduction. The coefficient estimate in stable across the specifications and significant at the 5% level with the exception of column (4), where estimating the state-by-party fixed effects causes a loss in power.

To compare magnitudes of the direct effect of blacklisting on deforestation and the reduction associated with reelection eligibility, Table 1.6 presents estimates from equation (1.10) while including and indicator variable equal to one if the particular county is blacklisted in a given year. Results indicate that blacklisting lead to a 25% reduction in deforestation. That the estimate of reelection eligibility remains significant and of the same magnitude as in Table 1.4 assuages potential concerns that the effect found in Table 1.6 could be driven by differential blacklisting of reelection eligible mayors.

## 1.6 Robustness Checks

### 1.6.1 Testing source of reelection eligibility

While term limited mayors may only arrive at their final term by winning two consecutive elections, reelection eligible mayors may be serving a first term either because they beat an incumbent mayor in the previous election or they won an election following a term limited mayor. To check whether the source of reelection eligibility has a differential effect on deforestation, I estimate equation 1.10 while separating reelection eligible mayors by those who followed a term limited mayor and those who defeated an incumbent. Table 1.7 presents the coefficient estimates which correspond to those in Table 1.4. The magnitude of the estimates are very similar by reelection eligibility source. I test whether the coefficient estimates are equal and report the corresponding p-values in the second panel of Tabel 1.7. All tests fail to reject that any of the point estimates are significantly different by reelection eligibility source.

### 1.6.2 2004 and 2008 winners

In Table 1.8, I present results from a difference in difference estimate of the effect of winning an election on deforestation in the following years. In doing so, for each election, I limit the sample to be only those counties with reelection eligible mayors in office prior to the election and estimate the following equation.

$$y_{i,t} = \beta_1 Win_i \times Post_t + \gamma X_{i,t} + \alpha_i + \eta_{s,t} + \epsilon_{i,t} \tag{1.11}$$

where $Win_i$ is an indicator if the mayor in county $i$ wins the particular election (2004 or 2008) and $Post_t$ is an indicator equal to one if the year is past the particular election. Columns (1) and (2) consider only mayors eligible for the 2004 election and cover years 2002-2008. Prior to 2004, both groups are reelection eligible and after 2004 some mayors win reelection and are thus term limited. The results indicate no difference in deforestation as a result of the change in reelection eligibility status. Columns (3) and (4) present results for the 2008 elections. Estimates indicate that after 2008, counties which won reelection deforested nearly 9% more than the counties that lost reelection and were thus reelection eligible. These findings still support the main hypotheses of the model even while imposing strong restrictions on the data.

### 1.6.3 Verification with other deforestation data

Deforestation is difficult to quantify and, as mentioned above, there are particular aspects of the INPE data that may warrant verification with other deforestation data. The recent release of satellite data on deforestation from Hansen et al. (2013) offers global estimates of deforested area at the 30 meter by 30 meter (pixel) scale from 2002-2012. I create an alternative measure of deforested by counting the number of pixels labeled as deforested in a given year, meaning the pixel had appreciable forest cover in the previous year and was determined to have no forest cover in the corresponding year. Unlike the INPE data, this source is not limited to tree cover changes in amazon forest but will also measure change in other types of forest (i.e. cerrado). This allows not only for an opportunity to validate the main findings but also a placebo test, by looking for this pattern of deforestation in areas of the Legal Amazon with no amazon forest which were not subject to blacklisting. In the absence of the threat of blacklisting, reelection eligible mayors should have no incentive to reduce deforestation beyond that of term-limited mayors.

Columns (1) and (2) in Table 1.9 replicate the main findings of paper with the alternative measure of deforestation (the natural log of pixels labeled as deforested in a given year) in the 639 counties that have in which INPE measures amazon forest deforestation. I find no difference in deforestation between term-limited counties and non term-limited counties prior to 2008, and I find deforestation in 10-11% higher in term limited counties after 2008.

Columns (3) and (4) present results from the same specification, but measuring deforestation on the 121 other counties in the legal amazon, but which have no amazon forest cover and are not measured by the INPE data and are not subject to blacklisting. In these specifications, I find no difference prior to the policy in term limited and non term limited counties. Importantly, In Columns (3) and (4) respectively, I find no difference and a difference in the opposite direction in deforestation after the 2008 policy.

While the findings in Table 1.9 are supportive of the story in this paper, they come with caveats. The cerrado forest does not provide an adequate control for amazon forest due to differences in dynamics and deforestation pressures as well as other policies that apply to the amazon forest only. Furthermore, the small sample size of the non amazon forest group as well as the small amount of forest cover imply that these measures are likely to contain more noise than the deforestation measures in of amazon forest counties. Regarding the data itself, it is unclear whether deforestation measured in this dataset corresponds to clearing of primary forest or secondary forests (plantations, etc.) and in some cases, soy farms have been classified as forest cover (Tropek et al., 2014).

### 1.6.4 Limiting Sample to 2005-2012

As discussed in Section 2, prior to the real-time satellite monitoring introduced in 2004, there was little de facto enforcement of deforestation. The estimates of pre-blacklist difference in deforestation by reelection eligibility in Table 1.4 are estimated from a period of little deforestation monitoring and a period of private enforcement (but no collective punishment). The model predicts that differences by reelection eligibility will only occur under the threat of collective punishment. Table 1.10 estimates equation (1.10) while limiting the sample to the 2005 to 2012 period to verify that the evidence in favor of the two model hypotheses regarding deforestation are not driven by a period of little enforcement. The coefficient estimates are similar in magnitude to those in Table 1.4 which indicates that the findings are not driven by the early 2002-2004 period.

### 1.6.5 Reelection outcomes and blacklisting

To show that blacklisting had a measurable impact on whether or not a politician won reelection, Table 1.11 presents estimates from a linear probability model of reelection in 2012 as a function of blacklisting status.[17] The columns in Table 1.11 are analogous to those in Table 1.5. Results indicate that mayors running for reelection were reelected 16% less often if they were blacklisted than mayors in counties that were not blacklisted.

[17]The linear probability model is used in place of a nonlinear probit model to allow for the inclusion of state by party fixed effects in the final column.

## 1.7 Concluding Remarks

In the report evaluating local institutions and deforestation in the Brazilian Amazon, May et al. (2010) state, "A chronic problem in the Amazon is elite capture of public institutions with regulatory responsibilities for access to and use of forest resources, for private economic and political interests, associated with land speculation, illegal logging, cattle ranching, tax evasion, drug trafficking, patron-client relationships and electoral campaigns." (page 36) Any policy aimed to slow deforestation in the Amazon (as well as other tropical forest rich countries) must face the reality that the local political context may impede effectiveness or even make certain policies untenable.

This paper seeks to better understand how local political incentives affect deforestation in Brazil. When the demands of the electorate are aligned with natural resource use that maximizes private rents, politicians face little trade off between private rents and reelection. However, when the demands of the electorate change - as in the case of the collective punishment policy - so as to no longer favor private rent maximizing resource use, reelection eligible politicians slow deforestation more than their term-limited counterparts. In the absence of political accountability, local political authority over forest resources may lead to patterns of forest use that are at odds with the electorate.

This paper presents a model of lobbying and deforestation with mayors who seek reelection. The equilibrium of the model suggests that, under the threat of blacklisting, reelection eligible mayors are more likely to limit deforestation in order to win reelection and reelection ineligible mayors are more likely to increase deforestation at the expense of the voters.

The data supports these predictions. Using variation in term limit status, there is no difference in deforestation by reelection eligibility prior to the blacklisting policy. After the policy, reelection eligible mayors reduce deforestation 10% more than their reelection ineligible counterparts. The evidence suggests that voters do in fact punish mayors when their county is blacklisted. They offer blacklisted incumbents less vote share in the following election and reelect them at lower rates.

Many have argued that the decentralization of forest authority to the local level may be effective to slow deforestation. The findings of this paper suggest that such policies should consider local political institutions as decentralization could create more rents for local politicians if voters are unable to hold politicians accountable for their actions.

Table 1.3: Covariate Balance by Election Cycle

| | | Reelection Eligible | Term Limited | RE-TL |
|---|---|---|---|---|
| **Year** | **Variable** | **Mean** (St. Dev) | **Mean** (St. Dev) | **[P-value]** |
| 2002 | Cattle per km^2 | 27.80 | 29.51 | -1.72 |
| | | (31.99) | (33.78) | [0.51] |
| | Planted Soy Area (%) | 7.85 | 5.31 | **2.54*** |
| | | (20.77) | (17.43) | [0.09] |
| | Soy Price | 0.047 | 0.043 | 0.004 |
| | | (0.10) | (0.10) | [0.61] |
| | Protected Area (%) | 22.07 | 17.82 | **4.26*** |
| | | (30.13) | (27.51) | [0.06] |
| | Share of Family Farming - Area | 0.44 | 0.45 | -0.01 |
| | | (0.26) | (0.26) | [0.69] |
| | Share of Family Farming - Value | 0.57 | 0.56 | 0.01 |
| | | (0.31) | (0.31) | [0.68] |
| | Share of Family Farming - Units | 0.86 | 0.85 | 0.01 |
| | | (0.13) | (0.13) | [0.43] |
| OBSERVATIONS | | 327 | 312 | |
| 2005 | Cattle per km^2 | 36.81 | 34.50 | 2.31 |
| | | (40.96) | (42.12) | [0.56] |
| | Planted Soy Area (%) | 10.23 | 14.66 | **-4.43*** |
| | | (21.87) | (26.40) | [0.07] |
| | Soy Price | 0.059 | 0.066 | -0.007 |
| | | (0.10) | (0.10) | [0.47] |
| | Protected Area (%) | 21.45 | 20.79 | 0.66 |
| | | (29.62) | (27.81) | [0.80] |
| | Share of Family Farming - Area | 0.44 | 0.45 | -0.01 |
| | | (0.25) | (0.27) | [0.79] |
| | Share of Family Farming - Value | 0.56 | 0.56 | 0.005 |
| | | (0.30) | (0.33) | [0.87] |
| | Share of Family Farming - Units | 0.86 | 0.85 | 0.006 |
| | | (0.13) | (0.14) | [0.63] |
| OBSERVATIONS | | 493 | 146 | |
| 2009 | Cattle per km^2 | 35.34 | 34.19 | 1.15 |
| | | (38.74) | (35.50) | [0.71] |
| | Planted Soy Area (%) | 10.54 | 9.99 | 0.55 |
| | | (23.38) | (21.86) | [0.77] |
| | Soy Price | 0.059 | 0.062 | -0.003 |
| | | (0.11) | (0.11) | [0.79] |
| | Protected Area (%) | 23.50 | 25.75 | -2.26 |
| | | (30.03) | (32.92) | [0.40] |
| | Share of Family Farming - Area | 0.45 | 0.44 | 0.01 |
| | | (0.26) | (0.25) | [0.62] |
| | Share of Family Farming - Value | 0.55 | 0.58 | -0.03 |
| | | (0.31) | (0.30) | [0.33] |
| | Share of Family Farming - Units | 0.85 | 0.87 | -0.02 |
| | | (0.14) | (0.13) | [0.20] |
| OBSERVATIONS | | 430 | 209 | |

Asterisks correspond to the following p-values: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.4: Difference in Differences Results from Model of Reelection Eligibility on Deforestation

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | ln(Defor) | SD_defor | ln(Defor) | SD_defor |
| Reelection Eligible | | | 0.0358 | 0.0244 |
| | | | (0.0282) | (0.0341) |
| Reelection Eligible x Post | -0.0964** | -0.103** | -0.138* | -0.132** |
| | (0.0474) | (0.0425) | (0.0599) | (0.0609) |
| Observations | 7,029 | 7,029 | 7,029 | 7,029 |
| R-squared (within) | 0.460 | 0.382 | 0.461 | 0.384 |
| Number of counties | 639 | 639 | 639 | 639 |

The dependent variable in columns (1) and (3) is the natural log of deforested area as measured in square kilometers, the dependent variable in columns (2) and (4) is a normalized measure of deforestation measured in standard deviations. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area and cover 2002 to 2012. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.5: Incumbent vote share in 2012 elections and blacklisting

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Blacklisted | -0.0546** | -0.0556** | -0.0570** | -0.0540* |
| | (0.0248) | (0.0243) | (0.0220) | (0.0311) |
| Prev Margin of Victory | | 0.229** | 0.211** | 0.244** |
| | | (0.0997) | (0.0883) | (0.109) |
| Number of Candidates | | | -0.0976*** | -0.0996** |
| | | | (0.0337) | (0.0431) |
| Number of Candidates$^2$ | | | 0.00663 | 0.00691 |
| | | | (0.00429) | (0.00574) |
| State x Party Fixed Effects | No | No | No | Yes |
| Observations | 365 | 365 | 365 | 365 |
| R-squared | 0.007 | 0.025 | 0.157 | 0.417 |

Standard errors are in parenthesis and clustered at the state level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.6: Results from Model of Reelection Eligibility and Blacklisting on Deforestation

| | (1)<br>ln(Defor) | (2)<br>SD_defor |
|---|---|---|
| Reelection Eligible x Post | -0.0964** | -0.104** |
| | (0.0474) | (0.0425) |
| Blacklist | -0.252*** | -0.192*** |
| | (0.0789) | (0.0519) |
| Observations | 7,029 | 7,029 |
| R-squared (within) | 0.462 | 0.384 |
| Number of counties | 639 | 639 |

The dependent variable in column (1) is the natural log of deforested area as measured in square kilometers, the dependent variable in columns (2) is a normalized measure of deforestation measured in standard deviations. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area and cover 2002 to 2012. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.7: Reelection Eligibility by eligibility type on Deforestation

| | (1)<br>ln(Defor) | (2)<br>SD_defor | (3)<br>ln(Defor) | (4)<br>SD_defor |
|---|---|---|---|---|
| Reel Elig via Term | | | 0.0298 | 0.0480 |
| | | | (0.0333) | (0.0463) |
| Reel Elig via nonTerm | | | 0.0386 | 0.0155 |
| | | | (0.0396) | (0.0409) |
| Reel Elig via Term x Post | -0.0826 | -0.118** | -0.127 | -0.182** |
| | (0.0591) | (0.0519) | (0.0804) | (0.0848) |
| Reel Elig via nonTerm x Post | -0.103** | -0.0955** | -0.144** | -0.110* |
| | (0.0513) | (0.0456) | (0.0666) | (0.0652) |
| *P-value of equality on Pre* | | | 0.839 | 0.493 |
| *P-value of equality on Post* | 0.704 | 0.630 | 0.823 | 0.336 |
| Observations | 7,029 | 7,029 | 7,029 | 7,029 |
| R-squared (within) | 0.460 | 0.384 | 0.461 | 0.385 |
| Number of counties | 639 | 639 | 639 | 639 |

The dependent variable in columns (1) and (3) is the natural log of deforested area as measured in square kilometers, the dependent variable in columns (2) and (4) is a normalized measure of deforestation measured in standard deviations. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area and cover 2002 to 2012. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.8: Reelection eligible counties, electoral outcomes and deforestation

| | (1) ln(Defor) | (2) SD_defor | (3) ln(Defor) | (4) SD_defor |
|---|---|---|---|---|
| | 2004 Election | | 2008 Election | |
| Election Won x Post Election | -0.0436 | -0.0260 | 0.0881* | 0.0951** |
| | (0.0630) | (0.0827) | (0.0468) | (0.0480) |
| Years | 2002 - 2008 | | 2005 - 2012 | |
| Observations | 2,289 | 2,289 | 3,944 | 3,944 |
| R-squared (within) | 0.326 | 0.292 | 0.434 | 0.362 |
| Number of counties | 327 | 327 | 493 | 493 |

The dependent variable in columns (1) and (3) is the natural log of deforested area as measured in square kilometers, the dependent variable in columns (2) and (4) is a normalized measure of deforestation measured in standard deviations. Columns (1) and (2) restrict the sample to only counties with mayors eligible for reelection in 2004. Columns (3) and (4) restrict the sample to only counties with mayors eligible for reelection in 2008. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.9: Results from Model of Reelection Eligibility on Deforestation in Amazon and Non Amazon forest

| | (1) ln(Defor) | (2) ln(Defor) | (3) ln(Defor) | (4) ln(Defor) |
|---|---|---|---|---|
| | Amazon Forest | | Non Amazon Forest | |
| Reelection Eligible | | -0.0155 | | 0.0278 |
| | | (0.0439) | | (0.0366) |
| Reelection Eligible x Post | -0.118*** | -0.0994** | 0.0975 | 0.0647 |
| | (0.0350) | (0.0439) | (0.0650) | (0.0831) |
| Observations | 7,029 | 7,029 | 1,331 | 1,331 |
| R-squared (within) | 0.486 | 0.386 | 0.392 | 0.392 |
| Number of counties | 639 | 639 | 121 | 121 |

The dependent variable in all columns is the natural log of deforested 30 meter by 30 meter pixels. Columns (1) and (2) include counties in other INPE specifications, and columns (3) and (4) include counties with no amazon forest but lie in the legal amazon. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area and cover 2002 to 2012. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.10: Difference in Differences Results from Model of Reelection Eligibility on Deforestation after 2004

| | (1)<br>ln(Defor) | (2)<br>SD_defor | (3)<br>ln(Defor) | (4)<br>SD_defor |
|---|---|---|---|---|
| Reelection Eligible | | | 0.0187 | 0.0170 |
| | | | (0.0498) | (0.0509) |
| Reelection Eligible x Post | -0.0784* | -0.0970** | -0.103 | -0.0744 |
| | (0.0430) | (0.0424) | (0.0787) | (0.0812) |
| Observations | 5,112 | 5,112 | 5,112 | 5,112 |
| R-squared (within) | 0.430 | 0.351 | 0.430 | 0.351 |
| Number of counties | 639 | 639 | 639 | 639 |

The dependent variable in columns (1) and (3) is the natural log of deforested area as measured in square kilometers, the dependent variable in columns (2) and (4) is a normalized measure of deforestation measured in standard deviations. All regressions include county fixed effects, state-by-year fixed effects, agricultural controls and protected area and cover 2005 to 2012. Standard errors are in parenthesis and clustered at the county level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

Table 1.11: 2012 Incumbent reelection indicator and blacklisting

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Blacklisted | -0.155 | -0.159* | -0.159* | -0.0598 |
| | (0.0999) | (0.0933) | (0.0946) | (0.119) |
| Prev Margin of Victory | | 1.070*** | 1.042*** | 1.097*** |
| | | (0.290) | (0.274) | (0.310) |
| Number of Candidates | | | -0.138 | -0.189* |
| | | | (0.0865) | (0.103) |
| Number of Candidates$^2$ | | | 0.00779 | 0.0120 |
| | | | (0.0108) | (0.0131) |
| State x Party Fixed Effects | No | No | No | Yes |
| Observations | 365 | 365 | 365 | 365 |
| R-squared | 0.006 | 0.044 | 0.078 | 0.336 |

Results from a linear probability model using an indicator of incumbent victory as the dependent variable. Standard errors are in parenthesis and clustered at the state level. Asterisks correspond to the following p-values: *** p<0.01, ** p<0.05, * p<0.1

# Chapter 2

# The Underground Economy of Fake Antivirus Software

## 2.1 Introduction

Over the past few years, electronic crimes revolving around a class of malware known as *scareware* have become extremely lucrative ventures. The concept is simple; design a ploy through social engineering that exploits a computer user's fear of revealing sensitive information, losing important data, and/or causing irreversible hardware damage. The most common form of scareware is *fake antivirus* (AV) software, also known as "rogue security software." More specifically, a fake AV program impersonates an antivirus scanner and displays misleading or fraudulent alerts in an attempt to dupe a victim into purchasing a license for a commercial version that is capable of removing nonexistent security threats. Some fake AV programs may also lock down system functionality to prevent victims from accessing files or web sites or from creating new processes, such as Windows Explorer, Task Manager, and a Command Prompt under the false pretense that it is for the victim's own protection. In addition, we have observed fake AV software that contains hidden backdoor capabilities, enabling the program to be used for other malicious purposes, such as launching distributed denial-of-service (DDoS) attacks against adversaries.

Over the past year, we have been able to acquire backend servers for several multi-million dollar criminal operations selling fake AV products. These fake AV businesses are run out of Eastern Europe and utilize affiliate networks known as *partnerka* to distribute the rogue software (Samosseiko, 2009). These partnerka networks use various pseudonyms, and operate by recruiting affiliates to install their software on as many computers as possible. In exchange, the affiliates receive a commission for driving traffic to landing pages, malware installations (also known as *loads*), and fake AV sales. Moreover, some partnerka offer additional incentives to the most successful affiliates with prizes including expensive cars, computers, and cell phones Krebs (2009a).

Since we have access to the servers used by these criminal organizations, we are able to directly analyze the tools that are used to create the fake AV products, including programs that assist perpetrators in controlling the malware's behavior and brand names, as well as custom *packers* that obfuscate the malware to evade detection by legitimate antivirus products. Some fake AV groups even make use of third-party commercial services to track the detection rates by the most popular antivirus vendors (e.g., McAfee, Symantec, and Trend Micro) Krebs (2009b), and they tweak their obfuscation algorithms until a low detection rate is achieved. We also have access to the instruments that are used to direct traffic to fake AV web sites, the infrastructure that prolongs the longevity of the operations, and a very detailed view of the financial profits that fuel these illicit enterprises. Interestingly, the miscreants behind fake AV products even offer refunds to victims who are persistent, in order to reduce the amount of credit card chargebacks, which we will discuss in more detail later.

Although various aspects of fake AV software have been studied, there are many facets of these operations that are not well understood, including the modus operandi of the criminals, the amount of money involved, the victims who purchase the software, the affiliate networks that promote the campaigns, and the flow of money from the victims' credit cards, to the payment processors, to the bank accounts controlled by the criminals. In this paper, we attempt to fill this void by presenting the analysis of several criminal organizations that sell fake AV products. More specifically, we make the following contributions: First, we provide an in-depth analysis of fake AV operations and present detailed statistics based on the analysis of more than a dozen servers belonging to several criminal organizations. This is the most comprehensive, large-scale study of fake AV campaigns that highlights different aspects of their operations from the infection process, to the financial complexities of maintaining a fraudulent business. Second, we examine how fake AV campaigns are managed and orchestrated, from the ringleaders' point of view. We discuss the software infrastructure that is utilized, the functionality it provides, and its role in the underground economy. lastly, we present an economic model that encapsulates financial patterns that are indicative of fake AV ventures. Our intent is to formalize the essential factors of these operations and to identify potential weaknesses that can be exploited to increase the criminals' functional and operational costs.

## 2.2 Technical Background

Before we present the financial logistics, we first discuss the methods that are utilized to infect machines with fake AV software and the infrastructure behind the process. In addition, we present details about three particular criminal operations running fake AV businesses. To protect ongoing law enforcement investigations, we refer to these three ventures as $AV_1$, $AV_2$, and $AV_3$. Note that we currently see ongoing activity (e.g., new malware samples, installations and online advertisements) from all three fake AV operations.

### 2.2.1 Infection Methods

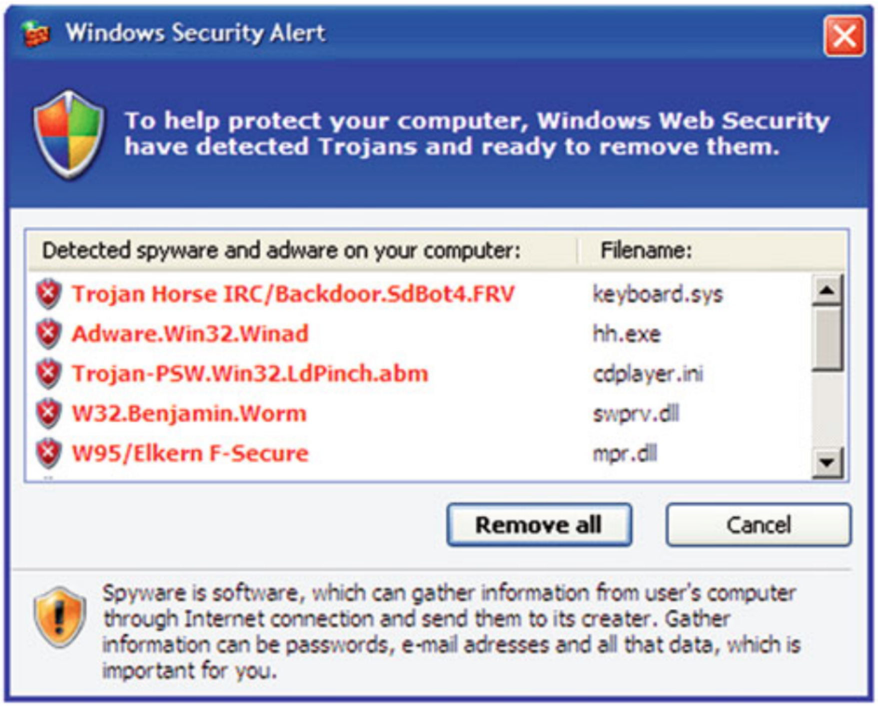
There are three primary infection methods used by fake AV distributors to propagate their malware: social engineering, drive-by-download attacks, and botnets. In this section, we present how these strategies are used to infect as many computers as possible with fake AV malware.

One of the most popular infection methods uses social engineering techniques to convince a victim to voluntarily install the fake AV. To launch this attack, a malicious web page displays a window in the browser (e.g., via JavaScript or Adobe Flash) that pretends that the machine has been infected with malware. An example is shown in Figure 2.1. To fix the security problem, the window also contains a link to a program that presumably helps to clean up the infection. Of course, this program is the fake AV software that attackers aim to install.

A second technique to install fake AV software is via drive-by download attacks. In a drive-by download attack, a web site is prepared with malicious scripts that exploit vulnerabilities in the web browser or one of its plugins. When the exploit is successful, the fake AV malware is installed automatically, without the user's knowledge or consent.

Both in the case of fake alerts and drive-by downloads, the initial goal of the attacker is to drive as many web visitors to their malicious web pages (sometimes called landing pages) as possible. In order to achieve this objective, attackers often make use of *blackhat search engine optimization* (SEO). Their intention is to poison search engine results by creating landing pages that contain popular search phrases. Many of these campaigns target current events such as the death of a celebrity, natural disasters, and holidays. Blackhat SEO relies on the fact that when search engine crawlers index a web site they

Figure 2.1: Alerts from a fake antivirus advertisement.




identify themselves through the HTTP `User-Agent` field (e.g., googlebot). Thus, a site under an attacker's control can serve content that contains popular keywords that a search engine will use in the computation of the page rank. If the process is done correctly, the landing page is ranked high in the search engine's results for these popular keywords.

When a user clicks on a search engine result that leads to a blackhat SEO landing page, the server analyzes the user's web browser (via the `User-Agent` header), and the referring web site (through the HTTP `Referer` field). The tools that are used to manage these SEO campaigns are known in the underground economy as a *traffic direction system* (TDS). These TDSs can leverage the header information to distinguish between search engine bots and web browsers. In order to avoid detection, TDSs often take additional countermeasures such as resolving the visitor's IP address to a geographic location and recording the number of accesses. Once the TDS has verified the traffic, a user is redirected a number of times to a landing page. This landing page will then launch a social engineering or drive-by download attack, as described previously.

Note that most TDSs also define a *time-to-live* (TTL) value that specifies how long a particular redirection URL will remain active. Most TTL values are very short, which makes it more difficult for security researchers to track active campaigns.

An alternative approach to using blackhat SEO techniques for traffic generation is to exploit the distribution systems and ubiquity of online ad networks. An attacker may compromise a legitimate ad network, or sign up as an advertiser to display malicious advertisements disguised as free pornography, missing audio/video codecs, or virus scans that perform similar social engineering attacks to con visitors into installing their malware. Online ad networks are also frequently used in conjunction with drive-by-download attacks, known collectively as *malvertisements*, to covertly install the fake AV software (without user interaction or permission).

A third infection method is through *botnets*, a collection of compromised computers under the control of an attacker. Several large botnets, such as Koobface, Conficker, and Bredolab, have been known to distribute fake AV software to machines under their control, which is believed to be one of their top sources of revenue (Kirk, 2010; Poulsen, 2009; Villeneuve et al., 2010).

Once fake AV software has been installed on the victim's machine (either voluntarily through social engineering or involuntarily through a drive-by attack or botnet), intrusive nags will be shown continuously to the victim, warning of "malware infections" or "intrusion attempts" that pose a risk to the user's system. At this point, the fake AV software usually advertises itself as a free trial version with limited functionality (i.e., detection only). If a victim wants to remove the malware infections, they must upgrade to a commercial version by purchasing a license key. When a victim clicks the software's purchase button, they are taken to one of the fake AV company's web sites. After a victim enters their personal information and credit card, they are sent a license key (e.g., through email) that essentially deactivates the bogus malware alerts, providing the user with a sense that their purchase was valuable.

### 2.2.2 Infrastructure

Similar to any other legitimate online business, when a fake AV company's servers are down, they lose potential revenue streams. Therefore, there are a number of measures that these organizations take to ensure the availability of their infrastructure. The first strategy is to deploy an array of proxy servers that are publicly visible. The sole purpose of these proxies is to relay content to one or more backend servers as shown in Figure 2.2. More specifically, these machines communicate directly with users that are redirected to a landing page or infected hosts that purchase a license. The proxy servers are typically partitioned depending on the specific role that they fulfill (e.g., TDS servers are not reused for relaying sales information). The main purpose of the front-end servers is to thwart mitigation efforts. Hence, taking down one, or even several, of these machines often has little impact, since the domain name address records that point to these servers can be changed quickly and easily. These front-end servers are designed to be lightweight and expendable, and typically have an automated deployment program that accelerates the process of creating new proxy nodes.

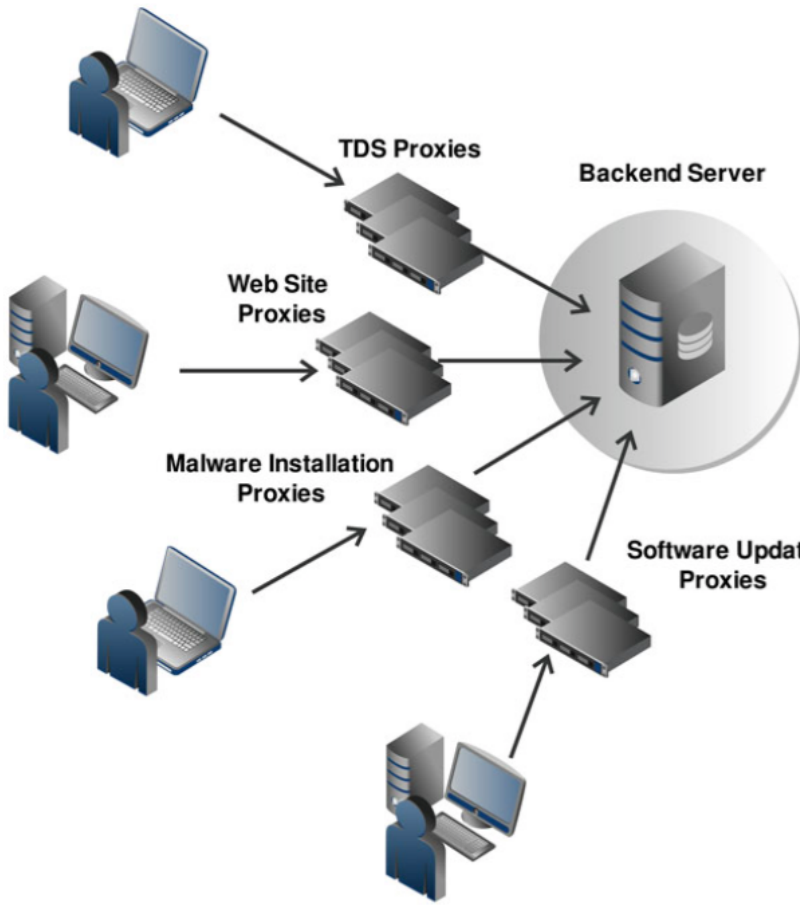
The main drawback of proxies (from an attacker's point of view) is that when a defender obtains access to one of these front-end servers (or monitors their ingress and egress network traffic), she can learn the location of the backend infrastructure. To address this problem and to further hide the location of the backend, the miscreants of fake AV operations may use multiple tiers of proxy servers. However, each extra tier will introduce additional network delay that could make a user who is purchasing a fake AV product more suspicious. In our experience, most fake AV operations use only one tier of proxy nodes. Thus, we were able to locate the backend infrastructure by tracking the network traffic from an infected host to a proxy node to the backend servers. By taking down the backend servers, the entire fake AV operation is disrupted (i.e., servers relaying sales, malware installations, and TDS become inoperable).

A second, important strategy is to register a large number of domain names. The domain names fulfill several purposes. First, it makes the fake AV web site look more legitimate (e.g., the domains are usually related to antivirus or security keywords). Second, the large number of domains makes takedown efforts more difficult, since the DNS records can be changed to point to any of their proxy servers. In addition, the reputation of a fake AV domain will decline as more people are defrauded, and many of the domains will become blacklisted. As a result, domain registrars may ultimately suspend some of the fake AV domains. Overall, the $AV_1$ crew purchased 276 domains, 17 front-end servers, and one back-end server. Similarly the $AV_2$ operation registered at least 188 domains, managed 16 front-end servers, and two back-end servers. We did not have complete visibility over the total number of domains used by $AV_3$, but from our observations, the infrastructure was similar to the others with a large number of free domains registered through the `co.cc` top-level domain (TLD), and approximately 20 front-end servers, and one back-end server.

## 2.3 Data Collection

In the following section, we describe the process that facilitated our efforts in obtaining access to these fake antivirus backend servers and the data we collected. The main tool that we utilized to analyze the fake AV malware was Anubis, a system that dynamically analyzes binary programs via runtime analysis (Bayer et al., 2009b). Anubis runs a Windows executable and documents the program's behavior, including system

Figure 2.2: Tiered infrastructure for many online criminal operations




We were able to obtain copies of three different fake AV organization's backend servers (in the shaded circle above) that control the entire operation.

modifications, processes creation, and network activity. ANUBIS is able to process on the order of tens of thousands of samples per day, providing us with a comprehensive view of the current malware landscape (Bayer et al., 2009a).

By searching through the network connections logged in the ANUBIS database, we were able to identify a number of unique network signatures commonly used by fake antivirus software. More specifically, when fake AV is installed, it often *phones home*, by connecting back to servers under the control of the fake AV criminal organization. For example, infected machines made an HTTP to notify the criminals of the installation and to credit the affiliate responsible for the infection. The parameters $p$ and $s$ provided details about the type and name of the malware

After observing network signatures associated with these fake AVs, we contacted the hosting providers whose servers were being used for controlling these operations. We provided them with network traces, malware samples, and other evidence that revealed the location of the servers that were situated within their network. The hosting providers responded by taking these servers down, and they provided us with direct access to the information stored on them. Note that we had previously collaborated with a number of these vigilant ISPs in the U.S. and abroad through FIRE (Stone-Gross et al., 2009b), our network reputation service that tracks where malicious content resides on the Internet.

In total, we were able to get a complete snapshot of 21 servers: 17 of which were proxy nodes, and 4 of which were backend servers. The information that we collected from these servers included data for $AV_1$ for approximately 3 months from January through April 2010, 16 months from January 2009 through May 2010 for $AV_2$, and from March 2008 through August 2010 for $AV_3$. From these data sources, we have a view of nearly the entire operation including web site source code, samples of the fake AV malware, and databases. The most interesting information is contained in the database records, which document everything from malware installations, fake AV sales, refunds, technical support conversations to the TDSs controlling the fake AV landing pages.

## 2.4 Following the Money Trail

Now that we have provided a summary of the fake AV infrastructure and our data sources, we will focus on the financial aspects that drive the sales of fake AV software. In particular, we analyze the flow of money from a victim to the criminals and their affiliates. In addition, we examine the ways in which the fake AV groups manage to stay under the radar when interacting with credit card payment processors.

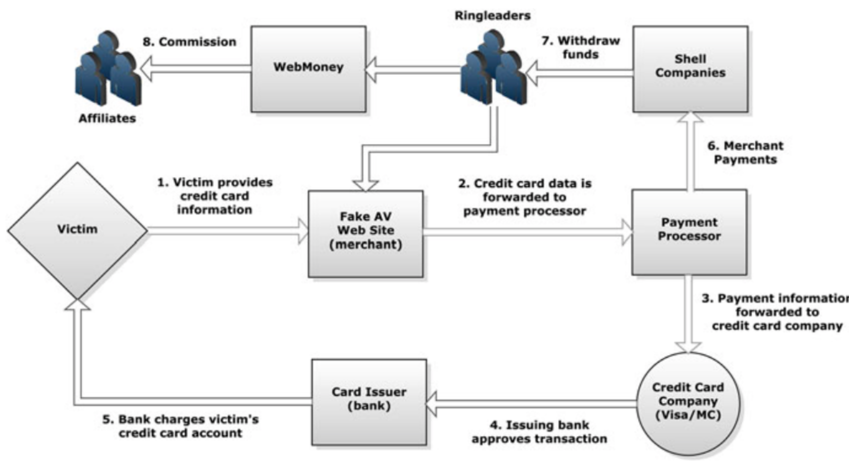
### 2.4.1 Transaction Process

Before we present the detailed statistics of sales, revenue, chargebacks and refunds, we introduce an overview of the various entities involved in a fake antivirus business. The transaction process, as shown in Figure 2.3, begins when a victim purchases the rogue AV software. This purchase is done through the fake AV company's web site (Step 1), where the victim enters her credit card information. The fake AV business (i.e., the merchant) then submits the credit card data to a third-party payment processor (Step 2). The payment processor forwards the information through one of the major credit card companies (Step 3), who requests authorization from the credit card issuer (Step 4). If the credit card issuer (i.e., a bank) approves the transaction, the victim's credit card is charged (Step 5), and the credit card company notifies the payment processor of the successful sale. Periodically (e.g., biweekly or monthly), the payment processor deposits funds into bank accounts set up by the fake AV businesses (Step 6). The ringleaders of the fake AV operation then withdraw the funds (Step 7) and pay a commission to their affiliates (Step 8). We will provide more details about this process in the following sections.

### 2.4.2 Sales

There are a number of factors that contribute to whether a victim purchases a license, such as the aggressiveness of the fake AV software (e.g., frequency of alerts, type of threats, and whether system performance is affected). In addition, the price and subscription models offered by most fake antivirus products play an interesting role, with subscriptions that range from 6-month licenses to lifetime licenses. The $AV_1$ operation

Figure 2.3: High-level overview of the transaction process for fake antivirus businesses.




offered licenses for 6-months at \$49.95, 1-year at \$59.95, and 2-years at \$69.95. These options were purchased almost uniformly with rates of 34.8%, 32.9%, and 32.3%, respectively. The $AV_2$ company's products also offered 6-month licenses at \$49.95, 1-year at \$69.95, and a lifetime license at \$89.95. The 6-month option was the most popular (61.9%), followed by the lifetime license (24.6%) and the 1-year license (13.5%). The products sold by $AV_3$ were priced at \$59.95 for a 1-year license and \$79.95 for a lifetime license. All of $AV_3$'s products were also bundled with a mandatory \$19.95 fee for 24x7 customer support services, bringing the total price to \$79.90 for the yearly license (purchased by 83.2% of victims) and \$99.90 (purchased by 16.8% of the victims) for the lifetime license.

In total, $AV_1$ "trial" products were installed 8,403,008 times, which resulted in 189,342 sales, or upgrades to the "commercial" version (a conversion rate of 2.4%) in only 3 months. Likewise, $AV_2$'s programs were installed 6,624,508 times, with 137,219 victims that purchased the fake antivirus over 16 months. That is a conversion rate of approximately 2.1%. The $AV_3$ business sold 1,969,953 licenses out of 91,305,640 installations from March 2008 through August 2010 (a conversion rate of approximately 2.2%).

The total victim loss from the three fake AV operations was \$11,303,494, \$5,046,508, and \$116,941,854 from $AV_1$, $AV_2$, and $AV_3$, respectively. Figure 2.4 shows the cumulative daily revenue for each of these fake antivirus operations. If we extrapolate these profits over one year, the $AV_1$ crew was on track to earn more than \$45 million dollars per year, while the $AV_2$ group earned approximately \$3.8 million per year. The largest and most profitable operation was $AV_3$, which raked in an average of \$48.4 million dollars per year.

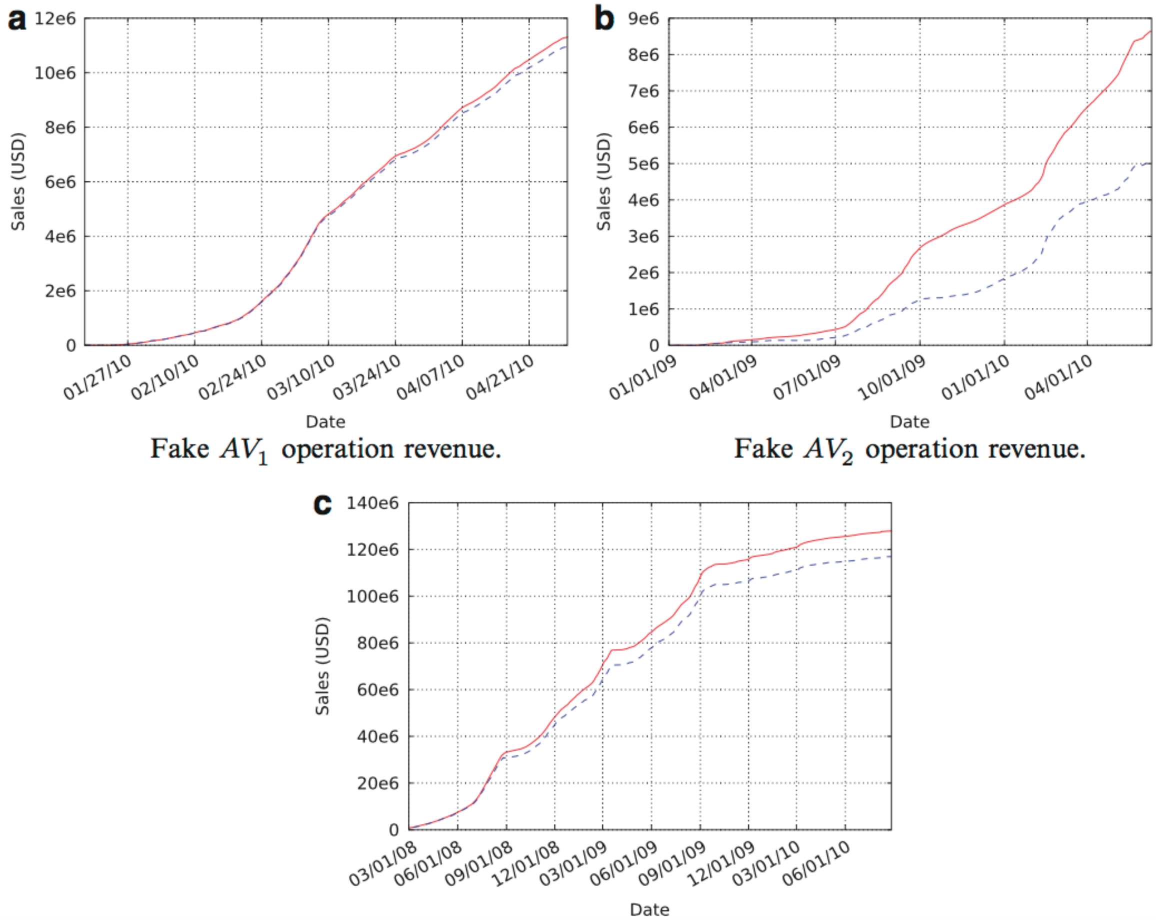
As we will discuss in Section 2.4.4, some credit card transactions were reported to be fraudulent and were credited back to the victim. Interestingly, victim complaints force these illegitimate firms into a complex position with their payment processors, as we will discuss in the following sections.

### 2.4.3 Payment Processors

An interesting facet of fake AV sales is the process in which credit card transactions are handled. In particular, payment processors (also known as payment service providers) are an integral part of every sale. Without these processors, fake AV operations would not be able to accept credit card payments. This would make it not only harder for a victim to purchase the product (i.e., they would have to use an alternative form of payment, such as cash, check, or money order), but it would also likely raise red flags that the software may be fraudulent. Note that payment processors must maintain a degree of legitimacy, or they risk losing the ability to accept major credit cards. For instance, a payment processor known as ePassporte lost the rights to accept Visa credit cards, due to a large amount of fraudulent transactions, money laundering, and other questionable activities (Krebs, 2010a). Note that the $AV_2$ crew at one point set up an ePassporte merchant account for processing credit card transactions.

Perhaps the most notorious payment service provider is Chronopay, which is head-

Figure 2.4: Three criminal organizations' revenue from fake antivirus sales.




Fake $AV_1$ operation revenue.

Fake $AV_2$ operation revenue.

Fake $AV_3$ operation revenue.

The solid line displays the total revenue, while the dotted line displays the revenue net of chargebacks and refunds.

quartered in the Netherlands and operated by Russian businessmen. Chronopay has long been associated with processing transactions for various forms of online criminal organizations (Mick, 2010). However, Chronopay also provides legitimate services to large organizations such as Electronic Arts, Kaspersky, and charities including the World Wildlife Federation, Greenpeace, and UNICEF. Because the volume of legitimate transactions from these businesses may far outweigh the fraudulent activities, major credit card companies may be hesitant to sever ties with Chronopay. Note that all three fake

AV businesses that we analyzed used Chronopay's credit card payment services.

There were several other, smaller payment processors that the fake AV operations used for credit card transactions. Interestingly, we found communications between one of these small payment processors and the fake AV perpetrators that revealed that the payment service provider was well aware of the fake AV business and even offered advice to help the group sell more products. There are a number of tricks that some of these dishonest payment service providers perform in order to benefit from fraudulent transactions. First, payment processors may offer *high-risk merchant accounts*, where the processor may earn close to 15% for each transaction. These are typically for questionable businesses that have significant problems with customer complaints (e.g., online pharmacies or pornography). Second, we observed that some of these payment processors allow an illicit company to create multiple merchant accounts in which transactions are periodically rotated (approximately every 30-45 days) through each account, such that a single account is never flagged for fraudulent activities, since the transactions are distributed over all of the accounts.

### 2.4.4 Chargebacks and Refunds

Interestingly, all three fake antivirus groups that we studied offered a certain number of refunds to individuals who requested them. At first, it may seem counter-intuitive for a criminal operation that is selling fraudulent products to provide refunds to victims. However, it is important to keep in mind that these criminal organizations have to use legitimate (or semi-legitimate) credit card payment processors for every transaction. In addition, payment processors are required by statutory (federal regulations) and contractual obligations (PCI) to provide various levels of consumer protection against theft and fraudulent purchases. When a victim reports a fraudulent transaction to their credit card issuer, they are issued a credit, which is known as a *chargeback*. If a business receives too many chargeback complaints, the payment processor may sever ties with the company and prohibit further credit card transactions. Therefore, it is important to minimize the number of chargebacks, which has the effect of extending the lifetime of the fake AV operation.

Overall, $AV_1$ granted 5,669 refunds (3% of sales) at a cost of $346,039 (in addition to 1,544 chargebacks worth $94,963). In comparison, $AV_2$ issued 11,681 refunds (or 8.5% of sales) at a cost of $759,666 (in addition to 3,024 chargebacks valued at $183,107). $AV_3$ refunded 151,553 (7.1% of sales) for a total of $10,951,191 (with 30,743 chargebacks valued at $2,225,430). Note that the primary credit card processor for $AV_3$ temporarily froze $AV_3$'s merchant account for approximately one month in March 2009, due to a high number of chargebacks. After this incident, $AV_3$ offered more refunds, and the number of chargebacks dropped accordingly.

Another important factor that has an impact on chargebacks and refunds is how frequently a fake AV business changes the name of their product. This is due to the fact that after a short interval (typically 3-7 days), victim complaints start appearing on consumer web forums that are in turn indexed by search engines. Thus, a victim may perform a Google search for the name of the fake AV and find that other users have similar grievances and complaints. Interestingly, we found that $AV_2$ had significant server problems and maintained the same product names for an extended period of time. As a result, they had the highest chargeback and refund rates.

As we will discuss in Section 2.6, the amount and timing of refunds follows an interesting pattern, which indicates that the criminals maximize their profits by refunding just enough sales to remain under a payment processors chargeback limit.

### 2.4.5 Affiliate Programs

The financial incentives for cybercrime play an important role both in the type and amount of fraud. In order to infect as many machines as possible and therefore maximize sales, fake AV businesses rely upon affiliate networks based primarily in Eastern Europe known as *partnerka*. The backend servers that we obtained contained payment records to these partners. The profits for some of the affiliates are immense, with members earning as much as 30-80% commission from sales leads. Remarkably, the top affiliate of $AV_1$ made more than $1.8 million dollars in approximately two months. Over the course of these two months, there were a total of 44 affiliates who were paid (out of 140 that enrolled), with four earning more than $500,000, 11 in excess of $100,000, and 15 more than $50,000. The average affiliate income was approximately $60,000 per month. In comparison, $AV_2$ had 98 active affiliates out of 167 total registered, and stored records for 9 months of payments to these affiliates. Overall, five of these affiliates made more than $300,000, 16 earned more than $100,000, and 22 earned more than $50,000. The $AV_3$ operation had a total of 1,107 affiliates with 541 who were active. The top $AV_3$ affiliate earned $3.86 million, and three others made more than $1 million. There were 15 $AV_3$ affiliates that earned over $100,000, and 23 that were paid more than $50,000.

By comparing the affiliate email addresses across the three different fake AV partnerka, we were able to determine that 70 affiliate members were involved in multiple groups. Interestingly, there was one affiliate who was associated with all three fake AV businesses.

The affiliate payments were made through WebMoney, a virtual electronic currency. There are several advantages that WebMoney provides for criminal activities. In particular, all transactions are anonymous and irreversible. That is, once a transfer has occurred it cannot be voided, regardless of whether it was fraudulent. Other benefits include a very low transaction fee (0.8%), and a large number of places, especially in Eastern Europe, that will exchange WebMoney for local currencies.

### 2.4.6 Shell Companies

One of the most important parts of the financial system from a fake AV company's perspective is the ability to *cash out* earned funds. Thus, a fake AV company must open one or more bank accounts to receive merchant remittances from their payment processors. These accounts are typically set up and registered to fictitious *shell companies*. We observed accounts registered primarily in Europe and Asia, including the Czech Republic, Finland, Cypress, and Israel. Once money is deposited into a shell account, the ringleaders can directly withdraw the funds. However, criminals who are more cautious may opt to use the services of *money mules*. A money mule is a person who is recruited (usually under the pretense of a work from home job) to accept a bank deposit, withdraw the funds, and wire the money (minus a service fee) back to the criminals. This greatly minimizes the risk that a criminal will be apprehended when receiving funds. Unfortunately, we were not able to determine the precise method used by these three fake AV groups to withdraw funds. Nevertheless, we believe the money was probably picked up directly by the ringleaders (or one of their close associates), based on the geographic locations of the bank accounts.

## 2.5 Victims

In this section, we analyze the victims that purchased fake AV software. In particular, we will study various characteristics of victims including: geographic location, operating systems, and institutions. In addition, we will examine the technical support and customer service provided by the three fake AV businesses.

The largest concentration of victims (by far) was in the U.S. (76.9%) followed by the U.K., Canada, and Australia. This is likely due to the fact that the fake antivirus products are primarily written for English speakers (only a few of them had been translated to other languages). The most popular, compromised operating systems were Windows XP (54.2%), Windows Vista (30.8%), and Windows 7 (14.8%). Internet Explorer 7 was the most commonly used browser (65.6%). The most frequently used email addresses of customers of fake AV products were Yahoo, Hotmail, AOL, Gmail, and Comcast. Other residential ISPs placed in the top 10 including AT&T, SBC Global, Verizon, and Bellsouth. This indicates that most victims probably purchased the fake AV software for their personal computers at home. However, there were a number of sales from victims at commercial, government, and military institutions.

All three of the fake AV companies offered various forms of customer service and technical support. Customer service for fraudulent products may seem contradictory, but its purpose is clear: to reduce the number of refunds and victim complaints. Overall, the fake AV groups offered two types of support systems. The first was an online system where victims could open tickets describing their problems, and technical support representatives would periodically reply to these tickets. The second type of support system was an interactive, live chat service, where a victim would talk in real-time with technical support personnel.

We were able to observe the communications in many of these support systems, and analyze how operators responded to questions, and how they handled irate customers. For the most part, victims were upset, realized that the fake AV software was a scam, and requested instructions for removing the malware from their system. The fake AV representatives typically responded with removal directions, but they warned users that their computer was still infected and made claims that competitors (i.e., legitimate antivirus vendors) were slandering their products.

We also performed automated data mining techniques to determine the relationship between complaints, sales, chargebacks, and refunds. To this end, we queried the fake AV groups' internal databases for patterns such as credit card numbers, unique identifiers (e.g., orders), email addresses, and various keywords (e.g., *fraud, scam, refund*, etc) that were relevant to disgruntled customer reactions. By correlating these database records, we examined whether a victim who purchased a fake AV product later filed a complaint through any of the support forums, and if a refund or chargeback was issued. Overall, only a small percentage (less than 10%) of victims actually sought refunds, and those who were issued refunds received their credit within 7 days on average. Note that the low rates of victim complaints that we discovered are similar to those reported by the computer security news investigation web site, KrebsOnSecurity (Krebs, 2010b).

## 2.6 Economic Model

In this section, we utilize the data that we have collected to identify behavior that is representative of a fake AV business. We then propose an economic model based on a key observation of refunds that may be used to detect other businesses that are engaged in illegal activities.

### 2.6.1 Refund Patterns

Fake antivirus software firms (hereafter, firms) act to maximize profits. To do so, the firms rely not only on the systematic transfer of funds to their accounts, but also on a return flow of refunds that mimics the behavior of legitimate providers. As this flow of refunds provides a clear pattern of behavior, we model the refund flow with consideration toward using it to detect and punish firms.

The flow of funds, and refunds, depends on two key players that act as intermediaries between the buyer of the fake software and the firm. As outlined in Figure 2.3, the payment processor is a key player that serves to transmit credit information from the buyer to the credit card network. The second key player is the credit card network, which incorporates both the actual card company (e.g. Visa) and the bank that issues the card (and thereby hosts the buyer's account). The payment flow is from the buyer, through the payment processor and then the credit card network, to the firm.

The trigger for a refund is a request, made by a purchaser, for return of payment upon discovery that the software is fake (or not what they expected). The purchaser may then issue a request for a refund at any point after the sale. To construct a model of requests, we let $s$ denote the number of sales in a given period and let $rq$ denote the number of refund requests that result from $s$. We model requests in period $t$ as a Poisson random variable:

$$rq_t = \lambda s_{t-1},$$

where $\lambda$ captures the expected portion of buyers from period $t-1$ who will issue a request for a refund in period $t$. Given the speed at which information is received and decisions are made, we are primarily concerned with periods corresponding to individual days.

When a refund request has been made, the firm can either ignore the request or grant a refund. If the firm ignores the request, then the buyer may contact the credit card network to obtain a refund. When the credit card network grants a refund to the buyer, the network must collect the funds from the firm by reversing the charge, hence refunds of this type are called chargebacks. This pattern is born out in the data as, for each of the firms under study, the average time to receive a chargeback is substantially longer than the average time to receive a refund (for $AV_1$, chargebacks average 23.7 days longer to process than refunds; the comparable numbers for the other firms are 21.4 days for $AV_2$ and 10.6 days for $AV_3$). For $AV_1$ and $AV_2$, 35-37% of all refunds occur within three days of sales. In contrast, only 1-6% of all chargebacks for $AV_1$ and $AV_2$ occur within three days of sales. For $AV_3$, only 12% of refunds occur within 3 days of sales but less than 1% of chargebacks occur within that same time.

If the firm ceases operations prior to a collection by the payment processor, then the processor must absorb the cost of the chargeback. Because a firm with a large number of sales in a period may decide to cease operations, leaving the processor at risk of absorbing a large number of chargebacks, the payment processor has an incentive to identify illegitimate firms and sever ties with them.

To model the interplay of requests, refunds (which are made directly by the firm to the buyer) and chargebacks, we must specify how payment processors monitor chargebacks to limit their risk. Let $\overline{cb}$ be a threshold, above which the credit card company denies all future transactions. In determining how many requests to refund, a firm that wishes to continue operations must balance the loss in current revenue from granting refunds against the loss of future revenue from being denied access to the credit card network. The number of refunds in a given period, $rf$, is thus an increasing function of the number of requests and a decreasing function of the number of chargebacks, $cb$,

$$rf = g\left(rq, cb\right).$$

Let the threshold $\overline{cb}$ apply to the sum of accumulated chargebacks over $T$ periods. The decision rule of the credit card network is to sever ties with a firm if $\sum_{s=1}^{t} cb_s > \overline{cb}$, for any period $t \in 1, \ldots, T$. As a consequence, a firm will increase the rate of refunds as the sum of accumulated chargebacks approaches the threshold $\overline{cb}$. That is, refunds follow the pattern

$$rf_t = \alpha \cdot rq_t + \beta \cdot rq_t \cdot \left\{ \overline{cb} - \sum_{s=1}^{t} cb_s < D \right\}, \tag{2.1}$$

where $\{A\}$ takes the value 1 if the event $A$ occurs and is 0 otherwise.

The desire to avoid crossing the threshold $\overline{cb}$ leads to a distinctive pattern of refunds and chargebacks. For a payment processor, Equation (2.1) provides several patterns to distinguish these firms from legitimate software providers. For example, refunds from firms may increase at the periodic interval corresponding to $T$ or may increase in reaction to an increase in chargebacks. Also, refunds should increase as the cumulated chargeback sum approaches $\overline{cb}$. For legitimate providers, no such dynamic pattern of refunds should emerge.

To understand the difference in the dynamic refund pattern between legitimate providers and fraudulent firms, note that in contrast to Equation 2.1, refunds for legitimate providers follow the pattern

$$rf_t = \alpha \cdot rq_t \tag{2.2}$$

Because refunds are not a function of chargebacks in Equation 2.2, refunds should depend only on requests for legitimate providers.

To provide evidence that a firm's refunds respond to chargebacks, we display daily refunds and chargebacks for the firms in Figure 2.5. For each of the firms, surges in daily chargebacks are closely followed by (or occur simultaneously with) surges in refunds. The only exceptions appear to be at the latter part of Figure 2.5.

While the figures reveal a dynamic pattern of refunds and chargebacks that is consistent with Equation 2.1, isolating the impact of chargebacks on refunds requires that we control for the level of sales. We must do so because refunds are positively related to sales, so it is possible that sustained increases in sales could lead to increases in both chargebacks and refunds. To estimate the isolated impact of chargebacks, we construct the ordinary least squares estimates of the coefficients in

$$rf_t = \beta_0 + \beta_1 cb_t + \beta_2 cb_{t-1} + \beta_3 \overline{s_t} + u_t. \tag{2.3}$$

The coefficients $\beta_1$ and $\beta_2$ capture the increase in refunds on day $t$ brought about by an increase in chargebacks on day $t$ and day $t-1$, holding previous sales constant. The coefficient $\beta_3$ captures the increase in refunds due to an increase in average sales over the past three days ($\overline{s_t}$). As we do not observe the number of refund requests each day, we use $\overline{s_t}$ as a proxy. The quantity $u_t$ is a random error that encompasses all other factors that influence refunds on that day.

Estimates of Equation 2.3 are contained in Table 1. The column labeled (I) corresponds to Equation 2.3 with $\beta_2 = 0$; that is, lagged chargebacks are not included (these lagged chargebacks are included in Column II). For each of the firms, chargebacks have a substantial impact on refunds after controlling for previous sales. For example, the estimate of 0.64 for firm $AV_1$ indicates that, after controlling for the average level of sales over the previous 3 days, an increase of 100 chargebacks leads to an increase of 64 refunds. In contrast, an increase in average sales of 100 leads to an increase of only 1 refund. The estimated standard errors describe the precision of our estimates: for this coefficient on chargebacks, the confidence interval of (0.16,1.12) indicates the range of plausible values for $\beta_1$. As the interval does not contain 0, the data is strongly supportive of a positive relationship between chargebacks and refunds.

In addition to controlling for sales, we also control for date of the month and day of the week to remove any monthly and daily trends. Column (III) in Table 2.1 corresponds to the coefficient estimates of Equation 2.3 while controlling for monthly and weekly patterns. This was possible with $AV_2$ and $AV_3$ but not for $AV_1$ due to limited data.

Table 2.1: Coefficient estimates for Equation 2.3

| $AV_1$ - **Refunds** | (I) | (II) |
|---|---|---|
| Chargebacks | 0.64 | 0.52 |
| | (0.24)* | (0.24)* |
| Lagged Chargebacks | - | 0.55 |
| | | (0.21)* |
| 3-day Average Sales | 0.008 | 0.009 |
| | (0.008) | (0.008) |

| $AV_2$ - **Refunds** | (I) | (II) | (III) |
|---|---|---|---|
| Chargebacks | 1.23 | 1.16 | 1.17 |
| | (0.14)* | (0.15)* | (0.14)* |
| Lagged Chargebacks | - | 0.26 | 0.25 |
| | | (0.12)* | (0.12)* |
| 3-day Average Sales | 0.043 | 0.041 | 0.041 |
| | (0.004)* | (0.004)* | (0.004)* |

| $AV_3$ - **Refunds** | (I) | (II) | (III) |
|---|---|---|---|
| Chargebacks | 0.72 | 0.71 | 0.72 |
| | (0.24)* | (0.23)* | (0.23)* |
| Lagged Chargebacks | - | 0.089 | 0.088 |
| | | (0.073) | (0.080) |
| 3-day Average Sales | 0.031 | 0.030 | 0.030 |
| | (0.004)* | (0.004)* | (0.004)* |

Note: * indicates significance at the 5% level. Heteroskedasticity-robust standard errors are reported in parenthesis. Our results are not sensitive to the choice of a 3-day average sales window.

Table 1 indicates significant correlation between chargebacks received and refunds granted while controlling for previous sales and monthly fluctuations among all three firms. Without knowing more firm-level details regarding their contracts with payment processors or restrictions from credit card networks further inference becomes difficult. However, we do interpret this as evidence that fraudulent firms seem to alter their refunds according to the chargebacks reported against them. Payment processors or credit card networks have more information and have a better understanding of the firm's chargeback constraints and may, therefore, be in a unique position to monitor these firms.

An important limitation to our analysis is that we lack comparable data for legitimate firms. Despite our findings above, we are unable to discern whether or not this pattern is distinctive to only illegitimate firms.

### 2.6.2 Detecting Fraudulent Firms

The previously described patterns in behavior could be observed by the payment processor since it knows the number of chargebacks against the firm at a particular time, the chargeback threshold faced by the firm, as well as the number of refunds the firm is offering (as these would have to pass through the payment processor). If the payment processor has an incentive to investigate its clients, the existence of this chargeback-responsive behavior could provide evidence that a particular antivirus company is fraudulent. The question is: Does the payment processor have an incentive to investigate its clients?

The payment processor (as noted in Section 2.4.3) receives a percentage of each transaction that occurs but faces a risk of losing business with a credit card company for too much fraudulent behavior. While losing a major credit card company like Visa would devastate a payment processor (as in the case of ePassporte), the credit card company may be hesitant to drop a payment processor if it does enough legitimate business (as in the case of Chronopay).

However, at any given time there is a risk that the fraudulent antivirus firm may be caught or may cease operations. In this case the firm will no longer be able to offer refunds and the payment processor will receive an increase in chargebacks from consumers who have no other way of receiving a refund. The payment processor would be forced to pay the entire amount of the chargeback (the chargeback fees as well as the entire refund amount) as it can no longer bill the firm. Depending on the volume of sales, the risk of future increases in chargebacks could be very costly. If this risk outweighs the revenue the payment processor receives from the firm's account, it may prefer to sever ties with the firm as to not be held liable for the potential chargebacks.

In the case when the firm is caught, credit card companies would have to pay the costs of the chargebacks if the payment processor is forced to shut down. The credit card companies may, therefore, be concerned if a small payment processor is serving an illegitimate firm that may be relatively large compared to the processor's overall volume. In these cases, credit card companies may have an incentive to investigate these firms if they are working with small payment processors. While the credit card company may not observe as much firm level information as the payment processor, it observes the chargebacks and refunds associated with a particular firm. Therefore, this could be a good technique for a credit card company to investigate fraudulent firms.

As mentioned above, we expect the rate of refunds offered by a fraudulent firm to vary in response to chargebacks incurred by the firm. As firms increase their sales, payment processors and credit card networks face increased risk of liability for future chargebacks if the firm ceases operations. This risk may warrant investigation of fraudulent firms using these observable patterns.

## 2.7 Ethical Considerations

The nature of the data that we collected raises a number of ethical concerns. In particular, we have a large amount of personal information for the victims who were defrauded by these three fake AV businesses. Thus, we took measures to protect the privacy and identity of the victims through the use of data encryption, automated program analysis, and by conducting our research according to established ethical principles in the field (Burstein, 2008; Dittrich et al., 2009; Garfinkel, 2008; Kenneally et al., 2010). We also obtained approval from the Institutional Review Board (IRB) at the University of California, Santa Barbara before performing our analysis. Finally, we provided all information that we obtained to U.S. law enforcement officials.

## 2.8 Related Work

In the past few years, there have been several studies that have analyzed various aspects of fraudulent businesses selling fake antivirus products. Researchers from Google described the techniques and dynamics used by cybercriminals to drive traffic to their sites via landing pages (Rajab et al., 2010). Other work analyzed the distribution and installation methods of rogue security software (Fossi et al., 2009). Various security vendors have reported on potential revenue from scareware operations based on the number of infections that they observed (Correll and Corrons, 2010; Micro, 2009). Cova et al. (2010) presented an analysis of the rogue antivirus structure and indirectly tried to measure the number of victims and profits based on poorly configured web servers used by several fake AV groups. They estimated the conversion rate of infections to sales at 1.36%, which is slightly lower than the rates that we observed. We also found a similar geographic distribution of victims in the U.S., and number of domains registered by larger fake AV groups. In comparison, our data provides a much more complete view of large-scale fake AV operations, with information dating back more than two years. We also had visibility of refunds and chargebacks from fake AV sales, which has never been studied before.

Techniques to identify drive-by-download attacks have been proposed that analyze web sites for malicious content in a virtual or emulated environment to detect exploits (Cova et al., 2010; Ikinci et al., 2008). The prevalence of malicious web sites has been examined through crawler-based approaches that analyzed billions of web pages (Mavrommatis and Monrose, 2008; Provos et al., 2007). Another study analyzed drive-by attacks via infiltration and provided insights into the compromised web servers used in the attacks as well as the security posture of potential victims (Stone-Gross et al., 2011a).

A number of recent papers have analyzed the reasons that cause users to fall victim to phishing scams, which include lack of knowledge and attentiveness to browser and other security related cues (Dhamija et al., 2006; Egelman et al., 2008). Several approaches have been proposed to detect phishing sites such as analyzing page content, layout, and other anomalies (Ludl et al., 2007; Pan and Ding, 2006; Rosiello et al., 2007). In addition, studies have analyzed the modus operandi of the criminal operations behind phishing (McGrath and Gupta, 2008), and the effectiveness of phishing defenses (Moore and Clayton, 2007).

Previous work has investigated the Internet's underground economy, through advertised prices of web forums (Zhuge et al., 2009) and IRC chat rooms (Franklin et al., 2007). Holz et al. (2009) studied the drop zones used by botnets to store stolen information from victims. Stone-Gross et al. (2009a) hijacked the Torpig botnet and studied the data exfiltrated from infected computers, and estimated the value of the compromised financial information (e.g., credit card numbers and bank account credentials). The underground economy of large-scale spam operations was examined in (Stone-Gross et al., 2011b). The paper analyzed the complexity in orchestrating spam campaigns, and explored an underground forum used by spammers to exchange goods and services. Another type of scam, known as *One Click Fraud*, was studied by Christin et al. (2010). The fraud works through intimidation (similar to fake AV) by threatening unsuspecting web site visitors with potential embarrassment (e.g., the victim was browsing pornographic content) unless a payment is received for a nonexistent service. The authors presented an economic model to determine the number of users that must fall victim to the scam in order to remain economically viable, and estimated losses in the tens to hundreds of thousands of U.S. dollars.

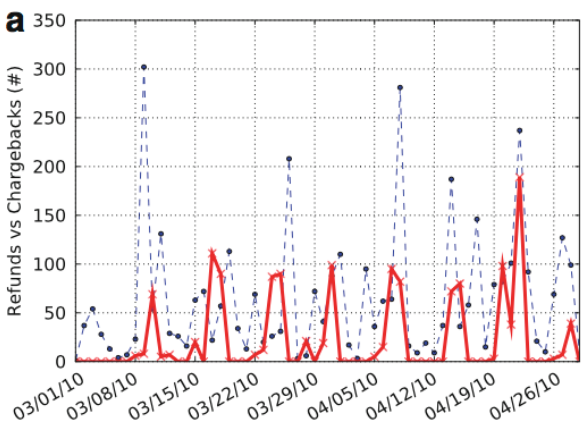
## 2.9 Conclusions

In this paper, we have presented an in-depth study of how a particular type of scareware, namely fake anti-virus software, is deployed and managed. Our work is unique in that it is based on the information contained on a number of key servers that were part of the criminals' infrastructure. This unprecedented access allowed us to obtain ground truth about the type and sophistication of the techniques used to lure victims into paying for scareware, as well as the amount of transactions performed, including refunds and chargebacks.

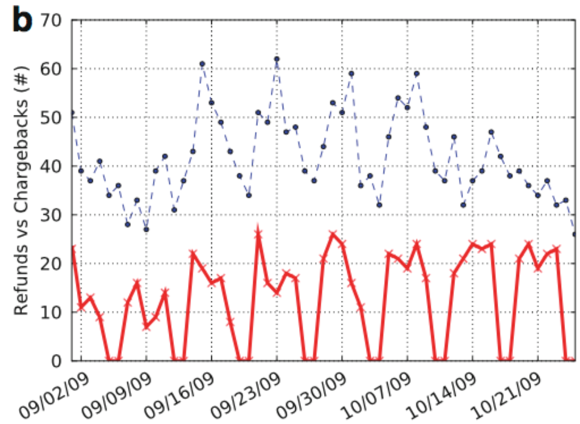
We leveraged this data to build an economic model that shows how cybercriminals are very careful in performing refunds and chargebacks in order to maintain a balanced financial posture that does not immediately reveal their criminal nature. Nonetheless, the economic model also outlines how these operations have distinct characteristics that may differentiate these criminal endeavors from legitimate business operations.

Future work will extend the current model with detection capabilities that can be directly applied to payment data streams. The goal is to develop a tool based on the model that can identify scareware operations automatically.

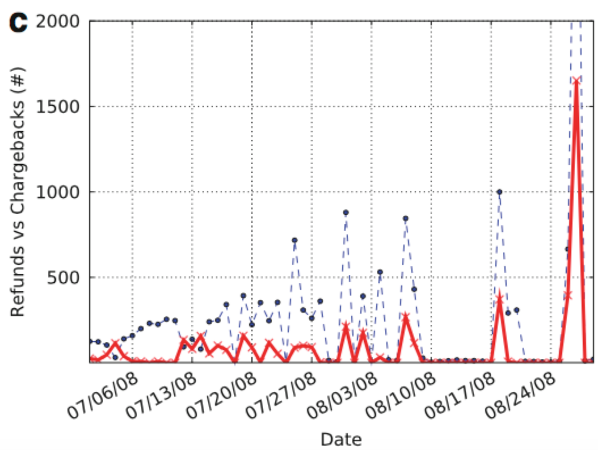
Figure 2.5: Daily refunds and chargebacks from fake AV sales.




Fake $AV_1$ refunds vs chargebacks.




Fake $AV_2$ refunds vs chargebacks.




Fake $AV_3$ refunds vs chargebacks.

The dashed line displays the number of refunds per day, while the solid line displays the number of chargebacks per day.

# Chapter 3

# Deforestation in Malawi: The Role of Agricultural Subsidies and Ethnic Favoritism

## 3.1 Introduction

More than one third of the rural population in developing countries is located on land with limited potential for agriculture (Barbier and Hochard, 2014). Programs aimed to improve food security and reduce poverty have attempted to improve land productivity through various means, yet the environmental consequences of such programs are ambiguous ex ante. Improvement of land productivity may increase the marginal return to clearing forest land for agriculture and thus increase deforestation. However, in a setting of subsistence farming on unproductive land, increased agricultural productivity may reduce or delay the need to shift cultivation, thus slowing deforestation.

Understanding the impact of increased land productivity on deforestation is an empirical challenge, as it requires spatial and temporal variation in the availability of input subsidies that are not strictly correlated with existing agricultural practices. In this paper, we overcome this obstacle by examining the case of a large-scale fertilizer subsidy program in Malawi and leveraging ethnic favoritism in government resource allocation. Exploiting a change in the ethnicity of the Malawi president following the 2004 election, we demonstrate that coethnic households are 10 - 12% more likely to have access to fertilizer subsidies than households of a different ethnicity. Local jurisdictions with majority or plurality populations of the same ethnicity as the president experience much less deforestation than others. Using district-level data on the quantity of fertilizer subsidies, we estimate the elasticity of deforestation with respect to fertilizer subsidies and find it to be consistently negative. Estimates instrumenting the provision of fertilizer subsidies with the coethnic relationship yield elasticities with larger magnitudes in nearly all specifications.

Our results suggest there exist significant environmental spillovers from development policies. Limiting deforestation has been a goal of development agencies for a number of years and has garnered increased interest as global climate change has become more salient. The United Nation‘s Millennium Development Goals stress the importance of slowing deforestation, especially in the developing world. One challenge to aid agencies and governments alike is that measures to slow deforestation are often seen to be at odds with poverty alleviation. Policies that can improve incomes and slow deforestation offer a coveted 'win-win' path to achieving these important ends.

The existing literature examining the effects of development policies, agricultural subsidies and input subsidies on deforestation have yielded mixed results. Some smaller case studies, such as Chibwana et al. (2012), have found that input subsidies, including fertilizer provision, can in fact slow deforestation by incentivizing farmers to switch from slash-and-burn agriculture to more intensive crops. Bulte et al. (2007) study agricultural subsidies in Latin America and argue that these subsidies encourage landowners to put more land into agriculture, thus increasing deforestation. Alix-Garcia et al. (2013) find evidence that poverty alleviation may exacerbate deforestation by increasing local demand for products that require the clearing of land.

Our paper differs from Bulte et al. (2007) and Alix-Garcia et al. (2013) in that it examines targeted input subsidies rather than acerage-based subsidies or income transfers. The setting also differs as Malawi (similar to much of sub-Saharan Africa) suffers from low land productivity in agriculture and the majority of those working in agriculture are subsistence-level farmers. Therefore, targeted input subsidies have real potential to improve yields and slow deforestation from shifting cultivation in Africa.

Our paper also relates to the existing literature on ethnic favoritism and resource allocation. While other papers are discussed below, Morjaria (2014) studies deforestation in Kenya during the democratic transition and finds districts of the same ethnicity as the president experience more deforestation relative to others. The mechanism in his paper is the allocation of permits to clear forested land, which explains the opposite relationship of deforestation and ethnic alignment with the president than that which we find in our paper.

The paper proceeds as follows. Section 3.2 provides important background on Malawi, including a discussion on the role ethnicity plays in resource allocation, the fertilizer subsidy program as well as agriculture and deforestation in this setting. In Section 3.3 we provide empirical evidence for our story. We demonstrate that land quality is negatively related to deforestation, that households ethnically aligned with the president are significantly more likely to receive fertilizer subsidies and that districts aligned with the president receive larger quantities. We demonstrate ethnic alignment of traditional authorities (a local jurisdictional unit) is negatively correlated with deforestation and we estimate district-level elasticities of deforestation and fertilizer subsidies via both ordinary-least-squares and instrumental variables and find them to be negative. Section

3.4 concludes with final remarks.

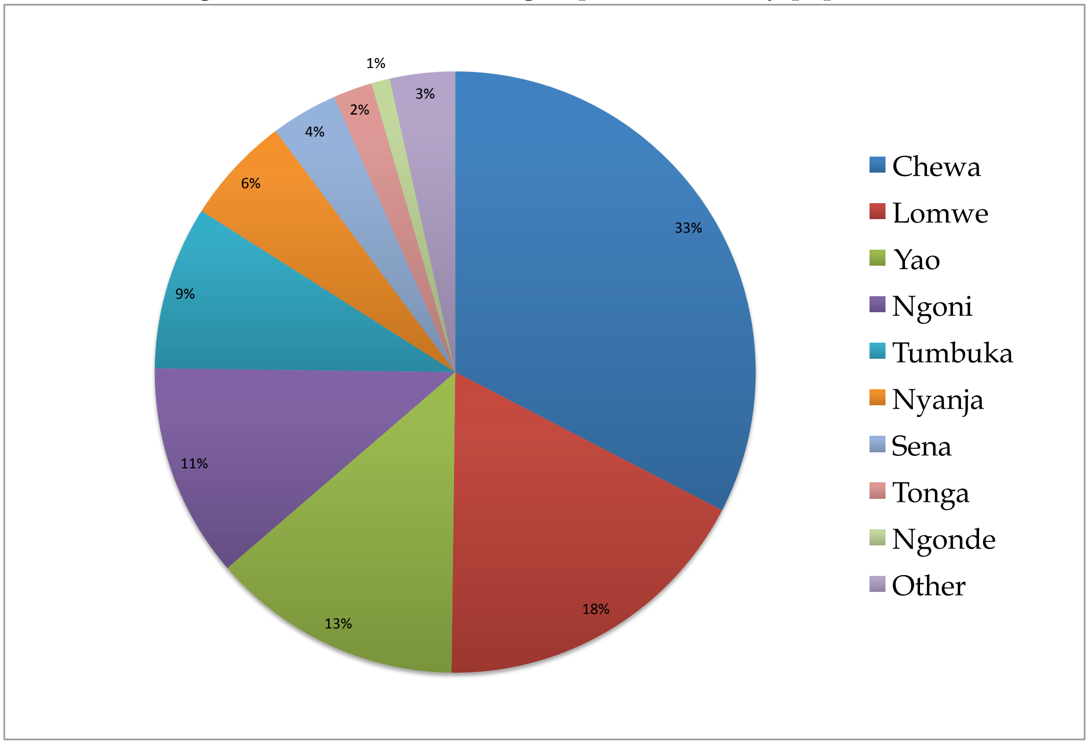
## 3.2 Ethnic favoritism, Deforestation and Fertilizer Subsidies in Malawi

Malawi is an extremely ethnically diverse nation, comprised of 9 major tribal groups: Chewa, Lomwe, Yao, Ngoni, Tumbuka, Nyanja, Sena, Tonga and Ngonde. The three dominant ethnic groups (Chewa, Lomwe and the Yao) account for approximately two-thirds of Malawi's total population (See Figure 3.1). Each group has their own language and traditions which continue to play an important role in Malawian society and politics. Existing geographic concentrations of ethnic groups prior to the drawing of national borders largely explain the spatial distribution of ethnicity observed today. Figure 3.2 presents the the dominant ethnic group in each traditional authority throughout Malawi.

The fact that there are few dominant ethnic groups has led to tensions within the political system. Posner (2004) finds ethnic tensions among the Tumbuka and Chewa much stronger in Malawi than in Zambia, where there are more than seventy ethnic groups, due to the greater relative political presence of these groups in Malawi. There is also limited trust of individuals from other ethnicities in Malawi. Robinson (2013) finds that ethnic divisions in Malawi also lead to economic fragmentation and limit price dispersion, arguing that many of the markets are trust-oriented. Given the strong role ethnic group identity plays in markets and politics, Malawi is susceptible to ethnic favoritism in public good provision.

Kramon and Posner (2013) study the role of ethnic favoritism throughout Africa and, while findings vary from country to country, they find evidence that, in Malawi, those sharing the ethnicity of the president in power benefit from improved infant care and

Figure 3.1: Share of ethnic groups in Malawi by population

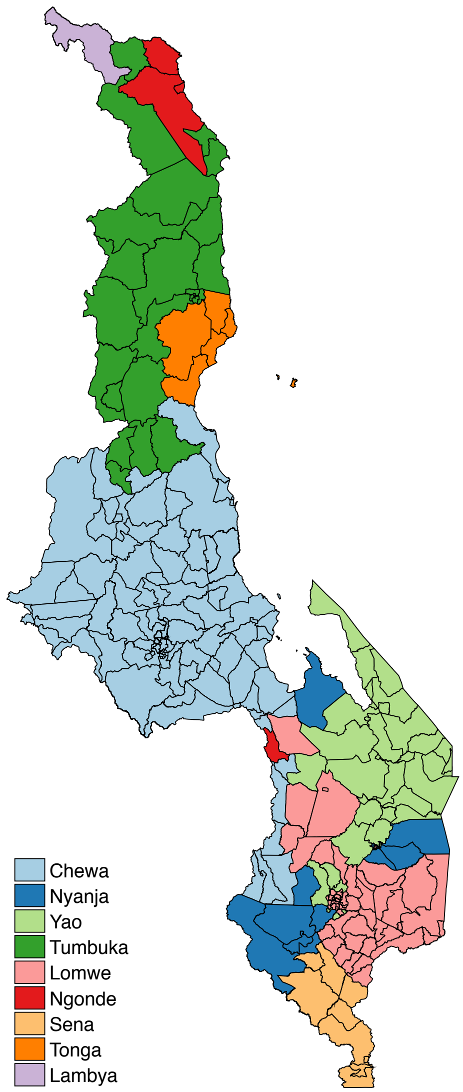



educational opportunities. Furthermore, in a new working paper, Ejdemyr et al. (2015) examine ethnic favoritism and targeting of distributive goods in Malawi. In their cross-sectional data, the authors find that there is a significant correlation between receiving a subsidy and ethnic alignment with the local politician. Additionally, the results suggest that local politicians in areas with more segregation are able to target their co-ethnics with public goods while politicians in less segregated areas rely on the use of subsidies, or private goods, to reward co-ethnics.

In this paper, we exploit the 2004 election in which the country saw a change in the ethnicity of the president from Bakili Muluzi, a member of the Yao tribe, to Bingu wa Mutharika, a member of the Lomwe tribe.[1] The Yao and Lomwe ethnic groups are both concentrated in Southeast Malawi and hold roughly equal population shares in the coun-

[1]This is the same change used by Kramon and Posner (2013) to study ethnic favoritism.

Figure 3.2: Traditional Authorities by Largest Ethnic Group



try.[2] In 2004, the Yao tribe lost a connection to the presidency, while the Lomwe gained it. We argue that this change led to important differences in the allocation of fertilizer subsidies, as this resource was targeted to coethnics and withheld from opposition ethnic groups.

### 3.2.1 Fertilizer and Seed Subsidies in Malawi

Malawi has historically been vulnerable to food insecurity due to its population density, limited access to resources and dependence on an agricultural sector centered on rainfed maize. To help increase food security, the government introduced a fertilizer subsidy program during the 1999/2000 season called the Starter Pack that targeted 2.8 million farm households, providing them enough seeds and fertilizer to cultivate about 0.1 hectares of their staple crop maize (Pauw and Thurlow, 2014). Cost concerns forced the government to revamp the program the next year and, under the new name Targeted Input Program (TIP), they cut the number of beneficiaries in half.

Following a food crisis in 2001/2002, the government again increased access to the subsidy program. The current program, the Farm Input Subsidy Program (FISP), was established during the 2005/2006 agricultural season and targets 1.5 million smallholder farmers throughout Malawi (Pauw and Thurlow, 2014). The Malawi agricultural subsidy program is one of the most important subsidy programs in Sub-Saharan Africa and provides an opportunity to explore the impacts of increased fertilizer and seed access on deforestation both because of its size and also its allocation.

Allocation of the fertilizer subsidies, distributed as vouchers, has always lacked transparency. While a targeted program, there was no defined criteria for defining which households qualified until 2007/2008.[3] The disbursement of fertilizer vouchers is central-

[2]See Figures 3.1 and 3.2.

[3]According to Dorward et al. (2008) prior to 2007/2008, eligible households were those who could not

ized with the Ministry of Agriculture and Food Security, a position held by the president, allocating vouchers to the district-level and, in some seasons, allocating them at a locality-level. Westberg et al. (2015), in a recent working paper, finds evidence of a politically motivated allocation of vouchers whereby they are directed towards swing voters and coethnic districts and away from those of the opposition prior to the 2009 election.[4]

### 3.2.2 Deforestation in Malawi

Deforestation in Malawi, where the majority of the population are subsistence farmers, is caused by slash and burn agricultural techniques. Malawi farmers face declining soil arability due to farming techniques and low land productivity and therefore slash and burn forested land to access more fertile soil. This agricultural practice has resulted in increased levels of deforestation in Malawi compared to the rest of Sub-Saharan Africa (UNFAO, 2010).

Economic models of shifting cultivation, such as Balsdon (2007), predict that improving soil quality and/or land productivity can increase the time a given field is cultivated before shifting to another. In this context, fertilizer should slow deforestation. However, economic models of land clearing for agriculture, such as Angelsen (1999) - among many others, predict that increasing the returns to agriculture should also increase the amount of land cleared for that purpose. These models would predict increased deforestation as a result of fertilizer subsidies. The effect of these subsidies on deforestation remains an empirical question.

There are a few papers that have found that agricultural subsidies lead to decreases in deforestation, but they are mostly small case studies. The article most relevant to the Malawi setting, Chibwana et al. (2012), finds that agricultural households in Malawi

afford one or two bags of fertilizer at current market prices and this was determined by local leaders.

[4]Due to strong ethnic identity and political overlap, the opposition districts are largely of the Yao ethnicity.

that received agricultural subsidies through FISP cleared less forest when compared to households that did not receive a subsidy. While this is only a small case study (N=380) that covers two districts (Kasunga and Machinga), the finding suggests that Malawian farmers moved away from slash and burn agriculture and focus on crop intensification of maize.

## 3.3 Empirical Analysis

As discussed above, the relationship between agricultural input subsidies and deforestation is theoretically ambiguous and depends on the constraints faced by the agricultural households of interest. Therefore, development programs, such as fertilizer subsidies, designed to improve agricultural productivity could increase or decrease levels of deforestation by raising the marginal productivity of agricultural land. Deforestation would increase if the fertilizer subsidies increased the demand for land clearing through increased productivity. However, in Malawi, where most farmers are subsistence farming on unproductive land, increasing agricultural productivity through fertilizer subsidies may reduce the need to shift cultivation to maintain the desired yields resulting in decreased levels of deforestation. Using data on subsidies, deforestation and ethnicity, we can empirically estimate the underlying relationship between fertilizer subsidies and levels of deforestation.

### 3.3.1 Soil Quality

The theory that increased productivity of agricultural land lowers deforestation assumes that, in our setting, agricultural households clear less land when their soil quality is higher. Agricultural inputs, such as fertilizer, play an important role in the overall productivity and quality of soil. The correct use of fertilizer has been shown to both increase yields and improve overall soil quality. Therefore, if our above assumption holds true, we would expect areas with lower soil quality to have higher levels of deforestation and areas with increased access to agricultural inputs to have lower levels of deforestation. We can test the first hypothesis about soil quality and deforestation by comparing levels of deforestation across areas with varying soil quality in Malawi.

The data for this analysis comes from two different sources. The first dataset is compiled by the United Nations Food and Agricultural Organization (FAO) constraints index. The index measures soil quality on a scale from 1-7, where 1 represents high levels of soil quality and 7 represents soil unsuitable for agriculture. We assign a score to each Traditional Authority (TA) in Malawi by taking the area weighted average of soil quality within the TA. The second data set is the deforestation data that comes from recently released data from researchers at the University of Maryland that provides estimates of forest cover for the entire terrestrial surface of the earth at a 30m×30m resolution (Hansen et al., 2013). Included in the dataset are estimates of the percentage of each 30m×30m grid cell in forest cover in the year 2000.[5] The dataset also provides annual indicators from 2001 to 2012 denoting that a grid cell containing nonzero tree cover in 2000 is estimated to have fallen to zero percent tree cover. For our deforestation measure, we limit the sample to cells that had at least 30% forest cover at baseline to reduce the noise in our deforestation measure. From those cells, we count the number of pixels deforested in each year within each traditional authority, thus creating a panel of annual deforestation at the traditional authority level from 2001-2012.[6]

To test the assumption that more deforestation occurs in areas with lower soil quality

[5] Forest cover is defined as area covered by vegetation greater than 5 meters in height.

[6] It should be noted that this definition of deforestation is not universally accepted. Tropek et al. (2014) point out that classifying forest as vegetation taller than 5 meters can lead to classification of different plantations as forest. Harvesting of these plantations may result in observed deforestation when, in reality, the land had been cleared and planted prior to the beginning of the study period.

we estimate the following equation:

$$y_{id} = \beta_0 + \beta_1 Soil_{id} + \beta_3 X_i + \gamma_d + \varepsilon_{id} \tag{3.1}$$

where $y_{id}$ is the average level of deforestation measured in traditional authority $i$ in district $d$ from 2001 to 2012. $Soil_{id}$ is an index of soil quality measured at the traditional authority level, $X_i$ are additional cross sectional controls at the traditional authority level, including population, population growth, electrification, average household size, fuel wood use, and percent of the population participating in agriculture, $\gamma_d$ is a district fixed effect, and $\varepsilon_{id}$ is the traditional authority error term.

Table 3.1 shows the regression results of soil quality on deforestation. Columns (1) through (3) use the level of deforestation as the outcome variable (as measured in number of pixels) and the results provide evidence that lower soil quality leads to significantly higher levels of deforestation. The significant and positive estimates can be interpreted as an increase in the number of soil constraints constraints to agriculture (or a decrease in soil quality) at the traditional authority level is correlated with higher levels of deforestation. The result holds up to the addition district fixed effects, but becomes insignificant with the addition of cross-sectional control variables. Instead using the natural log of deforestation as the outcome variable, the results in columns (4) through (6) again show that lower levels of soil quality are correlated with increased deforestation and this time the result is robust to the inclusion of year fixed effects, as well as additional controls. The results in Table 1 support the assumption that deforestation and agricultural productivity are negatively correlated in the case under study. Throughout our analysis, we prefer the log of deforestation as the outcome variable because it accounts for different levels of forest cover at baseline and also provides intuitive interpretations of regression coefficients.

Table 3.1: Soil Quality and Deforestation 2001-2012

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Mean Deforestation Pixels | | | Log of Mean Deforestation | | |
| Soil Constraints | 155.2*** | 209.7** | 130.2 | 0.424*** | 0.582*** | 0.259** |
| | (40.09) | (93.91) | (80.33) | (0.107) | (0.169) | (0.117) |
| District FE | NO | YES | YES | NO | YES | YES |
| Controls | NO | NO | YES | NO | NO | YES |
| Observations | 223 | 223 | 223 | 223 | 223 | 223 |
| R-squared | 0.020 | 0.386 | 0.434 | 0.067 | 0.509 | 0.745 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 3.3.2 Ethnic Favoritism and Fertilizer Subsidies

In order to identify the effect of fertilizer subsidies on deforestation in Malawi, we need to overcome the endogeneity of subsidy allocations. If all households growing the same crop with the same technology received the same amount of fertilizer subsidies at the same time, it would be difficult to empirically disentangle the effect of the subsidies on deforestation from other systematic changes such as variation in market price of maize, technology adoption, etc. Furthermore, if fertilizer subsidies were allocated based on some unobservable criteria that also effects the demand or cost of clearing forest (for example more productive households are able to get the subsidy easier than less productive ones) simple estimates of the relationship between subsidies and deforestation will be biased. To overcome this problem, we need a time-varying allocation criteria that is orthogonal to changes in unobservable criteria mentioned above.

We use a similar strategy as Kramon and Posner (2013), exploiting both the role of ethnic favoritism in government resource allocation and the change in the ethnicity of the Malawi president following the 2004 election. If there exist two households identical but for their ethnicity and the outcome of the 2004 election leads to the alignment in ethnicity of one household but not the other, in the presence of ethnic favoritism in subsidy allocation, this change will allow us to estimate the relationship between fertilizer subsidies and deforestation. For this to be a valid strategy, we need to show that coethnics before the election, members of or districts aligned with the Yao tribe, received more fertilizer subsidies compared to members of or districts aligned with other tribes prior to the 2004 election. Additionally, we need to show that coethnics after the election, members of or districts aligned with the Lomwe tribe, received more fertilizer subsidies compared to members of or districts aligned with other tribes following the 2004 election.

The data we use for this analysis comes from multiple sources. The first dataset contains information on use of agricultural subsidies at the household level and comes from two waves of the Malawi Integrated Household Survey (IHS). The IHS is part of the Living Standards Measurement Study (LSMS), which is a household survey program, run by the World Bank's Development Research Group, in partnership with national statistical offices (NSOs). The IHS waves in Malawi collected information on poverty and income equality, demographic characteristics, health, education, and agriculture. The IHS data used in the analysis comes from the IHS second and third waves for which data was collected in 2004 and 2010. The surveys collected information from a nationally representative sample. The sampling design is representative at both national and district level. The surveys ask households about agricultural subsidies received during the years 2001, 2002, 2003 and 2009. Additionally, they ask about the tribal language spoken within the household, which allows for the creation of a short household pseudo-panel with household characteristics, household ethnicity and an indicator variable for whether the household received an agricultural subsidy in a given year. Finally, indicator variables are created for both Yao and Lomwe households, as well as when a household is ethnically aligned with the president in power.

The second source of fertilizer data was provided by the International Food Policy Research Institute (IFPRI) and contains data on fertilizer subsidy allocations by district and year for 2001 through 2012 with the exception of 2006. The data reports the amount of fertilizer subsidies distributed to each district in kilograms. There are 24 districts over 11 years, which results in 264 observations for this part of the analysis.

The last dataset is the 2008 Malawi Census.[7] The census surveyed 298,607 households and 1,343,078 individuals (10% of the total population) and the ethnicity question allows us to measure tribal populations at the traditional authority level. Figure 3.3 illustrates the tribal distribution found in the Malawi census. Additionally, we create variables identifying traditional authorities and districts by the share of the population that belong to the Yao tribe, the share of the population that belong to the Lomwe tribe, the tribe that has a plurality, meaning that the tribe that makes up the largest proportion of the population, and the tribe that has a majority, meaning that the tribe that makes up over 50% of the population. Figures 3.2 and 3.3 show the ethnic plurality of each traditional authority and the share of population belonging to the Yao tribe and Lomwe tribe respectively. These maps lay out the spatial variation in ethnicity, which we rely on for identification.

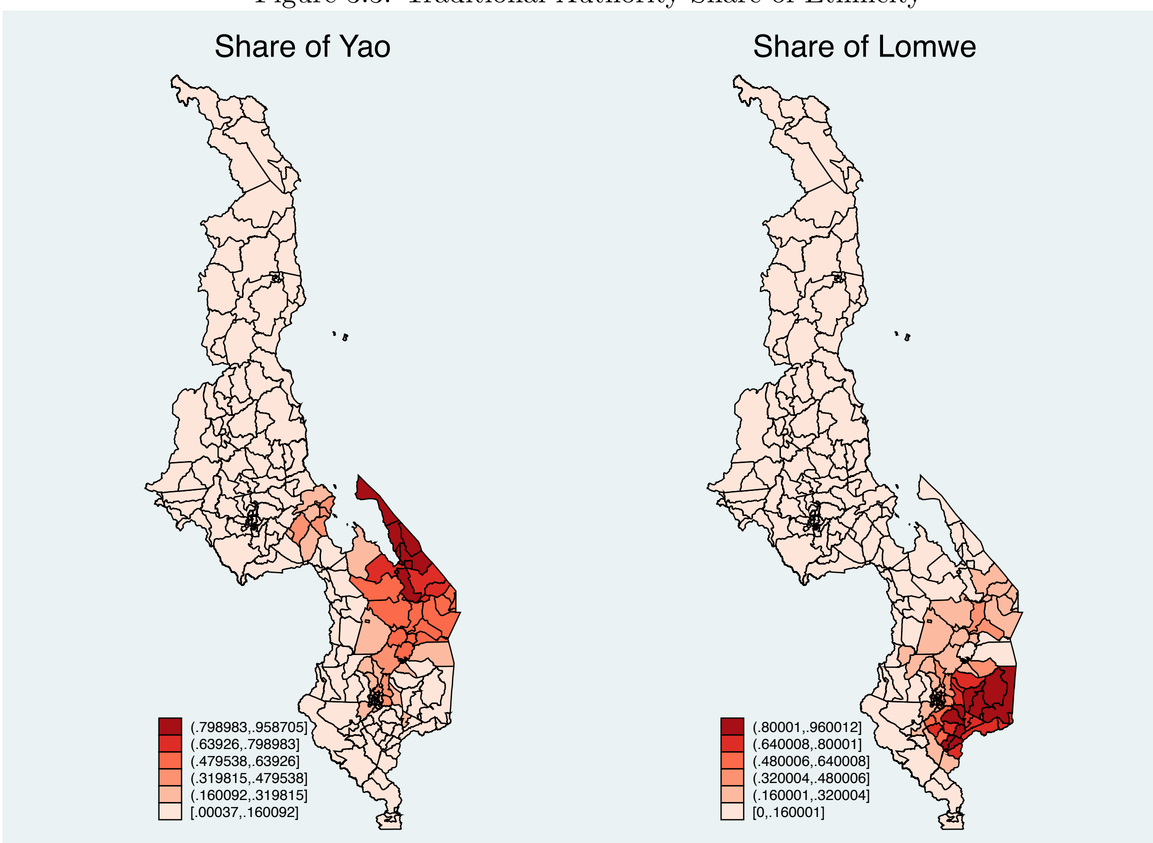
Using the psudeo-panel created by combining the IHS surveys, we can estimate the following model to test for ethnic favoritism in the allocation of fertilizer subsidies at the household level:

$$y_{jdt} = \beta_0 + \beta_1 Aligned_{jdt} + \gamma_d + \phi_t + \varepsilon_{jdt} \tag{3.2}$$

where $y_{jdt}$ is an indicator for whether household $j$ in district $d$ received an agricultural subsidy in year $t$. $Aligned_{jdt}$ is an indicator for whether a household is ethnically aligned

[7] We accessed the 2008 Malawi Census data via IPUMS.

Figure 3.3: Traditional Authority Share of Ethnicity



with the president (Yao before 2004 and Lomwe after 2004), $\gamma_j$ is a district fixed effect, $\phi_t$ is a year fixed effect and, finally, $\varepsilon_{jdt}$ is the household-year error term.

The coefficient of interest is $\beta_1$. If ethnic alignment influences subsidy allocation, then we would expect $\beta_1$ to be strictly positive. The results in Table 3.2 columns (1) to (3) show that ethnic alignment with the president leads to a significant increase in the probability of receiving a fertilizer subsidy by between 10 and 13 percentage points. The result is robust to the addition of year fixed effects and district fixed effects. This analysis suggests that ethnic favoritism does indeed impact government allocations of fertilizer subsidies at the household level.

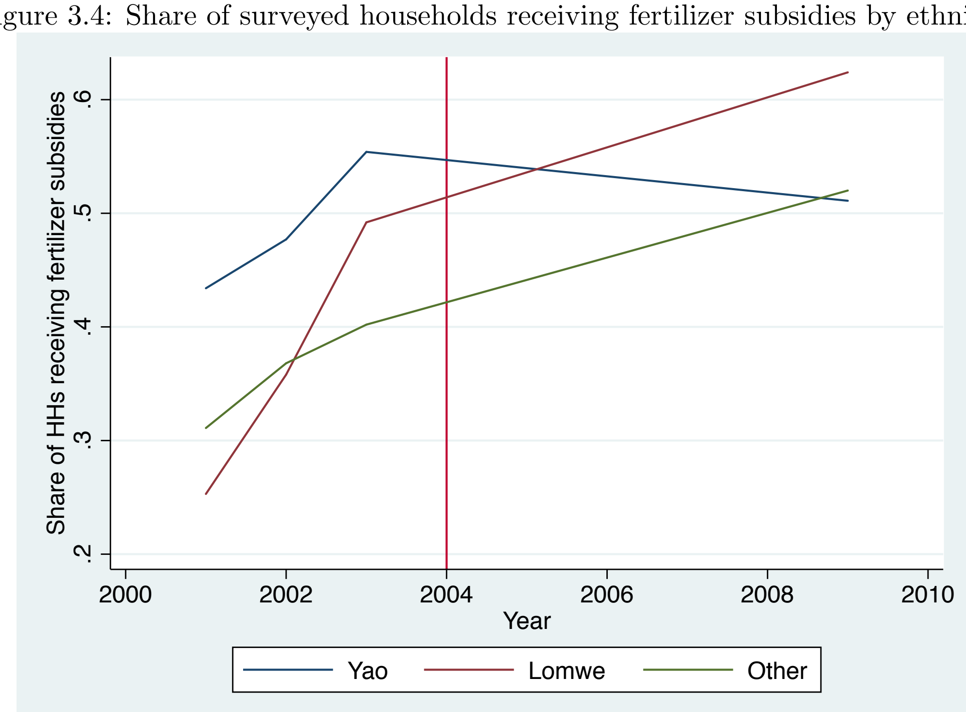
Additionally, we can explore the same question about ethnic favoritism and fertilizer subsidies using indicators for Yao households, Lomwe households, Post-2004, and the interactions between them. We estimate the following model:

$$y_{jdt} = \beta_0 + \beta_1 Yao_{jd} + \beta_2 Lomwe_{jd} + \beta_3 Post_t + \beta_4 Yao_{jd} \times Post_t + \beta_4 Lomwe_{jd} \times Post_t + \gamma_d + \varepsilon_{jdt} \tag{3.3}$$

where $y_{jdt}$ is as noted above, $Yao_{jd}$ is an indicator for whether a household belongs to the Yao tribe (the president‘s tribe before the 2004 election), $Lomwe_{jd}$ is an indicator for whether a household belongs to the Lomwe tribe (the president‘s tribe after the 2004 election), $Post_t$ is an indicator for after the election (this variable is dropped when year fixed effects are included), $\gamma_d$ is a district fixed effect and $\varepsilon_{jdt}$ is the household-year error term. The omitted reference group is all households of an ethnicity other than Yao or Lomwe.

The results from estimating the above equation are reported in Table 3.2 columns (4) through (6). Again, we can see that ethnic alignment with the president prior to the 2004 election led to an increase in the probability of receiving a fertilizer subsidy by between 13 and 14 percentage points for Yao households. Furthermore, Lomwe households had the same probability of receiving a subsidy as other ethnicities prior to 2004, but had an increase probability of receiving a subsidy between 7 and 11 percentage points after 2004. These results are robust to the addition of year fixed effects and district fixed effects. Additionally, Figure 3.4 provides a graphical presentation of ethnic favoritism in the allocation of fertilizer subsidies at the household level. The graph shows the share of surveyed households receiving a fertilizer subsidy by ethnicity and is consistent with Yao households receiving more subsidies prior to 2004 and Lomwe households receiving more subsidies following the 2004 election. These results provide evidence that ethnic favoritism plays a role in household-level fertilizer allocation.

Figure 3.4: Share of surveyed households receiving fertilizer subsidies by ethnicity




The above results on ethnic favoritism and the allocation of fertilizer subsidies at the household level are useful to demonstrate the relationship between the two, however, we do not observe deforestation at the household level. Therefore, we need to use aggregated data of fertilizer subsidies and show that ethnic favoritism also plays a role in how fertilizer

Table 3.2: Fertilizer Subsidies and Household Ethnicity

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Household Ethnically Aligned | 0.106*** | 0.125*** | 0.099** | | | |
| | (0.0084) | (0.027) | (0.038) | | | |
| Lomwe Household | | | | -0.0028 | -0.0028 | 0.0108 |
| | | | | (0.0123) | (0.0121) | (0.0134) |
| Lomwe x Post 2004 | | | | 0.107*** | 0.107*** | 0.0722*** |
| | | | | (0.0258) | (0.0257) | (0.0262) |
| Yao Household | | | | 0.128*** | 0.128*** | 0.0967*** |
| | | | | (0.009) | (0.00894) | (0.0121) |
| Yao x Post 2004 | | | | -0.137*** | -0.137*** | -0.134*** |
| | | | | (0.0192) | (0.0192) | (0.0194) |
| Post 2004 | | | | 0.160*** | - | - |
| | | | | (0.006) | - | - |
| Year FE | NO | YES | YES | NO | YES | YES |
| District FE | NO | NO | YES | NO | NO | YES |
| Observations | 44,241 | 44,241 | 44,241 | 44,241 | 44,241 | 44,241 |
| R-squared | 0.004 | 0.028 | 0.065 | 0.022 | 0.028 | 0.065 |

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

subsidies are allocated at the district level.[8] We estimate the following equation using the IFPRI fertilizer subsidy data and the Malawi Census data to examine the relationship between district alignment and fertilizer subsidies:

$$y_{dt} = \beta_0 + \beta_1 Yao_d + \beta_2 Lomwe_d + \beta_3 Post_t + \beta_4 Yao_d \times Post_t + \beta_5 Lomwe_d \times Post_t + \gamma_d + \varepsilon_{dt} \tag{3.4}$$

where $y_{dt}$ is amount of fertilizer subsidy allocated at the district and year level, $Yao_d$ is an indicator for whether a district's ethnic majority is the Yao tribe, $Lomwe_d$ is an indicator for whether a district's ethnic majority is the Lomwe tribe, $Post_t$ is an indicator for after the election (this variable is dropped when year fixed effects are included), $\gamma_d$ is a district fixed effect, $\phi_t$ is a year fixed effect and, finally, $\varepsilon_{dt}$ is the district-year error term. The omitted reference group is all districts of an ethnicity other than Yao or Lomwe.

Table 3.3 clearly shows that the districts where the majority of the population are members of the Yao tribe or Lomwe tribe received significantly more fertilizer subsidies compared to other districts prior to the 2004 election. Following the election, we can see that the Yao districts' fertilizer allocations dropped significantly compared to both the Lomwe districts and other districts.[9] These effects are robust to the addition of year fixed effects and district fixed effects. This result again provides evidence that ethnic favoritism plays an important role in how fertilizer subsidies are allocated in Malawi. Moreover, it seems that the impacts are driven by the significant decrease in fertilizer allocations to Yao districts following the 2004 election.

[8]While we are able to estimate deforestation at the traditional authority level, we only have data on the quantity of fertilizer distributed to the districts, so we must use this level of aggregation.

[9]The sum of the coefficient on $Yao$ and the coefficient on $Yao \times Post$ are effectively zero, which indicates that the Yao group benefitted from additional subsidies prior to the election, but returned to levels comparable to other ethnic groups after the election.

Table 3.3: Quantity of Fertilizer by Ethnic Majority of District

| | (1) | (2) | (3) |
|---|---|---|---|
| Lomwe District | 0.493*** | 0.493*** | 1.143*** |
| | (0.131) | (0.111) | (0.155) |
| Lomwe x Post 2004 | -0.133 | -0.133 | -0.133 |
| | (0.285) | (0.177) | (0.120) |
| Yao District | 0.890*** | 0.890*** | 1.607*** |
| | (0.159) | (0.0911) | (0.137) |
| Yao x Post 2004 | -0.250 | -0.250*** | -0.250*** |
| | (0.35) | (0.146) | (0.116) |
| Post 2004 | 1.893*** | - | - |
| | (0.138) | - | - |
| Year FE | NO | YES | YES |
| District FE | NO | NO | YES |
| Observations | 264 | 264 | 264 |
| R-squared | 0.419 | 0.616 | 0.853 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 3.3.3 Ethnic Alignment and Deforestation

While there seems to be a clear connection between ethnic favoritism and fertilizer subsidy allocations at both the household level and district level, the impacts on deforestation are still ambiguous. Using the panel data on deforestation and the census data on tribal plurality, tribal majority and ethnic share aligned with the president, we can examine the direct impacts of ethnic alignment on deforestation. As we do not observe fertilizer allocation at the traditional authority level, we cannot verify that the mechanism through which deforestation is affected is the fertilizer subsidies. We examine this explicitly in the next section at the district level, but we have more statistical power at the traditional authority level as we have many more observations. The empirical model we use employs a difference-in-differences strategy to identify the impacts. In order to do this, we construct three different measures of ethnic alignment at the traditional authority level. The first two are indicators for ethnic majority and plurality. Prior to 2004, traditional authorities with a majority or plurality of the population in the Yao tribe receive a one and all others a zero. Following 2004, traditional authorities with a majority or plurality of the population in the Lomwe tribe receive a 1one and all others a zero. The third measure uses the share aligned with the president in each traditional authority. We interact a pre-2004 indicator with the share of Yao in each traditional authority and interact a post-2004 indicator with the share of Lomwe in each traditional authority. Figures Figures 3.2 and 3.3 show the geographical variation of the alignment variables.

We estimate the following regression model to examine the direct effects of ethnic alignment with the president on deforestation:

$$y_{idt} = \beta_0 + \beta_1 Aligned_{idt} + \beta_2 X_{id} + \gamma_d + \phi_t + \varepsilon_{idt} \tag{3.5}$$

where $y_{idt}$ is the level or natural log of deforestation in traditional authority $i$ in district $d$ and year $t$. $Aligned_{idt}$ is one of the three measures of alignment with the president discussed previously, $X_{id}$ are additional controls for population and population growth, $\gamma_j$ is a district fixed effect, $\phi_t$ is a year fixed effect and finally $\varepsilon_{idt}$ is the traditional authority-year error term.

The parameter of interest is $\beta_1$, which is similar to a difference-in-difference estimator, because it compares the levels of deforestation of an aligned tribe‘s traditional authorities during periods when they are aligned with the president and when they are not. Therefore, the main threat to identification would be if the new president used his power to target areas of the previous president‘s tribe for deforestation. Given the fertilizer results, there doesn't seem to be any evidence that the new president punished members of the Yao tribe through fertilizer allocations. Moreover, this punishment mechanism is most likely not the case, as the transition following the 2004 election was peaceful and the previous and new presidents were members of the same political party.

The estimation results are displayed in Table 3.4. The top panel uses the level of deforestation as the outcome variable and the bottom panel uses the natural log of deforestation as the outcome variable, which provides more intuitive interpretations of the coefficients. Columns (1) and (2) use the tribal ethnic majority at the traditional authority level to measure alignment. Column (2) also includes a district level fixed effect. Columns (3) and (4) use the tribal ethnic plurality at the traditional authority level to measure alignment. Column (4) also includes a district level fixed effect. Finally, columns (5) and (6) use the share of the ethnically aligned tribe at the traditional authority level to measure alignment. Column (6) also includes a district level fixed effect.

The estimates indicate that traditional authorities with higher populations of co-ethnics aligned with the president experienced significantly less deforestation compared to traditional authorities not aligned with the president. Normalizing the level effects

Table 3.4: Ethnic Alignment and Deforestation in Traditional Authorities

| **Defor Pixels** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Maj Eth Align | -344.2*** | -95.93* | | | | |
| | (54.10) | (53.98) | | | | |
| Plur Eth Align | | | -393.8*** | -89.02** | | |
| | | | (44.74) | (39.22) | | |
| Share Aligned | | | | | -854.8*** | -187.1* |
| | | | | | (99.45) | (101.4) |
| Year FEs | YES | YES | YES | YES | YES | YES |
| District FEs | NO | YES | NO | YES | NO | YES |
| Observations | 2,676 | 2,676 | 2,676 | 2,676 | 2,676 | 2,676 |
| R-squared | 0.058 | 0.307 | 0.061 | 0.307 | 0.068 | 0.307 |

| **Log of Defor** | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|
| Maj Eth Align | -0.130 | -0.0577 | | | | |
| | (0.107) | (0.113) | | | | |
| Plur Eth Align | | | -0.777*** | -0.218** | | |
| | | | (0.0904) | (0.0944) | | |
| Share Aligned | | | | | -1.482*** | -0.616*** |
| | | | | | (0.165) | (0.187) |
| Year FEs | YES | YES | YES | YES | YES | YES |
| District FEs | NO | YES | NO | YES | NO | YES |
| Observations | 2,676 | 2,676 | 2,676 | 2,676 | 2,676 | 2,676 |
| R-squared | 0.183 | 0.567 | 0.199 | 0.568 | 0.205 | 0.569 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Deforestation by Ethnicity of Traditional Authority

| | (1) | (2) |
|---|---|---|
| Majority Yao | -196.7 | |
| | (191.2) | |
| Majority Yao x Post 2004 | 51.16 | |
| | (120.1) | |
| Majority Lomwe | -28.52 | |
| | (59.92) | |
| Majority Lomwe x Post 2004 | -67.55 | |
| | (84.72) | |
| Plurality Yao | | -165.8 |
| | | (133.9) |
| Plurality Yao x Post 2004 | | -2.198 |
| | | (108.5) |
| Plurality Lomwe | | -54.53 |
| | | (61.26) |
| Plurality Lomwe x Post 2004 | | -132.8* |
| | | (76.75) |
| Post 2004 | 178.7*** | 200.6*** |
| | (60.13) | (68.52) |
| Year FEs | NO | NO |
| District FEs | YES | YES |
| Observations | 2,676 | 2,676 |
| R-squared | 0.260 | 0.260 |

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

to hectares, we find that alignment with the president leads to between 8 and 75 fewer hectares being deforested. These effects make sense given that the agricultural subsidy provided to households by the government provides enough seed and fertilizer to farm 0.1 hectares. The bottom panel shows that ethnic alignment with the president decreases deforestation by 28% to 78%. These significant effects across specifications suggests that there is relationship of ethnic alignment on deforestation. However, ethnic alignment in the Malawi context leads to lower levels of deforestation, which goes against the hypothesis that politicians use deforestation as a good to reward co-ethnics or political allies (see Barbier et al. (2005) and Morjaria (2014)).

### 3.3.4 Fertilizer Subsidies and Deforestation

In order to estimate the effect of fertilizer subsidies on deforestation, we can use ethnic alignment with the president as an instrument. Using two-stage least-squares estimation, we estimate the impact of fertilizer subsidies on deforestation using the variation in fertilizer subsidy allocation attributable to ethnic favoritism. We aggregate both deforestation data as well as ethnicity data to the district level given our constraints on fertilizer data. We then estimate the following two-stage least-squares model where the first stage is:

$$Fert_{dt} = \beta_0 + \beta_1 Aligned_{dt} + \beta_3 X_d + \phi_t + \varepsilon_{dt} \tag{3.6}$$

and the second stage is:

$$y_{dt} = \alpha_0 + \alpha_1 \hat{Fert}_{dt} + \alpha_3 X_d + \phi_t + \epsilon_{dt} \tag{3.7}$$

where $Fert_{dt}$ is the natural log of fertilizer measured at the district-year level, $Aligned_{dt}$ is one of the three measures of alignment with the president discussed above, $X_d$ are additional controls for population and population growth, $\phi_t$ is a year fixed effect, $y_{dt}$ is the natural log of deforestation measured at the district and year level, and finally $\varepsilon_{dt}$ and $\epsilon_{dt}$ are the district-year error terms.

Using the variation in fertilizer allocations attributable to ethnic favoritism allows for estimates of $\alpha_1$ that are not confounded by time-invariant unobservable differences in agriculture or otherwise at the district-level. However, our instrument is invalid if alignment with the president also yields non-fertilizer related benefits that could impact forest clearing. For example, if other income transfers allow farmers to work less or increased public sector employment offered to coethnics incentivizes some to leave farming altogether, ethnic alignment would have an effect on deforestation other than through the channel of increased fertilizer subsidies. We believe these concerns are minimal for two reasons. First, the main source of employment in Malawi is overwhelmingly in agriculture. According to a 2008 survey by the National Statistics Office of Malawi, 84% of all workers reported their primary source of employment was in 'agriculture, fishing or forestry'. Furthermore, the magnitude of FISP was large. According to Dorward and Chirwa (2011), during the 2008/2009 growing season, expenditures on the program were approximately 16% of the national budget and 74% of the Ministry of Agriculture and Food Security's budget. These two factors lead us to believe the main effect of ethnic alignment on deforestation is through the agricultural subsidies.

The coefficients from estimating the above IV model are presented in Table 3.6. We use the natural log of deforestation and the natural log fertilizer in order to produce more easily interpretable elasticities. Columns (1) through (3) in Panel A are the results from an OLS regression of fertilizer and deforestation - estimating equation (3.7) replacing observed predicted fertilizer with observed fertilizer. Observed fertilizer is endogeneous, these estimates will be biased though the direction of the bias is theoretically ambiguous. The OLS results indicate a negative relationship between fertilizer and deforestation.

Table 3.6: District-level estimates of the elasticity of fertilizer and deforestation

| **Panel A** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Log Fertilizer | -0.0304 | -0.221** | -0.360** | -0.414 | -1.049 | 9.958 |
| | (0.0554) | (0.0994) | (0.147) | (0.355) | (0.723) | (28.18) |
| Year FEs | NO | YES | YES | NO | YES | YES |
| Controls | NO | NO | YES | NO | NO | YES |
| Observations | 264 | 264 | 264 | 264 | 264 | 264 |
| Instrument | None | None | None | Majority | Majority | Majority |
| IV F-Statistic | NA | NA | NA | 7.878 | 4.606 | 0.128 |
| R-squared | 0.001 | 0.125 | 0.499 | | | |

| **Panel B** | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|
| Log Fertilizer | -0.583** | -1.202** | -0.223 | -0.688* | -1.093** | 22.61 |
| | (0.280) | (0.489) | (1.423) | (0.356) | (0.466) | (291.6) |
| Year FEs | NO | YES | YES | NO | YES | YES |
| Controls | NO | NO | YES | NO | NO | YES |
| Observations | 264 | 264 | 264 | 264 | 264 | 264 |
| Instrument | Plurality | Plurality | Plurality | Share | Share | Share |
| IV F-Statistic | 15.18 | 11.52 | 0.980 | 10.24 | 11.76 | 0.00583 |
| R-squared | | | | | | |

Columns (4) through (6) use the ethnic majority aligned with the president as the instrument, but because we are using district-level data we have too few districts with a strict majority of Yao or Lomwe to provide sufficient power in the first stage regression - as evident by the low F-statistics. Panel B uses the plurality and share alignment variables, which provide more first-stage predictive power. The results in Panel B indicate that fertilizer allocations have a significant impact on deforestation even when aggregated to the district level. The preferred estimates are in columns (7) and (10) and suggest an elasticity of deforestation with respect to fertilizer between $-.59$ and $-.69$. Inclusion of district-level controls in the IV specifications dramatically reduces first-stage explanatory power as both ethnic composition and other controls are time invariant and the district-level aggregation leaves us with relatively few observations.

## 3.4 Final Remarks

In this paper, we provide evidence that ethnic alignment with the president played an important role in the allocation of fertilizer subsidies in Malawi. We further demonstrate that ethnic alignment with the president leads to a significant decrease in deforestation. At the district level, the elasticity between fertilizer subsidies and deforestation is negative and instrumental variable estimates utilizing ethnic favoritism in fertilizer allocation are larger in magnitude than the OLS estimates in most specifications.

These large and statistically significant effects show the importance of measuring the impacts of environmental spillovers of development programs to conduct an efficient cost benefit analysis. Additionally, policymakers should focus on implementing poverty alleviation programs, such as fertilizer subsidies, that have beneficial environmental impacts because they may be able to provide a 'win-win' scenario.

Many questions remain for future work. Given the spillover effect of avoided deforestation from fertilizer subsidies, how should these subsidies be targeted? The optimal allocation with the joint objective of poverty alleviation and environmental benefits may differ from the optimal allocation with the single objective of poverty alleviation. While we leverage ethnic favoritism in the allocation of fertilizer subsidies to examine their relationship to deforestation, it is likely the case that ethnic favoritism may have welfare consequences by diverting fertilizer away from households or districts where it could have a larger impact on both poverty and deforestation. What are the welfare consequences of this pattern of resource allocation? We look forward to pursuing these questions in future work.

# Bibliography


Alix-Garcia, J., McIntosh, C., Sims, K. R., and Welch, J. R. (2013). The ecological footprint of poverty alleviation: evidence from mexico's oportunidades program. *Review of Economics and Statistics*, 95(2):417–435.

Angelsen, A. (1999). Agricultural expansion and deforestation: modelling the impact of population, market forces and property rights. *Journal of development economics*, 58(1):185–218.

Assunção, J., Gandour, C., and Rocha, R. (2012). Deforestation slowdown in the legal amazon: prices or policies. *Climate Policy Initiative (CPI) Working Paper, Pontífica Universidade Católica (PUC), Rio de Janeiro, RJ, Brazil. p*, 3.

Assunção, J., Gandour, C., and Rocha, R. (2013). Deterring deforestation in the brazilian amazon: Environmental monitoring and law enforcement. *Climate Policy Initiative Report, PUC-Rio, May.*

Balakrishnan, A. (2008). Brazilian government faces criminal charges over amazon deforestation. *The Gaurdian*, September 30.

Balsdon, E. M. (2007). Poverty and the management of natural resources: A model of shifting cultivation. *Structural Change and Economic Dynamics*, 18(3):333–347.

Barber, C. P., Cochrane, M. A., Souza Jr, C., and Veríssimo, A. (2012). Dynamic performance assessment of protected areas. *Biological Conservation*, 149(1):6–14.

Barbier, E. B., Damania, R., and Léonard, D. (2005). Corruption, trade and resource conversion. *Journal of Environmental Economics and Management*, 50(2):276–299.

Barbier, E. B. and Hochard, J. P. (2014). Poverty and the spatial distribution of rural population. *World Bank Policy Research Working Paper*, (7101).

Bayer, U., Habibi, I., Balzarotti, D., Kirda, E., and Kruegel, C. (2009a). A view on current malware behaviors. In *USENIX workshop on large-scale exploits and emergent threats (LEET)*.

Bayer, U., Kruegel, C., and Kirda, E. (2009b). Anubis: Analyzing unknown binaries.

Bernheim, B. D. and Whinston, M. D. (1986). Menu auctions, resource allocation, and economic influence. *The quarterly journal of economics*, pages 1–31.

Besley, T. (2006). Principled agents?: The political economy of good government. *OUP Catalogue*.

Bohn, H. and Deacon, R. T. (2000). Ownership risk, investment, and the use of natural resources. *American Economic Review*, pages 526–549.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *American Economic Review*, 103(5):1759–96.

Bulte, E. H., Damania, R., and Lopez, R. (2007). On the gains of committing to inefficiency: corruption, deforestation and low land productivity in latin america. *Journal of Environmental Economics and Management*, 54(3):277–295.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., and Sieber, S. (2012). The political economy of deforestation in the tropics. *The Quarterly Journal of Economics*, 127(4):1707–1754.

Burstein, A. J. (2008). Conducting cybersecurity research legally and ethically. *LEET*, 8:1–8.

Chibwana, C., Fisher, M., and Shively, G. (2012). Cropland allocation effects of agricultural input subsidies in malawi. *World Development*, 40(1):124–133.

Christin, N., Yanagihara, S. S., and Kamataki, K. (2010). Dissecting one click frauds. In *Proceedings of the 17th ACM conference on Computer and communications security*, pages 15–26. ACM.

Cisneros, E., Hargrave, J., Kis-Katos, K., et al. (2013). Unintended consequences of anti-corruption strategies: public fiscal audits and deforestation in the brazilian amazon. *Unpublished working paper.*

Correll, S. and Corrons, L. (2010). The business of rogueware: analysis of the new style of online fraud.

Cova, M., Kruegel, C., and Vigna, G. (2010). Detection and analysis of drive-by-download attacks and malicious javascript code. In *Proceedings of the 19th international conference on World wide web*, pages 281–290. ACM.

De Janvry, A., Finan, F., and Sadoulet, E. (2012). Local electoral incentives and decentralized program performance. *Review of Economics and Statistics*, 94(3):672–685.

Deacon, R. T. (1994). Deforestation and the rule of law in a cross-section of countries. *Land Economics*, pages 414–430.

Dhamija, R., Tygar, J. D., and Hearst, M. (2006). Why phishing works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM.

Dittrich, D., Bailey, M., and Dietrich, S. (2009). Towards community standards for ethical behavior in computer security research. Technical report, Technical Report 2009-01, Stevens Institute of Technology, Hoboken, NJ, USA.

Dorward, A. and Chirwa, E. (2011). The malawi agricultural input subsidy programme: 2005/06 to 2008/09. *International journal of agricultural sustainability*, 9(1):232–247.

Dorward, A., Chirwa, E., Kelly, V. A., Jayne, T. S., Slater, R., and Boughton, D. (2008). Evaluation of the 2006/7 agricultural input subsidy programme, malawi. final report. Technical report, Michigan State University, Department of Agricultural, Food, and Resource Economics.

Egelman, S., Cranor, L. F., and Hong, J. (2008). You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1065–1074. ACM.

Ejdemyr, S., Kramon, E., and Robinson, A. L. (2015). Segregation, ethnic favoritism, and the strategic targeting of distributive goods1.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2).

Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274–1311.

Ferreira, S. and Vincent, J. R. (2010). Governance and timber harvests. *Environmental and Resource Economics*, 47(2):241–260.

Fossi, M., Turner, D., Johnson, E., Mack, T., Adams, T., Blackbird, J., Low, M., McKinney, D., Dacier, M., Keromytis, A., et al. (2009). Symantec report on rogue security software. *Whitepaper, Symantec, October.*

Franklin, J., Perrig, A., Paxson, V., and Savage, S. (2007). An inquiry into the nature and causes of the wealth of internet miscreants. In *ACM conference on Computer and communications security*, pages 375–388.

Fujiwara, T. (2010). Voting technology, political responsiveness, and infant health: evidence from brazil. *Department of Economics, University of British Columbia. http://grad. econ. ubc. ca/fujiwara/jmp. pdf.*

Garfinkel, S. L. (2008). Irbs and security research: myths, facts and mission creep. *UPSEC*, 8:1–5.

Grossman, G. M. and Helpman, E. (1994). Protection for sale. *The American Economic Review*, 84(4):pp. 833–850.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., Loveland, T., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853.

Hansen, M. C., Shimabukuro, Y. E., Potapov, P., and Pittman, K. (2008). Comparing annual modis and prodes forest cover change data for advancing monitoring of brazilian forest cover. *Remote Sensing of Environment*, 112(10):3784–3793.

Hargrave, J. and Kis-Katos, K. (2013). Economic causes of deforestation in the brazilian amazon: A panel data analysis for the 2000s. *Environmental and Resource Economics*, pages 1–24.

Holz, T., Engelberth, M., and Freiling, F. (2009). *Learning more about the underground economy: A case-study of keyloggers and dropzones*. Springer.

Ikinci, A., Holz, T., and Freiling, F. C. (2008). Monkey-spider: Detecting malicious websites with low-interaction honeyclients. In *Sicherheit*, volume 8, pages 407–421.

INPE (2013). Prodes: Monitoramento da floresta amaznica por satlite. *www.obt.inpe.br/prodes/. Accessed on 15 July 2013*.

Kenneally, E., Bailey, M., and Maughan, D. (2010). A framework for understanding and applying ethical principles in network and security research. In *Financial Cryptography and Data Security*, pages 240–246. Springer.

Kirk, J. (2010). Bredolab-infected pcs downloading fake antivirus software.

Kramon, E. and Posner, D. N. (2013). Who benefits from distributive politics? how the outcome one studies affects the answer one gets. *Perspectives on Politics*, 11(02):461–474.

Krebs, B. (2009a). Massive profits fueling rogue antivirus market. *Washington Post*.

Krebs, B. (2009b). Virus scanners for virus authors.

Krebs, B. (2010a). Following the money; epassporte edition.

Krebs, B. (2010b). Rogue antivirus victims seldom fight back.

Ludl, C., McAllister, S., Kirda, E., and Kruegel, C. (2007). On the effectiveness of techniques to detect phishing sites. In *DIMVA*, volume 7, pages 20–39. Springer.

Mahdavi, P. (*Forthcoming* 2015). Explaining the oil advantage: Effects of natural resource wealth on incumbent reelection in iran. *World Politics*, 67(2).

Mavrommatis, N. P. P. and Monrose, M. A. R. F. (2008). All your iframes point to us. In *USENIX Security Symposium*, pages 1–16.

May, P. H., Millikan, B., Gebara, M. F., et al. (2010). The context of redd+ in brazil: drivers, agents and institutions. *CIFOR Occasional Paper*, (55).

McGrath, D. K. and Gupta, M. (2008). Behind phishing: An examination of phisher modi operandi. *LEET*, 8:4.

Mick, J. (2010). Russian anti-spam chief caught spamming.

Micro, T. (2009). The business of cybercrime-a complex business model. *A Trend Micro White Paper*.

Moore, T. and Clayton, R. (2007). An empirical analysis of the current state of phishing attack and defence. In *WEIS*. Citeseer.

Morjaria, A. (2014). Is democracy good for the environment in developing countries? evidence from kenya. *Working Paper, Harvard University*.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., da Motta, R. S., et al. (2014). Slowing amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–1123.

Nolte, C., Agrawal, A., Silvius, K. M., and Soares-Filho, B. S. (2013). Governance regime and location influence avoided deforestation success of protected areas in the brazilian amazon. *Proceedings of the National Academy of Sciences*, 110(13):4956–4961.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., et al. (2011). A large and persistent carbon sink in the worlds forests. *Science*, 333(6045):988–993.

Pan, Y. and Ding, X. (2006). Anomaly based web phishing page detection. In *Computer Security Applications Conference, 2006. ACSAC'06. 22nd Annual*, pages 381–392. IEEE.

Pauw, K. and Thurlow, J. (2014). *Malawi's farm input subsidy program: Where do we go from here?*, volume 18. International Food Policy Research Institute.

Pfaff, A., Robalino, J., Lima, E., Sandoval, C., and Herrera, L. D. (2013). Governance, location and avoided deforestation from protected areas: Greater restrictions can have lower impact, due to differences in location. *World Development*.

Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., Perz, S., Bohrer, C., Arima, E., Laurance, W., et al. (2007). Road investments, spatial spillovers, and deforestation in the brazilian amazon. *Journal of Regional Science*, 47(1):109–123.

Pfaff, A. S. (1999). What drives deforestation in the brazilian amazon?: evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management*, 37(1):26–43.

Posner, D. N. (2004). The political salience of cultural difference: Why chewas and tumbukas are allies in zambia and adversaries in malawi. *American Political Science Review*, 98(04):529–545.

Poulsen, K. (2009). Conficker doomsday worm sells out for 49.95.

Provos, N., McNamee, D., Mavrommatis, P., Wang, K., Modadugu, N., et al. (2007). The ghost in the browser analysis of web-based malware. In *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, volume 10, pages 4–4.

Rajab, M. A., Ballard, L., Mavrommatis, P., Provos, N., and Zhao, X. (2010). The nocebo effect on the web: an analysis of fake anti-virus distribution. In *USENIX workshop on large-scale exploits and emergent threats (LEET)*.

Robinson, A. L. (2013). Internal borders: ethnic diversity and market segmentation in malawi. In *Working Group in African Political Economy National Meeting, Massachusetts Institute of Technology*, pages 3–4.

Rondoniagora (2014). Prefeito de nova mamor e empresrios so denunciados por fraude para legalizar madeira clandestina. *Rondoniagora Newspaper*, June 10, 2014.

Rosiello, A. P., Kirda, E., Kruegel, C., and Ferrandi, F. (2007). A layout-similarity-based approach for detecting phishing pages. In *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*, pages 454–463. IEEE.

Samosseiko, D. (2009). The partnerkawhat is it, and why should you care. In *Proc. of Virus Bulletin Conference*.

Stone-Gross, B., Abman, R., Kemmerer, R. A., Kruegel, C., Steigerwald, D. G., and Vigna, G. (2013). The underground economy of fake antivirus software. In *Economics of Information Security and Privacy III*, pages 55–78. Springer.

Stone-Gross, B., Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., Kruegel, C., and Vigna, G. (2009a). Your botnet is my botnet: analysis of a botnet takeover. In *Proceedings of the 16th ACM conference on Computer and communications security*, pages 635–647. ACM.

Stone-Gross, B., Cova, M., Kruegel, C., and Vigna, G. (2011a). Peering through the iframe. In *INFOCOM, 2011 Proceedings IEEE*, pages 411–415. IEEE.

Stone-Gross, B., Holz, T., Stringhini, G., and Vigna, G. (2011b). The underground economy of spam: A botmasters perspective of coordinating large-scale spam campaigns. In *USENIX workshop on large-scale exploits and emergent threats (LEET)*, volume 29.

Stone-Gross, B., Kruegel, C., Almeroth, K., Moser, A., and Kirda, E. (2009b). Fire: Finding rogue networks. In *Computer Security Applications Conference, 2009. ACSAC'09. Annual*, pages 231–240. IEEE.

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., and Storch, D. (2014). Comment on "high-resolution global maps of 21st-century forest cover change". *Science*, 344(6187):981–981.

UNFAO (2010). *Global forest resources assessment 2010: Main report.* Food and Agriculture Organization of the United Nations.

Villeneuve, N., Deibert, R., and Rohozinski, R. (2010). *Koobface: Inside a crimeware network.* Munk School of Global Affairs.

Wendland, K. J., Lewis, D. J., and Alix-Garcia, J. (2014). The effect of decentralized governance on timber extraction in european russia. *Environmental and Resource Economics*, 57(1):19–40.

Westberg, N. B. et al. (2015). Exchanging fertilizer for votes? *Working Paper.*

Zhuge, J., Holz, T., Song, C., Guo, J., Han, X., and Zou, W. (2009). *Studying malicious websites and the underground economy on the Chinese web.* Springer.

# Appendix A

# Reelection Incentives, Blacklisting and Deforestation in Brazil

### A.0.1 Proof of Proposition 1

Proof by contradiction. Let $\boldsymbol{h}^*$ be the vector of deforestation allocations that maximizes the utility of a corrupt politician conditional on $\sum_{i=1}^{N} h_i = H$ for a given level of $H$ and given a set of truthful bids from landowners, $\{s_i^*\}_{i \in N}$. This implies

$$\boldsymbol{h}^* = ARGMAX_{\boldsymbol{h}}\{\sum_{i=1}^{N} s_i^*(h_i)\} \tag{A.1}$$

$$S.T. \sum_{i=1}^{N} h_i = H.$$

Now assume some other allocation vector, $\tilde{\boldsymbol{h}} \neq \boldsymbol{h}^*$ yields a higher payoff for the benevolent mayor than does $\boldsymbol{h}^*$ and also accords with the constraint of $\sum_{i=1}^{N} h_i = H$. This implies

$$\sum_{i=1}^{N} u_i(\tilde{h_i}) + \sum_{i=1}^{N} s_i^*(\tilde{h_i}) > \sum_{i=1}^{N} u_i(h_i^*) + \sum_{i=1}^{N} s_i^*(h_i^*). \tag{A.2}$$

Substituting equation (1.2) for $u_i(h_i)$ and canceling the positive sum of $s_i^*(h_i)$ with the negative sum of $s_i^*(h_i)$ from (1.2), yields the following inequality.

$$\sum_{i=1}^{N} P\alpha_i q(\tilde{h_i}) > \sum_{i=1}^{N} P\alpha_i q(h_i^*) \tag{A.3}$$

Subtracting $\sum_{i=1}^{N} P\alpha_i q(0)$ from both sides yields

$$\sum_{i=1}^{N} P\alpha_i q(\tilde{h}_i) - \sum_{i=1}^{N} P\alpha_i q(0) > \sum_{i=1}^{N} P\alpha_i q(h_i^*) - \sum_{i=1}^{N} P\alpha_i q(0). \tag{A.4}$$

Which is equivalent to

$$\sum_{i=1}^{N} \Big( P\alpha_i q(\tilde{h}_i) - \alpha_i P q(0) \Big) > \sum_{i=1}^{N} \Big( P\alpha_i q(h_i^*) - P\alpha_i q(0) \Big).$$

$$\sum_{i=1}^{N} s_i^*(\tilde{h}_i) > \sum_{i=1}^{N} s_i^*(h_i^*) \tag{A.5}$$

Which contradicts equation (A.1). Therefore, any allocation vector such that $\sum_{i=1}^{N} h_i = H$ that maximizes the utility of the corrupt politician will also maximize the utility of the benevolent politician.