ABSTRACT

Title of dissertation: AVANÇA BRASIL AND DEFORESTATION IN THE AMAZON

Timothy Scott Thomas, Doctor of Philosophy, 2003

Dissertation directed by: Professor Marc Nerlove
Department of Agricultural and Resource Economics

The goal of this study is to predict the change in deforestation in the Brazilian Amazon resulting from road paving under *Avança Brasil.* Predictions are reported in Chapter 5, where I find that the effect on on-farm deforestation is a moderate 2.5 percent increase, and in the very long run, a 7.9 percent (entire Amazon) or 11.5 percent (forest biome only) increase. When off-farm deforestation is considered as well, net deforestation increases 10.3 percent and gross deforestation increases 14.4 percent. In the very long run, both net and gross deforestation increase by around 25 percent.

Chapter 1 explains why researchers should treat the Legal Amazon as two separate ecozones, forest and *cerrado.* I then estimate the magnitude of forest regrowth in the Amazon and, finally, demonstrate the importance of defores-

tation “seeds” from which deforestation spreads. Chapter 2 presents a simple dynamic model of land clearing. It includes several regressions which confirm the importance of starter deforestation and show how the effect of roads on deforestation appears substantially stronger when deforestation “seeds” are not controlled for. The regressions also show the importance of agroclimatic variables and the effectiveness of protected areas.

Chapter 3 demonstrates a new spatial disaggregation technique that allows one to map the likely distribution of farm activity inside the cross-sectional units. The technique is more efficient than standard regression analysis on uncensored data, but not necessarily so with censored data.

Chapter 4 uses satellite data for 1996 to 1999 to show that after a slight rise, rates of deforestation decline with increases in deforestation levels, and that this effect is quantitatively large and statistically significant. Rates of deforestation computed in Chapter 4 show that clearing of agricultural land might take place over 60 years until steady state is reached.

Deforestation rates on land already settled are not responsive to farmgate prices, though deforestation levels are. New settlement locations are very sensitive to farmgate prices. Even after controlling for prices and agroclimatic suitability, new farm establishments are much more likely to locate close to already established farms than far from them.

AVANÇA BRASIL AND DEFORESTATION IN THE AMAZON

by

Timothy Scott Thomas

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2003

Advisory Committee:

Professor Marc Nerlove, Chair
Professor Nancy Bockstael
Professor Anna Alberini
Professor John List
Professor Maureen Cropper

DEDICATION

This work is dedicated to my father, James Franklin (Frank) Thomas, who died just two months short of its completion. Only after his death did I take time to reflect on the accomplishments of his life. One of the most remarkable, in light of my experience over the last several years, was his earning a law degree while being married and working full time—and graduating magna cum laude, to boot. Never one to brag about himself or to complain, I never heard stories of how he managed to pull it off, against the odds. In the final stretch of my own academic marathon, I would have benefited from the lessons he could have taught me.

ACKNOWLEDGEMENTS

This dissertation would not have been written without the encouragement, advice, and assistance of a number of people. While I would not want to try to rank those to whom I am indebted, there is no doubt that my colleague and friend Ken Chomitz is among the group of those most responsible for its completion. He took a risk in hiring me four years ago to work with him on a year-long study of Amazon deforestation data, when I knew little about the Amazon, and nothing about Arc View and Stata, the tools of the trade. Little did I know that I would work with him over the course of several years, and that I would eventually shift the focus of my dissertation from soil fertility and fallow systems to agro-economic potential and forest conversion in the Brazilian Amazon. Thanks, Ken, for teaching me, mentoring me, being patient with me while I tried to verbalize many half-baked ideas—and for the many ways you showed kindness and respect.

My colleagues and friends at IMAZON have been so patient and supportive through the years. My understanding of many of the finer points of deforestation (and data issues) would not have been possible without their help. I particularly thank Eugenio Arima, who has repeatedly gone out of his way to answer obscure questions and keep me on the right path. Perhaps I can return the favor as you head down the last stretch of your own dissertation. I must also thank Carlos Souza and Rodney Salomão who showed a great deal of patience with me as I pursued the intricacies of some of the GIS datasets they kindly shared with me. Thanks to Paulo Barreto who helped coordinate much of the interaction between Ken and me at the World Bank, and Eugenio, Carlos, Rodney, and Marky Brito at IMAZON in Belém.

I am grateful to other colleagues at the World Bank. For helping me with the finer points of GIS, I am grateful to Uwe Deichmann who is such a vast ocean of knowledge and experience. Over the course of the last several years I have benefited from the help and advice from Sheila Wertz and Andy Nelson, and in my very early days of learning GIS and spatial econometrics, from Jyotsna Puri and Somik Lall.

In the area of understanding Amazon deforestation, I have learned much from Bob Schneider and Sergio Margulis. Bob, in particular, always managed to look at tables and maps that I had worked on, and ask deep questions that caused me to delve even deeper into the data, to see what more we might learn.

Moving away from the work I did at the World Bank, there are a number of others that I want to thank. My adviser, Marc Nerlove, has shown a remarkable combination of patience and persistence, knowing when to push, and when to back off. I am sure that when I started working at the World Bank, he was skeptical that I would finish this dissertation. But he also did not give up hope, and gently prodded me to action. My regret about working at the World Bank is that I had less time to interact with Marc, from whom I learned a great deal in my early graduate school years as his teaching assistant and research assistant.

Many of my dearest friends from church and prayer groups have prayed for me (for years!) to be able to complete this dissertation. Prayer is powerful—so thanks for helping me have the strength to persevere! The people in my congregation have been remarkably understanding of their pastor, who has dropped many balls while trying to devote as much energy to writing this thesis as possible. Thanks for picking up some of the dropped balls and lightening my load—including volunteering to give some of the sermons so that I could have more time for writing other things. Thanks to my mom and dad who kept faith in me through all these years.

Thanks to my dear wife, Steph, and my precious children, Josiah and Esther. Through the years, Steph has consistently encouraged me to keep at it until it is done. And toward the end, she sacrificed much time and energy, covering for me while I was pounding away at the computer in the basement. Josiah and Esther are too young to understand why it is important for Daddy to work hard and not play with them, but they have been loving and forgiving, nonetheless. I hope to make up to you soon for missed time. Thanks also to Steph and our dear friend Linda Price, for proofreading my drafts, and offering helpful and insightful advice on how to say things more clearly. Any remaining confusing sentences are probably due to me not having sense enough to take their advice!

Finally, I must thank my God and King, who makes all things work for the good of those who love him (Romans 8:28). Little did I know that this doctorate was going to turn into a character-building experience—for which I am grateful. Thanks, Father, for your kindness and faithfulness and love—without which my life (and even this dissertation) would be without meaning!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

BSRSI	Basic Science and Remote Sensing Initiative is a research program in the Department of Geography at Michigan State University focused on global changes at both regional and global scales.
CI	Conservation International.
CID	Center for International Development, located at Harvard University.
EOS	Earth Observing System-Amazon, a project at University of Washington to understand the biogeochemistry, hydrology, and sedimentation of the Amazon River and its drainage basin.
ESRI	Environmental Systems Research Institute.
ETM	Enhanced Thematic Mapper, one of the newest sensors used to detect deforestation from satellites.
FAO	Food and Agriculture Organization.
GIS	Geographic Information System.
IBAMA	*Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis*, Brazilian Institute of Environmental and Natural Resource Protection.
IBGE	*Instituto Brasileiro de Geografia e Estatística*, Brazilian Institute of Geography and Statistics.
IFPRI	International Food Policy Research Institute, located in Washington, D. C.
IMAZON	*Instituto do Homem e Meio Ambiente da Amazônia*, the Amazon Institute of People and the Environment, located just outside of Belém, Brazil.

INPE	*Instituto Nacional de Pesquisas Espaciais*, the Brazilian Space Agency.
IPAM	*Instituto de Pesquisa Ambiental da Amazônia*
IPCC	Intergovernmental Panel on Climate Change.
ISA	*Instituto Socio-Ambiental*.
JERS	Japanese Earth Resources Satellite.
MSS	Multi Spectral Scanner, an older sensor used to detect deforestation from satellites.
NASDA	National Space Development Agency of Japan.
NIMA	National Imagery and Mapping Agency.
PAM	*Produção Agrícola Municipal*, annual survey of crops.
PEV	*Produção Extrativa Vegetal*, annual survey of extracted forestry products.
PPM	*Pesquisa Pequária Municipal*, annual survey of livestock.
SAR	Synthetic Aperture Radar
TM	Thematic Mapper, a sensor used to detect deforestation from satellites.
TRFIC	Tropical Rain Forest Information Center, part of BSRSI at Michigan State University.
UNEP	United Nations Economic Programme
USDA	U. S. Department of Agriculture
USGS	U. S. Geological Survey
WCMC	World Conservation Monitoring Centre, affiliated with UNEP.
WHRC	The Woods Hole Research Center.

WIDER	World Institute for Development Economics Research, part of the United Nations University. Located in Helsinki.

Chapter 1: The State of the Brazilian Amazon Forest

Why Study Deforestation in the Amazon?

Tropical forests appear to be gravely threatened. This is a global concern, because the consequences of their reduction could have wide-reaching impact. First of all, burning of tropical forests releases vast amounts of greenhouse gases, which are stored in the upper atmosphere. The build up of these gases is believed to contribute to a global rise in temperatures that could potentially shift the patterns of rainfall throughout the world, causing floods and droughts of greater severity than normal. It could also lead to significant melting of the polar ice caps, causing the oceans to rise, and raising havoc on coastal populations (IPCC).

Furthermore, tropical deforestation appears to be responsible for the permanent loss of plant and animal species, many of which are lost before they can even be catalogued as having existed (Mittermeier at al.). Driving species into extinction might be argued to be alarming because of moral reasons. But even those who do not accept the moral argument might be concerned about their loss on the grounds of not understanding the larger and perhaps irreplaceable role each species plays in an ecosystem, and therefore the possible “domino” effect that species loss might have on the rest of the local or even global environment. We may also be concerned that we are losing valuable genetic stock that could be part of the solution to health, agricultural, environmental, and commercial problems.

The Amazon Forest is the largest tropical forest in the world, covering between 5.5 (Andersen et al.) and 6.6 (Wood) million square kilometers, depending upon who you ask. Like other tropical forests, the Amazon is under severe pressure from people desiring to profit from it—some of them very wealthy already, others very poor. The Amazon may be the best known tropical forest in the world and—perhaps because of its size and popularity—is one of the most studied. The Amazon spans six (Wood) or seven (EOS) countries, again, depending on who is counting. The largest portion—greater than the sum of all the portions in other countries—lies in Brazil.

Most of the studies of the Amazon focus on Brazil. This is probably in part due to the fact that there are more data available on the Brazilian Amazon than on any other part of the Amazon. The Brazilian government deserves much credit for making annual data on forest and agricultural products by *município* (i.e., county) readily available through the internet.[1] They have also provided detailed deforestation data on the internet, showing the precise location of patches of deforestation as small as a hectare that have

---

[1] For example, http://www.sidra.ibge.gov.br, last accessed July 2003.

taken place in the high deforestation areas of the Amazon since 1996.[2] This disclosure of valuable data has allowed researchers to study the causes of deforestation, and propose solutions resulting from their analyses.

The ultimate goal of this study is to estimate the impact on Amazon deforestation of the road paving projects proposed by the Brazilian government in their $45 billion investment program called "*Avança Brasil*". A secondary objective is to use the data in this study to contribute to the understanding of the land use and land change processes in the Brazilian Amazon. Another secondary objective is to introduce a technique to be used with GIS data in regression analysis that will enable the researcher to predict the geographic distribution of the dependent variable inside each cross-sectional unit.

What follows in this chapter is a brief overview of *Avança Brasil.* Then I present an argument for and evidence supporting the idea that researchers should treat the Legal Amazon as two separate ecozones: the forest and the *cerrado.* Next, I review satellite and survey data useful for studying deforestation in the Brazilian Amazon, followed by simple calculations and cross-tabulations which address two issues of importance throughout this study: the issue of forest regrowth, and the issue of "seed" deforestation.

---

[2] http://www2.ibama.gov.br/desmata/.

Chapter 2 describes the explanatory variables used throughout the rest of the study, then presents a simple dynamic model of land clearing on farms that forms the underlying framework for the regressions of Chapters 2 through 5, and especially the analysis done in Chapter 4. Chapter 2 concludes with several regressions using agricultural census data. These regressions explore the effects on farm area cleared of roads, protected areas, "seed" deforestation, and agroclimatic suitability.

Chapter 3 repeats the main regression from Chapter 2, presenting a new spatial disaggregation technique. Chapter 4 uses satellite data for 1996 to 1999 to look at the dynamics of deforestation, investigating how the rate of deforestation is influenced by farmgate prices for beef and milk, level of deforestation, agroclimatic variables, and protected areas.

Chapter 5 reviews three well-known predictions and a fourth less well-known prediction for the effects of road paving—as planned by *Avança Brasil*—on deforestation. I then present my own predictions and some policy recommendations. Chapter 6 summarizes and concludes.

Avança Brasil

In the early 1960s, the Brazilian government took a major step by proposing a bold plan for the development of the Amazon which included the establishment of major roads as a necessary precursor to development. Since then, the government has revised its plans for Amazonian development every few years, following a pattern of ebb and flow, with some plans being very conservative, and others being very expansive. The Brazilian government's $45 billion plan for Amazonian development during the 1999-2006 period is called *Avança Brasil* (Cattaneo). Relative to recent years, the plan appears to be quite bold in its proposal to expand infrastructure, perhaps the boldest since the initial plans to develop infrastructure in the Amazon.

The plan proposes expanding paved roads from 11,900 kilometers to 18,145 kilometers (IPAM et al.).[3] Most of the expansion would be paving of already existing major roads, including the Santarém-Cuiabá highway (BR-163); the Humaitá-Manaus highway (BR-319); the Transamazônica highway between Marabá and Rurópolis (BR-230); the Manaus-Boa Vista highway (BR-174, which was paved in 1997 and 1998); the Cuiabá-Porto Velho highway (BR-364); and more than 2,000 kilometers of other roads (IPAM et al.). The first two highways cut through the heart of the Amazon, but have low population densities and are currently in such poor condition that the areas along them are inaccessible through most of the year (Cattaneo).

Avança Brasil also includes plans for ports, waterways, and a gas pipeline. This dissertation, however, focuses on road paving.

---

[3] Cattaneo agrees that new paving will increase the stock of paved roads by 6,245 kilometers, but writes that the current stock of paved roads is only around 7,000 kilometers.

Ecological Zones and the Challenge of Defining the Study Area

A legislative compromise expanded the political boundaries of the Amazon region in Brazil so that parts of Maranhão, Tocantins, and Mato Grosso lying outside the tropical forest might also benefit from regional development incentives. The resulting area, called the "Legal Amazon", includes the states of Pará, Amapá, Roraima, Amazonas, Acré, Rondônia, Mato Grosso, almost all of Tocantins, roughly half of Maranhão, and a tiny sliver of Goias. When studying the Brazilian Amazon, researchers have typically chosen to study the Legal Amazon, perhaps because extra data are available for that region, and perhaps to be in agreement with the "legal" definition of the Amazon (examples include Pfaff 1997, 1999; Andersen et al.; Reis and Margulis; Perz; and even my own work with my colleague, in Chomitz and Thomas). However, Figures 1 through 6 show that there is virtually no scientific reason to study the Legal Amazon as a single unit.

In Figure 1, note that the Amazon Basin proper—as determined by the U.S. Geological Survey from their hydrological work based on the digital elevation model—excludes a large portion of the eastern and southern Legal Amazon. Indeed, while most of the non-Amazon river basins in the Legal Amazon empty into the Atlantic Ocean near the mouth of the Amazon river, it is interesting to note that the river basin in southern Mato Grosso empties into the Atlantic Ocean on the other end of the continent, thousands of kilometers away from the mouth of the Amazon. Figure 1 also includes the Amazon Basin boundary created by the EOS (Earth Orbiting System) Amazon Project at the University of Washington. The names assigned to each river basin are those given by me for ease of reference, and are not part of the USGS dataset.

While it is helpful to understand where the different river basins lie, I do not believe that the study area should be limited by the river basins. What is really of interest is the boundaries for the ecozone which contains the tropical forest in and around the Amazon basin. When we look at the maps in Figures 2 through 5, we see that all of the major ecozone systems for Brazil distinguish the forest biome from the *cerrado* (i.e., savanna) biome. While these systems (WCMC, WWF, and two from IBGE) do not agree precisely, they agree generally that a large portion of the southern and southeastern Legal Amazon is in the *cerrado* biome.

This would not be important to a study of tropical deforestation, except that economic development in the *cerrado* biome is vastly different than in the Amazon forest biome. This is probably because the *cerrado* gets less rainfall on average than the Amazon forest, and therefore is more suitable for agriculture. Furthermore, there are less trees and less dense vegetation in the *cerrado*; therefore, the land is easier to convert to agriculture than land in the forest. Studying the *cerrado* portions of the Legal Amazon to understand tropical deforestation is akin to a doctor examining one patient in order to make a diagnosis on an unrelated patient. It makes more sense to examine each patient, and treat each based on its own symptoms.

This is not to say that the *cerrado* is unworthy of study. Indeed it is! Conservation International (CI) ranks the Brazilian *cerrado* as the second most critical environmental hotspot in the world (Mittermeier at al.), mostly due to its large number of threatened species. The Brazilian Amazon is not even listed among CI's top ten. But much of the Brazilian *cerrado* lies outside the Legal Amazon. To do a study of the Legal Amazon as a single ecosystem does a disservice to both the Amazon forest and the *cerrado*.

To be convinced that the *cerrado* area of the Legal Amazon is genuinely different than the Amazon forest area, look at the image in Figure 6 made by the University of Maryland from satellite images taken in 1992 and 1993. Note how the *cerrado* has considerably less vegetation than the rest of the Legal Amazon. While this could be in part from settlement and deforestation, it is largely due to the vegetation naturally being much less dense.

Upon closer inspection, we note other lightly vegetated areas within the Legal Amazon. In fact, just as the *cerrado* biome is considered savanna, so are some of these areas (see Roraima, for instance). Many systems of delineating ecozones declare these as part of the Amazon forest. Why is it reasonable to consider these savannas as part of the Amazon forest, while the other savanna (*cerrado*) is considered a separate ecozone? The answer lies in the fact that these patches are surrounded on all sides by forest, while the *cerrado* is a much larger region, and is not surrounded by one vegetation type. While this study includes these patches in the analysis of deforestation in the Brazilian Amazon, these microvariations have very important explanatory power as to how the settlement pattern has become what it has within the forest biome.

In Figure 2, we see that the *babaçu* forests[4] extend beyond the limit of the Legal Amazon, making a case to likewise extend our study area to include the full extent of the *babaçu* forests. However, some of the important data are not available beyond the limit of the Legal Amazon, and for this reason, we likewise will exclude the portions of the *babaçu* forests which lie beyond the limits of the Legal Amazon.

Sources of Data on Deforestation and Agricultural Expansion

There are two categories of datasets which provide information on deforestation: satellites and surveys (including censuses). Each has strengths and weaknesses. Satellite data are relatively inexpensive and the coverage is wide when compared to surveys and censuses. They can be re-sampled relatively easily. However, the best satellite interpretation can be costly, because it in-

---

[4] *Babaçu* trees are the natural secondary vegetation to spring up in large portions of Maranhão. The *babaçu* is a palm tree that produces a kernel from which an oil is extracted, which is useful for soap and foods. The nut shells are used for charcoal for cooking (Porro).

volves a lot of visual inspection. Moreover, satellite data are sometimes difficult to interpret, in the sense that there are not always clear cutoffs as to where a certain density of vegetation can be thought of as forest, and anything less can be thought of as deforestation. Because *cerrado* looks like deforestation, these areas are not easily examined by satellite images. Finally, clouds keep us from seeing what is being done to the land in the Landsat images (but not radar images).

Surveys are by nature limited in scope, and so usually do not provide sufficient geographic coverage. Censuses, on the other hand, have complete geographic coverage, but they are usually expensive, and as a result, can only be done infrequently. Furthermore, they may miss important agents of deforestation. In the case of the agricultural census, only farmers are surveyed, not loggers. Even if loggers were surveyed, those doing illegal logging would not likely report accurate information. Farmers also may have motives to report their land uses incorrectly.[5] Even when farmers intend to report land use honestly, because of the vast size of many farms and the irregular shape of tracts of land, it might be difficult for farmers to report land areas

---

[5] There are legal limits to the amount of deforestation that can be done on farms, but the survey seems to indicate that most farmers are willing to report that they have far exceeded those limits. There may be motives for inflating clearing figures: farmers need to show that land is being put to productive use, or it can be confiscated.

accurately. Censuses do, however, allow us to have insight into *cerrado* land use, and this makes the agricultural census extremely important in understanding deforestation.

Despite the downside of both types of data, in that both are subject to some measurement error, they are the best and only indicators of deforestation that we have, and therefore we should not discount them too much. Furthermore, they can be particularly valuable in tracking deforestation over time.

Satellite data

INPE

The Brazilian space agency INPE (*Instituto Nacional de Pesquisas Espaciais*) has been charged with monitoring deforestation in the Amazon. They do this by studying LANDSAT images of the Brazilian Amazon. In their preliminary analysis, they examine between 44 and 49 images, representing the "hotspots" of deforestation—roughly 80 percent of the total deforestation in the Amazon. Then, in the following year, they examine the remainder of the 229 images, making corrections to their original estimate. Only patches of deforestation that are at least 6.25 hectares in size are tracked, though the satellite reports data for each 0.09 hectare unit (INPE).

INPE's report from June 2002 states that out of the approximately 4 million square kilometers of forest formation in the Legal Amazon, as of August

2001, 603,514 square kilometers have been deforested (INPE). Since 1997, deforestation has averaged 17,164 square kilometers per year—a rate of 0.5 percent of the remaining forest per year—an area just smaller than the state of New Jersey. This is down from the 1977 to 1988 average of 21,130 square kilometers per year, but is higher than the average for 1989 to 1994, which was 13,703 square kilometers per year. These numbers represent gross deforestation—that is, they do not adjust for forest added through regrowth in areas already cleared. It is difficult to know how much, if any, regrowth is taking place each year, though later in this chapter, we will examine other data that will give some indication of the magnitude of regrowth.

IBAMA

IBAMA's work is similar to that of INPE, in that it uses satellite data to monitor deforestation. IBAMA (*Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis*) is probably more useful to the researcher, because their data are more accessible, with both visual and tabular data available on the web.[6] The smallest patch of deforestation that IBAMA measures is 1 hectare, which is considerably smaller than that of INPE. IBAMA focuses on ap-

---

[6] http://www2.ibama.gov.br/desmata/index.htm, last accessed July 2003. However, as of June 2003, INPE began making their data available, which is now potentially more useful than that of IBAMA. Their data can be accessed from http://www.obt.inpe.br/prodesdigital/cadastro.php.

proximately 50 scenes[7] per year, so it is not possible to research deforestation in the entire Amazon from their data, but only what they deem to be the most active scenes. Like INPE, they do not attempt to keep track of reforestation.

TRFIC

The Tropical Rainforest Information Center (TRFIC) is part of the Basic Science and Remote Sensing Initiative (BSRSI) at Michigan State University. They have a large number of satellite products available for downloading from the web.[8] At the time of this writing, one can download interpreted satellite information for 1986, 1992, and 1996 for almost any scene in the Legal Amazon. The 1999 data should be available soon. Since TRFIC uses an algorithm to interpret the deforestation data, there are more likely to be some errors in interpretation that would not have occurred with more time-consuming interpretation by analysts. On the other hand, the data are very accessible here. Furthermore, their algorithm allows for regrowth of forest, and in that sense is superior to the data and subsequent analysis by INPE or IBAMA. It is also superior in that the data allow for detection of deforestation in a patch as small as 0.09 hectares,[9] and that the time horizon is much

---

[7] A scene is a Landsat image in the shape of a rhombus, approximately 180 kilometers on edge.

[8] http://www.bsrsi.msu.edu/trfic/, last accessed July 2003.

[9] They do not report deforestation patches smaller than 2.93 hectares, however.

longer, with some data available from the 1970s. The weaknesses in this data source relative to INPE are some missing scenes, and the scenes not being well georeferenced through time. That is, a gridcell can move up to two or three kilometers going from the 1986 scene to the 1992 scene. In 1999 a new Landsat was launched which georeferences the data much better than the previous ones, and this problem will disappear with time. But if one is trying to do some type of panel analysis with this dataset, the deforestation is likely to have to be aggregated to at least 8 kilometer square units of analysis to minimize measurement error—or the researcher will have to "rubbersheet" the layers of data carefully to make them align more precisely.

JERS and other radar

The JERS-1 (Japanese Earth Resources Satellite) Synthetic Aperture Radar (SAR) of the National Space Development Agency of Japan (NASDA) was used in late-1995 through mid-1996 to study deforestation in the Amazon. Radar has the advantage that it can look through clouds and can see at night. The resolution on this specific product is approximately 90 meters, but other radar images have resolutions of 20 meters. Interpreting the image sometimes presents a challenge, but it can be useful for deforestation analysis. Nevertheless, even though a complete coverage of the Amazon is available to the public, I am unaware of any land use studies that have used the data.

Survey and census data

Agricultural census

The last agricultural census (*Censo Agropecuário*) was conducted from August 1995 to July 1996, with the results published in 1998. The preceding census was done in 1985. The agricultural census surveys all farms in Brazil, collecting farm-related income and expenditure data, as well as data on agricultural operations during the year (inputs and outputs), and land use allocation.

Table 1 contains a summary of land use and agricultural productivity data from the 1996 agricultural census by ecological zone, from aggregated *município* data. We see that the part of the Legal Amazon that I have designated as the forest ecozone (using the WWF ecozones as the primary guide) has approximately 81 percent of the total land area, and 82 percent of the farm establishments. On average, however, farms are more than three and a half times larger in the *cerrado* ecozone, and a higher percentage of the land is converted to agricultural uses, so that the forest ecozone has only 56 percent of the land claimed for farms, and only 46 percent of the land that was or is now recently used in agriculture. The lower proportion of a farm converted for agricultural use in the forest ecozone may reflect a respect for the legal requirement to keep more land in forest in the forest ecozone than the *cerrado* ecozone, or it may simply reflect that with less vegetation in the *cerrado* eco

Table 1. Summary of Agricultural Activity in the Legal Amazon in 1996, by Ecozone

	Cerrado	**Forest**	**Total**
Area (hectares)	93,417,560	414,261,472	507,679,040
# of farms	145,770	676,886	822,656
Area of farms (hectares)	52,856,133	67,821,668	120,677,801
Farm area in total	56.6%	16.4%	23.8%
Mean farm size	363	100	147
Natural forest	28.4%	51.1%	41.2%
Natural pasture	22.5%	9.2%	15.0%
Planted pasture	30.1%	26.1%	27.9%
Cropland	5.8%	2.3%	3.9%
Tree plantations	0.1%	0.4%	0.3%
Perennials	0.2%	1.2%	0.8%
Fallow	2.1%	1.9%	2.0%
Abandoned	6.0%	5.1%	5.5%
Nonproductive	4.7%	2.5%	3.5%
Subtotal "natural"	*50.9%*	*60.3%*	*56.2%*
Subtotal actively used for agric.	*58.8%*	*39.3%*	*47.8%*
Subtotal formerly used for agric.	*8.1%*	*7.0%*	*7.5%*
Subtotal actively or formerly used	*65.6%*	*43.8%*	*53.3%*
Total cattle	15,782,492	20,313,332	36,095,824
– per hectare of pasture	0.57	0.85	0.70
– per farm	108	30	44
Value of production	2,219,251	2,782,321	5,001,572
– per farm (R$)	15,224	4,110	6,080
– per ha. of farm (R$/ha)	41.99	41.02	41.45
– per ha. of agric. land (R$/ha)	64.01	93.67	77.69
– per ha. of active agr. land (R$/ha)	71.46	104.30	86.63
Value of livestock production	826,142	1,217,021	2,043,163
– per hectare of pasture (R$/ha)	29.73	50.74	39.46
Value of crop production	1,393,108	1,565,300	2,958,409
– per hectare cropland (R$/ha)	426.61	582.11	496.83

Source: Dataset is from IBGE (1998b). Based on *município* data.
Notes:
1) Unless otherwise noted, values are in thousands of R$.
2) The exchange rate in June 1996 was US$1.00 = R$ 1.0013.

zone it is easier to use the land for agriculture. Later in this chapter, I use values of fallow and abandoned land from this table to estimate forest re-growth.

Table 1 also reports summary statistics dealing with cattle in the Amazon. We see that 56 percent of the cattle are in the forest ecozone, even though 54 percent of the pasture is in the *cerrado* ecozone. As a result, the stocking density per unit of pasture is higher in the former. On the other hand, since farms on average in the *cerrado* ecozone of the Legal Amazon are so much larger in land size, the average farm there has more than three and a half times the number of cattle found on the average farm in the forest ecozone.

We also see information on the gross value of agricultural production in Table 1. The mean gross value of production per farm is much higher in the *cerrado* portion of the Amazon, and while it is difficult to net out labor and other costs, it appears that this translates into greater per capita income in the *cerrado.* However, the forest biome farms are much more productive per unit of land, both in regard to livestock and crops.

Table 2 shows the distribution of farms by size in the two ecological zones of the Legal Amazon, based on aggregated *município* data from the 1996 agricultural census. Perhaps the most notable observation from this table is how the extremely large farms dominate the land use in the Amazon. Farms greater than 10,000 hectares in size hold 28 percent of the total farmland, and this is approximately the same regardless of which of the two ecozones one focuses on. Furthermore, we note that almost 53 percent of the total farmland is in farms greater than 2,000 hectares in size. Here we note a difference between the two ecozones: the *cerrado* has almost 58 percent in this size farm, while the forest only has 49 percent. This difference arose entirely from the farms in the 2,000 to 10,000 hectare range, which account for 31 percent of the land in the *cerrado*, but only 20 percent of the land in the forest ecozone.

Table 2. Farm Size Distribution, by Ecozone

	Cerrado	**Forest**	**Total**
# of farms		*percent of total*	
Undeclared size	0.5%	2.2%	1.9%
	cumulative % of total of declared farm sizes		
0 - 5	25.7%	39.6%	37.1%
5 - 20	35.9%	54.9%	51.5%
20 - 100	65.4%	85.7%	82.0%
100 - 500	88.0%	97.7%	95.9%
500 - 2,000	96.7%	99.4%	98.9%
2,000 - 10,000	99.6%	99.9%	99.8%
farm area (hectares)	*cumulative % of total of declared farm sizes*		
0 - 5	0.1%	0.6%	0.4%
5 - 20	0.4%	2.1%	1.4%
20 - 100	4.5%	16.2%	11.1%
100 - 500	18.4%	36.1%	28.4%
500 - 2,000	41.7%	51.3%	47.1%
2,000 - 10,000	72.6%	71.8%	72.1%

Source: Dataset is from IBGE (1998b).

Looking at the number of farms rather than the land area of farms, we see that almost 40 percent of the farms in the forest are smaller than 5 hectares, while just over one-quarter of the *cerrado* farms are in that size range. Similarly, the forest zone has 15 percent of its farms in the 5 to 20 hectare range, while the *cerrado* only has 10 percent.

The agricultural census provides some information on agricultural trends in the Amazon, since it is repeated every five to ten years. The census preceding the one in 1995-1996 was conducted in 1985. Unfortunately, this information is only available to me at the state level. There are essentially nine states in the Legal Amazon: Rondônia, Acré, Roraima, Amapá, Amazonas, Pará, Maranhão, Tocantins, and Mato Grosso. A tiny sliver of Goias is also in the Legal Amazon, but most researchers ignore the sliver, as they do the sliver of Tocantins that is outside of the Legal Amazon. While the majority of Maranhão is inside the boundary, enough of it is outside that when looking at state-level data, researchers are mixed on how to treat this state. Here we will treat it as part of the Amazon. For the purpose of trying to distinguish between true Amazon forest biome and *cerrado* biome that happens to be in the Legal Amazon, and when looking at only state-level data, the first six in the list will be considered Amazon forest states, and the latter three will be considered mixed *cerrado*-forest states.

Table 3 compares land use in 1985 and 1996. The first thing we note is that the number of farm establishments has decreased dramatically, while the total area of farms has increased only slightly. However, this is not to say that

Table 3. Farmland Use in the Legal Amazon, 1985 and 1996

Category	1985	1996
	All States	
# of farms	1,153,047	893,128
Area of farms (hectares)	115,950,634	120,759,203
Cleared area (sum next 5)	39,281,611	48,515,674
Cropland	6,114,674	5,737,591
Planted pasture	18,631,098	32,932,124
Tree plantations	220,075	349,832
Fallow (< 4 yrs old)	3,971,230	2,603,055
Fallow (> 4 yrs old)	10,344,534	6,893,072
Natural pasture	24,096,271	18,217,069
Forest	46,786,448	49,824,336
Unproductive land	5,786,304	4,202,124
Cattle	18,758,770	35,617,365
Stocking density (cattle / ha. Pasture)	0.44	0.70
	Forest States	
# of farms	496,393	401,262
Area of farms (hectares)	45,212,310	41,593,164
Cleared area (sum next 5)	12,402,468	14,279,911
Cropland	2,025,769	1,704,827
Planted pasture	5,824,165	9,485,653
Tree plantations	162,594	254,163
Fallow (< 4 yrs old)	1,325,006	732,175
Fallow (> 4 yrs old)	3,064,934	2,103,093
Natural pasture	4,401,371	3,822,813
Forest	26,652,451	22,466,464
Unproductive land	1,756,020	1,023,976
Cattle	5,361,795	12,058,479
Stocking density (cattle / ha. Pasture)	0.52	0.91
	Mixed Cerrado-Forest States	
# of farms	656,654	491,866
Area of farms (hectares)	70,738,324	79,166,039
Cleared area (sum next 5)	26,879,143	34,235,763
Cropland	4,088,905	4,032,764
Planted pasture	12,806,933	23,446,471
Tree plantations	57,481	95,669
Fallow (< 4 yrs old)	2,646,224	1,870,880
Fallow (> 4 yrs old)	7,279,600	4,789,979

Category	**1985**	**1996**
Natural pasture	19,694,900	14,394,256
Forest	20,133,997	27,357,872
Unproductive land	4,030,284	3,178,148
Cattle	13,396,975	23,558,886
Stocking density (cattle / ha. Pasture)	0.41	0.62

Source: Dataset is from IBGE (1998b). Based on state-level data.
Note: Forest states are Rondonia, Acre, Roraima, Amapa, Amazonas, and Pará. The mixed *cerrado*-forest states are Maranhão, Tocantins, and Mato Grosso.

the land use has remained static. What we observe is a sharp increase in cleared land (i.e., land that was not allowed to remain as natural forest or natural pasture[10]). The largest component of this change is in planted pasture, where we see a 77 percent increase. This increase was reflected in corresponding decreases in natural pasture, fallow, and abandoned land. Because these are aggregate statistics, it is not clear whether farmers brought into production land that was in these other land uses, or whether a large number of farmers abandoned their land entirely while at the same time, other farmers settled new land with larger quantities of planted pasture. The answer is that it was probably a mixture of the two, but the data limitations do not let us determine the answer.

---

[10] This is IBGE's definition of cleared land, which is also the one used by Andersen et al. However, in the regression work done in the following chapters, natural pasture is treated as cleared land, since it is likely to be degraded from its natural state if it is used as pasture.

We also see in Table 3 that changes in land use were more intense in the mixed *cerrado*-forest states. Farm area actually decreased in the forest states, while increasing by 12 percent in the *cerrado*-forest states. Cleared land increased more dramatically in the latter, especially in regard to planted pasture. More observations related to this table will be made in a later section dealing with approximating forest regrowth.

Figure 7 shows trends in total land in farm establishments from 1975 to 1996. We see that two states are set apart by their phenomenal growth—Rondônia and Mato Grosso. We also see that the two states that distinguished themselves with substantial negative growth—Amazonas and Acré—are among the forest states. Three other states—Amapá, Maranhão, and Tocantins—had virtually no change; and two other states—Roraíma and Pará—had total growth near 50 percent during the 21-year period.

Table 4 shows the gross value of agricultural production, contrasting *municípios* of the forest biome with those of the *cerrado*. Curiously, the value of cattle and other large animals (which, in fact, is almost made up entirely of cattle) is almost the same percent of total value in both biomes: approximately one-third of the total. Their value per head of cattle in the herd (not per head harvested) is also almost identical, differing by only 2.7 percent.

Table 4. Value of Production from Agricultural Census, 1996

	Forest		**Cerrado**	
	000s R$	**% of all**	**000s R$**	**% of all**
All farm production	2,782,321		2,220,989	
Cattle and other large animals	970,638	34.9%	736,004	33.1%
Medium animals	49,561	1.8%	25,605	1.2%
Chickens and other birds	196,822	7.1%	65,852	3.0%
Perennial crops	320,530	11.5%	45,232	2.0%
subtotal specified items	*178,807*		*25,883*	
percent of all	*55.8%*		*57.2%*	
banana	91,051	3.3%	20,252	0.9%
coffee	63,323	2.3%	2,670	0.1%
oranges	24,434	0.9%	2,961	0.1%
Annual crops	811,612	29.2%	1,316,801	59.3%
subtotal specified items	*697,826*		*1,285,985*	
percent of all	*86.0%*		*97.7%*	
soybeans	10,834	0.4%	756,848	34.1%
manioc	438,455	15.8%	14,133	0.6%
rice	154,403	5.5%	148,346	6.7%
sugarcane	16,909	0.6%	225,060	10.1%
maize	74,320	2.7%	121,996	5.5%
cotton	2,906	0.1%	19,602	0.9%
Horticultural crops	50,605	1.8%	6,388	0.3%
Silvicultural products	45,441	1.6%	2,105	0.1%
subtotal specified items	*43,659*		*0*	
percent of all	*96.1%*		*0.0%*	
Logs for paper	43,658	1.6%	0	0.0%
Extracted products	337,113	12.1%	23,004	1.0%
subtotal specified items	*255,190*		*18,668*	
percent of all	*75.7%*		*81.2%*	
Logs	152,451	5.5%	5,466	0.2%
Firewood	36,875	1.3%	8,876	0.4%
Charcoal	30,887	1.1%	4,325	0.2%
Açaí	34,976	1.3%	1	0.0%
Cattle (number)	20,313		15,816	
Value per head of all cattle	47.78		46.53	

Source: Dataset is from *Censo Agropecuário 1995-1996*.
Note: The exchange rate in June 1996 was US$1.00 = R$ 1.0013.

What is most notable in regard to agricultural production is that annual crop production represents about twice the percent of total production in the *cerrado* biome—59.3 percent—compared to the forest biome. Furthermore, soybeans represent 57 percent of the value of annual crop production in the *cerrado* portion of the Amazon, with sugarcane placing a distant second with 17 percent, and rice and maize following.

While the forest biome is not as heavily dominated by annual crops, they do represent 29.2 percent of the total value of agricultural production. Just as in the *cerrado*, annual crops in the forest biome are dominated by one crop—manioc, instead of soybeans—which provides 54 percent of the value of all annual crops.

While agriculture in the *cerrado* is dominated by cattle and annual crops (especially soybeans), the situation in the forest biome is much more diverse. Not unexpectedly, extracted products play an important role in forest agriculture, as do perennial crops and chickens. Since logs are considered an extracted product, one might actually be surprised to see that they only provide 5.5 percent of the value of agriculture in the forest biome. However, these tabulations are based on the agricultural census, which surveys farmers, and not loggers. Of the perennial crops in the forest, bananas are the most important, representing 28 percent of the value of all perennials harvested.

Annual surveys of agriculture and forestry

There are a number of annual surveys which provide insight as to the current trends in agricultural expansion. These include the *Produção Agrícola Municipal* (PAM), which among other things provides estimates as to the total land area in each of the major crops, as well as the annual harvest of each crop by *município*; the *Levantimento Sístematico da Produção Agrícola* (LSPA), which reports estimates of the most recent harvests, by month; the *Pesquisa Pequária Municipal* (PPM), which gives information on livestock numbers and production of livestock products such as eggs and milk; the *Pesquisa Trimestre de Abate de Animais*, the *Pesquisa Trimestre do Leite*, and the *Pesquisa Trimestre do Couro*, which give information on livestock processed at slaughterhouses, milk processed, and animal hides, by month; *Silvicultura*, which catalogs annual production of tree plantation products, such as charcoal; and the *Produção Extrativa Vegetal*, which gives annual information on extracted forest products, including logs. All of these are available on the web.[11]

Tables 5 through 7, and Figures 8 through 11 show trends in agriculture over the last decade. Figures 8 and 9 show trends in production of major crops in the forest and cerrado biomes of the Legal Amazon, respectively. Each figure consists of three graphs: the value of production from 1994 to 2000, and the hectares and quantity harvested from 1990 to 2000. While data on value of production are available back to 1990, exchange rate volatility in the early nineties and changes in Brazilian currency make converting these values to either constant Brazilian currency or U. S. dollars difficult.

Table 5. Yields of Important Crops, 1990 to 2000

year	Soybeans cerrado	Cotton cerrado	Rice cerrado	Rice forest	Sugarcane cerrado	Manioc cerrado	Coffee forest	Maize cerrado	Maize forest
1990	2.0	1.3	1.0	1.1	61.7	11.4	1.3	1.4	1.0
1991	2.3	1.2	1.6	1.4	61.4	11.5	1.2	1.8	1.1
1992	2.5	1.4	1.3	1.0	60.3	11.5	1.2	1.8	1.1
1993	2.4	1.5	1.2	1.2	62.0	11.0	1.2	1.9	1.2
1994	2.6	1.5	1.8	1.5	69.7	10.8	1.2	2.1	1.2
1995	2.3	1.3	1.9	1.4	70.2	11.0	1.2	2.2	1.2
1996	2.6	1.4	1.7	1.5	70.1	6.7	1.1	2.5	1.3
1997	2.7	2.0	1.9	1.4	73.5	11.8	1.0	2.4	1.3
1998	2.7	2.7	1.8	1.4	69.8	11.6	1.0	2.0	1.3
1999	2.8	3.2	2.3	1.6	67.8	12.4	1.1	2.4	1.4
2000	3.0	3.9	2.4	1.8	61.7	12.3	1.1	2.5	1.4

Source: PAM.

Another warning about the data is that the figures show abrupt changes between 1995 and 1996, especially for hectares harvested. I believe this means that the estimates were adjusted to the more accurate counts taken in the 1996 agricultural census. Andersen et al. tell us that the "PAM data are intended to provide timely information so they are not estimated from a random sample of producers, but are rather based on a 'subjective survey' of ex-

---

[11] http://www.sidra.ibge.gov.br/, last accessed July 2003.

perts' opinions" (pp 46-47). This statement, combined with my own observation of the data, indicates that the estimates likely reflect some type of predicted growth based on the preceding year estimates, and calibrated back to the preceding census.[12]

Table 6. Mean Prices of Important Crops, 1994 to 2000

year	Soybeans *cerrado*	Cotton *cerrado*	Rice *cerrado*	Maize *cerrado*	Maize forest	Sugarcane *cerrado*
1994	0.18	0.44	0.15	0.09	0.14	0.016
1995	0.14	0.40	0.14	0.10	0.14	0.023
1996	0.18	0.37	0.16	0.09	0.15	0.016
1997	0.20	0.47	0.17	0.08	0.14	0.020
1998	0.16	0.45	0.20	0.09	0.14	0.019
1999	0.11	0.34	0.12	0.06	0.10	0.010
2000	0.13	0.33	0.11	0.09	0.12	0.010

year	Manioc *cerrado*	Manioc forest	Logs forest	Bananas forest	Coffee forest	Peppercorn forest
1994	0.11	0.10	0.046	1.51	1.20	1.16
1995	0.16	0.12	0.046	1.74	1.32	1.25
1996	0.24	0.11	0.039	1.16	0.97	1.56
1997	0.14	0.09	0.034	1.10	0.98	3.55
1998	0.14	0.08	0.035	1.09	0.99	3.91
1999	0.09	0.05	0.026	0.70	0.79	4.28
2000	0.11	0.05	0.025	0.64	0.63	2.31

Source: Dataset is from PAM.

---

[12] See Nerlove (1958) for a similar critique of U. S. agriculture data.

The last comment regarding the structure of these graphs is that the quantity graphs reflect indices set to 100 for 1996. The indices were necessary because the units were not identical across crops (e.g., logs were in cubic meters, while manioc was in tons).

Table 7. Value of Production per Hectare of Important Crops, 2000

	Biome	**R$/ha**
soybeans	*cerrado*	401
soybeans	forest	377
cotton	*cerrado*	1,307
rice	*cerrado*	269
rice	forest	237
maize	*cerrado*	230
maize	forest	158
sugarcane	*cerrado*	619
manioc	*cerrado*	1,295
manioc	forest	633
banana	forest	1,778
coffee	forest	668
peppercorn	forest	5,685

Source: Dataset is from PAM.

While it is hard to detect at first from the graphs, the land devoted to annual and perennial crops has expanded more rapidly between 1996 and 2000 in the *cerrado* portion of the Legal Amazon, with an increase of 41 percent compared to the 27 percent of the forest portion. The rapid expansion in the *cerrado* is fueled mostly by the expansion of soybeans, which in the four years has expanded by 50 percent; while the increase in the forest is led by rice, which has expanded in the same years by 41 percent. Because soybeans dominate all other crops in the *cerrado* portion of the Amazon, their influence on total crop area expansion was much greater than the effect of rice in the forest portion, since maize and manioc occupy comparable areas.

It will be of great interest in coming years to observe the area devoted to cotton, since this area has quintupled over the last four years, leading it to be second only to soybeans in total value of production in the *cerrado* portion. In the forest portion, we note rapid increase in area devoted to soybeans (though the total area is still small), coffee (since 1998), and maize. I think it is reasonable to be skeptical about rapid increase in maize and rice, since these crops showed a marked adjustment downward between 1995 and 1996, indicating a propensity for agriculture offices to overestimate the area cultivated and harvested. The forest zone soybean increase is almost completely in Mato Grosso, and most of this is in four municípios located adjacent to the *cerrado* zone. Soybeans are constrained to areas with rainfall below 2,000 millimeters per year, which rules out most of the forest zone for its cultivation.

In Table 4, we saw that according to the agricultural census, logs only represented a small portion of total value of agricultural production in the forest portion of the Amazon. However, the annual survey includes logs from all sources, not just farmers. According to the estimates of the *Produção Extrativa Vegetal*, the value of logs extracted in this biome totaled R$1.67 billion, far exceeding the value from the agricultural census of R$0.15 billion, and equivalent to 63 percent of the value of all agricultural production (excluding extracted logs), or 39 percent of agriculture plus logging.

tracted logs), or 39 percent of agriculture plus logging. However, note the dramatic decline in logging after 1995, with total volume in 2000 equal to only 29 percent of the volume of 1995.

All other things being equal, we would expect that trends in crop values and quantities would be similar to those for hectares harvested. However, we find many differences. Tables 5 and 6 can help us better understand the reason for these differences, because the reason must be related to either yield or price changes. One thing we note from Table 5 is the increase in soybean yields from 2.0 tons per hectare in 1990 to 3.0 tons per hectare in 2000. A more startling contrast appears when observing trends for cotton in the *cerrado*: yields are flat between 1990 and 1996, and then suddenly in 1997 there is a 43 percent increase, followed by a stream of dramatic increases through the end of available data. In 1996, yields were 1.4 tons per hectare, but by 2000, the yields averaged 3.9 tons per hectare.

While less remarkable, we note that in 1990, maize yields were similar in forest and *cerrado*, 1.0 and 1.4 tons per hectare; yet by 2000, there is large difference, with the *cerrado* averaging 2.5 tons per hectare, and the forest averaging only 1.4. We note a similar trend for rice, 1.0 tons per hectare and 1.1 tons per hectare for *cerrado* and forest in 1990, to 2.4 and 1.8 in 2000. Finally, we note only small trends for the other crops: sugarcane appears to have increased yields through 1999, but then experienced a poor harvest in 2000; manioc increased only slightly; and coffee appears to have actually declined.

Table 6 shows the effects of exchange rates on prices, with the prices for all crops but one dramatically lower in 1999 and 2000 compared to the preceding years. The one exception was for peppercorn, which perhaps has a price set by national supply and demand, rather than by the international market.

Table 7 shows the average value of major crops per hectare. Peppercorn leads the list at $5,685 per hectare. This is followed by bananas, cotton, and manioc. Among the grains and legumes, soybeans leads the pack. It is not difficult to see why the area of some crops is expanding. The anomaly here is manioc, because it has high value and is easy to grow. Perhaps it has been overvalued by the surveys, due to high self-consumption but low purchases and sales.[13]

Figure 10 shows the differences between herd growth in the forest and *cerrado* biomes of the Amazon. We see that the herd in the forest is growing much more rapidly than the one in the *cerrado*—37 percent to 14 percent between 1996 and 2000—an annual rate of 8.2 percent against 3.2 percent. Fig-

---

[13] That is, perhaps there are significant quality differences between manioc which is sold and that which is self-consumed, and the self-consumed manioc is over-valued by IBGE researchers, because it is valued as if it were the high-quality manioc sold on the market.

ure 11 shows the differential growth in cattle herds across states. We note a 52 percent increase in herd size in Pará, between 1996 and 2000; a 43 percent increase in Rondônia; and a 42 percent growth rate in the Mato Grosso forest area.

Population census

Table 8 shows some of the trends in population growth between 1991 and 2000, broken down by forest and *cerrado* ecozones within the Legal Amazon. Perhaps most surprising to many people is the degree of urbanization in the Amazon, with almost 80 percent of the population in the *cerrado* portion living in urban areas in 2000, and almost 70 percent of the population in the forest portion living in urban areas.[14] Both are increases of approximately 15 percent since 1991. While the overall population of the Amazon expanded at a rate of 1.7 percent per year during that nine-year time span—adding 4.3 million people—the size of the rural population declined in both biomes: 4.1 percent per year in the *cerrado*, and 2.5 percent year in the forest.

Since 1991, the two fastest growing Amazon states have been Amapá at 5.6 percent per year, and Roraíma at 4.5 percent per year. In both cases, changes in rural population have been small, with the bulk of the population change observed in annual urban growth rates of 6.6 and 6.3 percent annu-

---

[14] Designations of “urban” and “rural” are from IBGE. Areas designated in each category are updated for each census or population count.

ally. Other states with large changes in urban populations are Maranhão, Tocantins, and Pará, with rates of 5.9, 5.3, and 5.1 percent per year; while at the same time observing rapid rural depopulation at rates of 2.9, 3.0, and 2.4 percent per year.

Table 8. Population Trends in the Legal Amazon

	Cerrado	**Forest**	**Total**
Urban population, 2000	2,787,006	10,666,085	13,453,091
Rural population, 2000	796,509	4,808,308	5,604,817
Percent urban, 2000	77.8%	68.9%	70.6%
Urban population, 1991	2,074,438	7,079,774	9,154,212
Rural population, 1991	1,159,393	6,057,574	7,216,967
Percent urban, 1991	64.1%	53.9%	55.9%
Rate of change in urban	3.3%	4.7%	4.4%
Rate of change in rural	-4.1%	-2.5%	-2.8%
Rate of change in total	1.1%	1.8%	1.7%
Mean population per sq km, 2000	3.8	3.7	3.8
Mean population per sq km, 1991	3.5	3.2	3.2

Source: Data are from the 1991 and 2000 population censuses.

In the period between 1980 and 1991, Rondônia and Roraima experienced very high total population growth rates of 7.4 and 8.8 percent per year, with rapid urban growth of 9.2 and 9.4 percent per year. Even during that period, the states of Amapá, Tocantins, and Acre were experiencing rural depopulation, with annual rates of decline of 2.6, 1.2, and 0.7 percent. During the period from 1970 to 1980, Rondônia experienced the most rapid inter-census population growth rate of any Brazilian state since 1950, when it grew 14.6 percent per year, more than quadrupling in size to just over half a million people. In fact, between 1970 and 2000, Rondônia had almost a 12-fold increase in population. During the same period, Roraíma increased almost eight times its 1970 size, though even in 2000 it had slightly less than a third of a million people. Mato Grosso has quintupled between 1970 and 2000, and both Pará and Amazonas have almost tripled in the same time frame. Taking the nine states of the Legal Amazon as a unit, they have grown from 7.8 million people in 1970 to 21.0 million people in 2000, yet the rural population has only grown from 5.0 million to 6.7 million.

Caution is needed when interpreting rates of urbanization. For example, between 1991 and 2000, the *município* of São Luís more than tripled its urban population, yet its total population increased by less than 25 percent. This appears to have resulted from the census redesignating large areas as urban, and not the case of population influx or internal migration (though it is not impossible that in the 832 square kilometers represented by this *município,* there was some internal migration).

The rapid growth of the urban population in the Legal Amazon seems to support Faminow's hypothesis that much of the herd growth and pasture expansion in the Amazon was in response to the increased urban demand for beef, due both to rapid population growth and increased income of urban residents.[15] Given the total herd growth rate of 6.1 percent per year for all of

---

[15] Kaimowitz (2002) criticizes Faminow for overstating his case, but the basic point of Faminow's argument is correct: local demand for beef contributed to

the Legal Amazon between 1996 and 2000, it seems that the 2.9 percent per year population growth rate of the nine Amazon states lends support to Faminow's hypothesis.

Table 9 shows the list of cities within the Legal Amazon that according to the 2000 census had urban populations of 100,000 or more. In addition to reporting the total and urban population of each *município* in the table, I included a broader measure of each municipal area by summing all of the urban population within the census microregion (a collection of *municípios*). This broader measure is important for places like Belém, where there are 1.3 million urbanites in the *município*, but a total of 1.8 million, when including the metropolitan area. The Belém metropolitan area—urban and rural together—grew at an annual rate of 2.8 percent between 1991 and 2000, while the other Amazon megacity, Manaus, grew at an annual rate of 3.5 percent over the same period[16].

---

the expansion of beef supply. During the 1970s, urban growth in Amazon states was 6.5 percent; in the 1980s, it was 5.1 percent; and in the 1990s, 4.7 percent. High urban growth rates such as these could explain a significant portion of the growth in the cattle population. While I do not have figures for cattle populations in the 1970s and 1980s, between 1996 and 2000 the cattle population grew by 6.1 percent annually.

[16] For more on urbanization in the Amazon, see Browder and Godfrey. They explore the mechanics of urbanization in the Amazon in great detail, dividing urbanization based on its causes.

Forest Regeneration

Table 9. Growth of Large Cities in the Legal Amazon

Name	Urban population, 1991	Urban population, 2000	Total population, 2000	Urban population of metropolitan area, 2000
Manaus	1,006,585	1,394,724	1,403,796	1,472,912
Belém	849,187	1,271,615	1,279,861	1,781,871
São Luís	246,244	835,325	868,047	874,931
Cuiabá	395,662	476,178	483,044	706,308
Ananindeua	74,051	391,994	392,947	see Belém
Porto Velho	229,788	273,496	334,585	315,105
Macapá	154,063	270,077	282,745	360,770
Rio Branco	168,679	226,134	252,885	249,701
Imperatriz	210,051	218,556	230,451	382,427
Várzea Grande	155,307	211,283	215,276	see Cuiabá
Boa Vista	120,157	196,942	200,383	205,691
Santarém	180,018	186,567	262,721	251,341
Rondonópolis	113,032	141,660	150,049	194,661
Marabá	102,435	134,258	167,873	155,597
Palmas	19,246	133,877	137,045	195,490
Castanhal	92,852	121,198	134,442	173,281
Araguaína	84,614	105,822	113,090	182,904
Caxias	84,331	103,416	139,689	122,744

Source: Data are from the 1991 and 2000 population censuses.
Notes:
1) Population is by *município*.
2) Metropolitan area population is computed by microregion.

We see from Table 1 that in the forest biome roughly 2 percent of farm area is in fallow, and that between 5 and 6 percent is vacant or abandoned (meaning that it has been in fallow for more than 4 years). We can try to use these numbers as a first pass at computing the magnitude of forest regrowth, and as such, perhaps provide a crude correction to the gross deforestation calculation provided by INPE, in order to derive a measure of net deforestation.

In the forest ecozone, when cropland, plantations, perennials, and planted pasture are added together, we see that they total 30.0 percent of total farmland (ignoring natural pastures). If we want to answer the question, “What percentage of already cleared farmland is abandoned each year?”, we might begin by reasoning that since the category of fallow covers up to four years, and since fallow represents 1.9 percent of farmland, then maybe 0.5 percent of farmland is abandoned each year. If that is true, then the land in the abandoned category could possibly be land that is aged between 5 and 14 years of abandonment, since in the forest ecozone this is 5.1 percent of farmland, and dividing by 0.5 percent we get approximately 10 years. There are, of course, so many implicit assumptions in this calculation that we do not want to take it too seriously, except as a sort of “back of the envelope” calculation.

Since we have gone this far with the crude estimation, we should at least finish it. Approximately 30 percent of farmland is actively used for agriculture (net deforestation) and 37 percent is now or once was used for agriculture (gross deforestation). Since 0.5 percent per year is abandoned, then the annual rate of forest regrowth on farms is 0.5 divided by 37, or 1.3 percent of gross on-farm deforestation. If this calculation is repeated for the *cerrado* portion of the Legal Amazon, the result is 1.2 percent of gross deforestation allowed to regrow each year.

I have already cited INPE's annual report that as of August 2001, 603,514 square kilometers of gross deforestation had occurred[17]. If forest regrowth off-farm occurs at the same rate as that which occurs on active farms (i.e., farms that have not been abandoned), then there are possibly 7,800 square kilometers of reforestation every year. This is very significant, especially in comparison to the average annual gross deforestation of 17,164 square kilometers.

Table 1 shows that on-farm gross deforestation (planted pasture, cropland, plantations, perennials, fallow, and abandoned) in the forest ecozone only accounts for approximately 250,000 square kilometers of gross deforestation. The value computed by INPE for August 1996 (when the agricultural census ended) was 517,000 square kilometers. So what happened to the other 267,000 square kilometers of gross deforestation? It is possible that not all land in the *municípios* that are part of the *cerrado* biome are interpreted as *cerrado* from satellite imagery. That is, perhaps we should count some of the land from the *cerrado municípios* in gross deforestation for the Amazon. But how much? If we only counted fallow and abandoned land as land that was

---

[17] According to Table 3, cleared area in 1996 totaled 485,000 square kilometers, using the traditional definition of clearing, which I used in the calculations above. If we limit our focus to the forest biome, this figure drops to 143,000 square kilometers.

once forested, we add another 43,000 square kilometers to gross deforestation from farms.

Some can be accounted for by pointing to deforestation that did not take place on farms (such as urban and suburban areas, and deforestation near mines), though quantitatively this amount should be small. Another perhaps larger amount of this can be accounted for by farm abandonment. Table 3 indicates that in 1996 there were around 19 percent fewer farms in the forest ecozone states than in 1985. This number is the net change in farm establishments, so the abandonment rate may have been higher, since there were likely new farms established during that period, offsetting the gross abandonment. Unfortunately, this figure is further muddied by considering the possibility that groups of small farms may have been consolidated into larger farms. More important, however, is that Table 3 also shows that the cleared farmland between 1985 and 1996 only increased by 18,700 square kilometers, which is only slightly greater than the average annual gross deforestation reported by INPE. Even if we acknowledge that the rate of deforestation is high in the forest ecozone portion of what I have called the "mixed *cerrado*-forest states", Table 3 shows that for all Amazonian states, the area cleared is only 92,000 square kilometers greater in 1996 than in 1985, which is only 8,400 square kilometers per year, less than half of the gross deforestation computed by INPE, and in reasonable agreement with the calculations above.

To shed more light on the concept of gross deforestation versus net deforestation, a look at the TRFIC data would be helpful. The results are in Table 10. The table shows that as of 1992, satellite imagery analysis showed that there were 239,000 square kilometers of deforestation, and 126,000 square kilometers of regrowth. The TRFIC dataset was missing about seven and a half scenes, representing approximately 240,000 square kilometers of total land (mostly in Amapá, but also in the heart of Roraima and the fringes of Pará and Tocantins), and of course, unknown quantities of deforestation and regrowth. Nevertheless, adding deforestation and regrowth together gives 365,000 square kilometers of measured gross deforestation in 1992, and INPE reports 440,000 square kilometers of gross deforestation for the same year. That is to say, the TRFIC and INPE values are in reasonable agreement. If we are to believe the TRFIC data, approximately one-third of gross deforestation is reforesting! If we assume that it takes 15 years for forest regrowth to be indistinguishable from native forest in satellite imagery, then we find that annual regrowth must be around 8,400 square kilometers, which is in remarkable agreement with the amount of regrowth estimated from agricultural census calculations.

Brondízio et al. find that in the 3,718 farm lots in their study area west of Altamira along the Transamazon Highway (in the *municípios* of Altamira, Brasil Novo, and Medicilândia), 20 percent of the area was in secondary forest

Table 10. Deforestation and Regrowth, 1992

State	Forest	Deforest	Regrowth	Water	Cloud or shadow	*Cerrado*	All
Pará	960,100	73,698	48,948	31,588	33,020	43,996	1,191,350
Maranhão	83,447	45,964	24,280	2,053	8,902	104,464	269,108
Amazonas	1,391,705	9,000	10,037	51,692	80,461	37,754	1,580,648
Rondônia	175,205	22,056	10,004	2,102	12,590	16,250	238,207
Acré	143,448	6,508	1,925	385	1,210	113	153,587
Mato Grosso	467,113	65,052	11,187	6,840	6,617	346,098	902,906
Tocantins	35,859	16,013	331	1,933	1,234	216,687	272,057
Subtotal	*3,256,877*	*238,290*	*106,711*	*96,592*	*144,033*	*765,362*	*4,607,865*
Roraima	170,078	710	19,259	38	1,930	1,585	193,602
Amapá	43,287	22	17	246	12,818	2,248	58,638
Total	***3,470,242***	***239,022***	***125,987***	***96,876***	***158,782***	***769,195***	***4,860,104***

Source: Author's tabulation of GIS data provided by the Tropical Rainforest Information Center (TRFIC).

succession, while 17 percent was in production. Their analysis distinguished age of plot (i.e., date of arrival of farmer). They found that across all arrival cohorts, secondary succession medians were larger than productive agricultural land medians; and that there was a trend to have larger areas in production and larger areas in secondary succession the earlier the farmer settled in the region. That is, it seems to be part of the normal land management system of farmers in this region to have more than 50 percent of the land in forest regrowth. This result is stronger than and supportive of the macro-level results that I just reported using both the satellite data from the Tropical Rain Forest Information Center (TRFIC) and the agricultural census data.

Brondízio et al. also report that there is a tendency to use land with shorter regrowth periods for pasture, and land with longer regrowth periods for crops. They also write that "the ability to balance the amount of fallow in different stages of regrowth is an important element of farm management in the frontier" (p. 158). The implication of this statement is that there is likely to be little or no forest regrowth on active farms that is not destined to be cleared again later. This changes how I think about regrowth, at least about that which is on active farms: it is not area that is destined to return to native forest, but land that is "scheduled" (however flexible that schedule might be) to be cleared again for productive agriculture.

Deforestation and Accessibility

One of the most compelling reports dealing with Amazonian deforestation and accessibility by roads comes from Diógenes Alves, a researcher at INPE. He compares deforestation maps from the 1973 to 1978 period to deforestation maps from the 1991 to 1996 period, and discovers that 47 percent and 73 percent of deforestation between the two periods took place within 25 and 50 kilometer bands around major roads (mostly federal roads, but also some important state roads). He also noted that 86 percent of new deforestation occurred within 25 kilometers of areas which were deforested by 1978. This seems to imply that major roads and early deforestation are also highly correlated. Since some Amazonian cities were settled in the 1600s, it may be that many of these patches of early deforestation (i.e., those noted by satellite for 1973 to 1978) represent well-established towns and cities. Since major roads connect the well-established towns and cities, it is difficult to separate the effects of accessibility from the effects of urban sprawl and providing food for the local urban centers.

Alves also reports that he was unable to include in his report information on 30 percent of the Amazon due to cloud cover. Depending on the correlation between deforestation, roads, and settlements, on the one hand, and cloud cover on the other, there might be some significant error in the figures just cited.

Other researchers reproduce Alves's tabulations with other datasets (see Laurance et al. 1991), but they do not consider nearness to earlier settlement in their analysis, and therefore likely attribute a much heavier weight to accessibility than they should. They also ignore official government settlement and relocation schemes, which are usually near roads, and which interfere with the ability to predict deforestation apart from government intervention.

Since I have the TRFIC data for 1992, as well as a GIS road map and a GIS human disturbance (Portuguese: *antropísmo*) map for 1976 (from IBGE 1997a, which I believe was the one used by Alves), I decided to test the influence of roads versus previous deforestation and settlement. Table 11 shows the results of the cross-tabulation. As the reader can see, deforestation is more likely to occur near roads and near *antropísmo*, rather than far from them. In order to get some idea of which has a stronger effect, we see that being less than 25 kilometers from a major road but greater than 100 kilometers from *antropísmo* shows only 3.7 percent deforestation, while being less than 25 kilometers from (but not inside) an area disturbed by humans by 1976 but greater than 100 kilometers from a major road shows 9.1 percent deforestation. In fact, the table shows a general trend that deforestation rates are almost identical for both roads and *antropísmo* up to 50 kilometers away from the actual road or disturbance, but then the influence of *antropísmo* fades less rapidly than the influence of roads at distances greater than 50 kilometers.

These results seem to point to the tendency for new deforestation to take place near older deforestation. That is, new deforestation was more likely to take place near the 1976 deforestation than far away from it. The table also seems to imply that a major road built in an area not previously deforested may not lead to a high amount of deforestation. We must be cautious before we embrace this interpretation. The reason for caution—not even taking into consideration the potentially grave consequences to tropical forests if this hypothesis is wrong—is that we have not accounted for the way in which major roads may have led to the pre-1976 deforestation, and so—referring back to Table 11—it is possible that the long-run expectation of the effect of roads is 20.5 percent deforestation within 25 kilometers of a road, rather than 3.7 percent (the value for deforestation within 25 kilometers of a road, but greater than 100 kilometers from previous deforestation). A second reason for caution is that we have not controlled for other factors that may influence the suitability of a site for timber harvesting or agriculture. If one of the other factors were likely to influence deforestation, building a road through such an area could lead to high deforestation rates, even if it is not near previous deforestation.

Table 11. Effect of Major Roads and Pre-1976 Human Disturbances of the Land on Deforestation

	Road category				
Human disturbance category	**< 25 km**	**25-50 km**	**50-100 km**	**> 100 km**	**Total**
			proportion deforested		
Inside area	0.433	0.342	0.248	0.183	0.375
< 25 km	0.289	0.225	0.198	0.091	0.215
25-50 km	0.237	0.143	0.108	0.033	0.119
50-100 km	0.134	0.073	0.044	0.023	0.050
> 100 km	0.037	0.012	0.010	0.007	0.010
Total	0.205	0.119	0.063	0.014	0.062
			area of forest + regrowing + deforested (000s of square kilometers)		
Inside area	64.1	25.9	15.8	4.6	110.4
< 25 km	162.8	103.5	95.2	93.9	455.3
25-50 km	76.3	95.6	93.4	118.6	383.9
50-100 km	80.6	80.8	221.9	275.5	658.8
> 100 km	149.7	144.3	296.8	1,631.1	2,221.9
Total	533.5	450.0	723.1	2,123.7	3,830.4

Source: Author’s tabulation of GIS data provided by the Tropical Rainforest Information Center (TRFIC) for 1992.

Table 12 shows the results of the cross-tabulation of distance to nearest pre-1996 government settlement scheme vs. distance to nearest major road. These results must be interpreted with caution because the deforestation data are from 1992, but the time period notation for the settlement data only indicated "before 1996" and "1996 or later" as the categories. Thus we can at least limit the study to pre-1996 settlement areas, but some of these settlements will have been established after 1992. Nevertheless, since many of them were established before 1992, we may use this as an indicator of the importance of settlement areas in causing deforestation.

Table 11 seemed to show that new deforestation is more likely to happen near old deforestation, and this table seems to support that hypothesis (in the sense that we would naturally assume a settlement area would be deforested, and here we note that even being near—say less than 50 or 100 kilometers—leads to higher deforestation rates). If we believe this interpretation, then the implication is that government land settlement schemes actually open new patches of deforestation that over several years may become extremely large patches of deforestation—opening up areas that might never have been deforested, had market processes been left to themselves! As before, I must urge caution in adopting this interpretation, because not only were there problems with the data, but we also need to consider the possibility that settlement schemes are planted in areas already undergoing deforestation, rather than being the cause of deforestation.

Table 12. Effect of Major Roads and Government Settlement Schemes on Deforestation

	Road category				
Settlement category	**< 25 km**	**25-50 km**	**50-100 km**	**> 100 km**	**Total**
	proportion deforested				
Within, or < 25 km	0.311	0.210	0.129	0.075	0.201
25-50 km	0.202	0.138	0.087	0.042	0.112
50-100 km	0.084	0.053	0.036	0.024	0.038
> 100 km	0.044	0.010	0.014	0.003	0.007
Total	0.205	0.119	0.063	0.014	0.062
	area of forest + regrowing + deforested (000s of square kilometers)				
Within, or < 25 km	226.4	144.9	160.8	117.4	649.5
25-50 km	136.4	120.2	161.5	163.2	581.3
50-100 km	92.9	104.9	237.5	414.0	849.3
> 100 km	77.9	79.9	163.4	1,429.1	1,750.3
Total	533.5	450.0	723.1	2,123.7	3,830.4

Source: Author’s tabulation of GIS data provided by the Tropical Rainforest Information Center (TRFIC) for 1992.

To test further whether previous deforestation can lead to greater deforestation, we can look at another cross-tabulation contained in Table 13: that of cities with populations larger than 100,000 (as of 1996) vs. major roads. Here we see a very strong effect of large cities on deforestation. In the less than 25-kilometer category, this is perhaps as much or more from urban and suburban development than from agricultural activities. However, as we go farther away from the city, we would expect that most deforestation is due to agricultural and forestry activities, perhaps in supplying goods for that particular city. This cross-tabulation seems to confirm the hypothesis that deforestation tends to lead to further nearby deforestation.

Table 13. Effect of Major Roads and Cities with 100,000 People on Deforestation

	Road category				
Settlement category	**< 25 km**	**25-50 km**	**50-100 km**	**> 100 km**	**Total**
		proportion deforested			
< 25 km	0.426	0.247	NA	NA	0.423
25-50 km	0.254	0.157	0.079	NA	0.215
50-100 km	0.213	0.130	0.094	0.017	0.152
> 100 km	0.194	0.116	0.061	0.014	0.055
Total	0.205	0.119	0.063	0.014	0.062
	area of forest + regrowing + deforested (000s of square kilometers)				
< 25 km	13.0	0.2	0.0	0.0	13.2
25-50 km	23.7	15.0	0.3	0.0	39.0
50-100 km	63.1	46.1	45.6	1.8	156.6
> 100 km	433.8	388.6	677.2	2,121.9	3,621.5
Total	533.5	450.0	723.1	2,123.7	3,830.4

Source: Author's tabulation of GIS data provided by the Tropical Rainforest Information Center (TRFIC) for 1992.

Table 14 shows the results of regressing the proportion deforested on road, *antropísmo,* and settlement distances, in the first set of regressions; and road and *antropísmo* distances alone (excluding settlement in case it is endogenous or the data are faulty), in the second set of regressions. The regressions confirm the findings of Table 11 and Table 12: the effect of building a new road on deforestation is potentially small, as long as there is not some type of “seed deforestation” already present or about to be implanted by a new settlement scheme.

Table 14. OLS Regressions on Proportion Deforested Using TRFIC 1992

Variable	Coef.	t-stat	Coef.	t-stat
major road, < 25 km	0.0751	6.433	0.1023	6.864
major road, 25-50 km	0.0204	1.721	0.0411	2.684
major road, 50-100 km	0.0025	0.252	0.0165	1.314
1976 *antropísmo*, within	0.2698	12.411	0.3054	10.701
1976 *antropísmo*, < 25 km	0.1390	11.388	0.1675	10.795
1976 *antropísmo*, 25-50 km	0.0669	5.403	0.0866	5.397
1976 *antropísmo*, 50-100 km	0.0154	1.553	0.0287	2.248
settlement scheme, within or < 25 km	0.0932	7.964		
settlement scheme, 25-50 km	0.0361	3.170		
settlement scheme, 50-100 km	0.0012	0.134		
constant	-0.0062	-1.177	-0.0022	-0.340
log likelihood	202.0		171.2	

Source: Author’s tabulation of GIS data provided by the Tropical Rainforest Information Center (TRFIC) for 1992.

Notes:

1) Regressions weighted by area of forest, regrowing, and deforested.

2) Each of the 99 observations in the regressions represent the sum of all gridcells with the unique combination of values from each of the three categories. For example, one of the categories is “within 25 km of a major road, 25 to 50 km from *antropísmo*, and greater than 100 km from a settlement scheme”.

Chapter 2: Agriculture and Deforestation

Brief Introduction

The preceding chapter gave a general overview of some issues involved in studying Amazon deforestation, including a review of various types of deforestation data, the importance of distinguishing between gross and net deforestation, the difference between deforestation measured by the agricultural census and satellite data, changes in agriculture and demographics over the past several decades, and the importance of distinguishing between the influence of roads and that of previous deforestation.

This chapter continues the work of the previous chapter in laying some important foundations for the rest of the dissertation. These include reviewing many of the datasets which provide variables used to explain deforestation, and providing a simple static and a simple dynamic model of agricultural land clearing. The main contribution of this chapter is a number of tobit regressions examining the impact of roads, cities, and previous settlement on agricultural land use (i.e., deforestation on farms), controlling for agroclimatic factors. These regressions form the basis of later regressions in Chapter 5.

This chapter begins with a defense of the agricultural census as a source of data for a study of deforestation, then follows with a brief review of similar studies. Then it describes a number of datasets used to supply explanatory variables. That section is followed by a presentation of the two models of deforestation. The chapter concludes with a presentation of the regression results.

Using the Agricultural Census to Study Deforestation

In Chapter 1, I cited the difference between timber harvested on farms and total timber harvested. The difference is quite large, and points to the importance of loggers in the removal of trees. On the other hand, farmers remove many trees using fire—trees that are not reported in timber harvest figures, because they are not sold or used for any productive purpose other than their ash providing some valuable nutrients to the soil. While there have been some studies of the relationship between logging and ranching or farming that show that loggers are important for the process of opening up land for settlers, not all land that is logged gets settled (Veríssimo et al. 1995).

Land that has been logged but not settled in many cases has a reasonable chance of having trees regrow. Regrowth is possible for a number of reasons. Firstly, loggers are often selective in their harvesting. In the first wave, they tend to take the high-value species such as mahogany. Then they may come back later for the medium-value species, and finally even later for the low-value species (Uhl et al.). Or, they may not come back. Secondly, even if the land is clear-cut, trees can still regrow, especially if there are forest stands nearby—though the land may have a lower vegetative cover for a longer period of time than with selective harvesting.

On the other hand, even selective logging changes the dynamics of the forest. Much of the understory that is normally protected by the taller trees becomes exposed to sunlight after logging and dries out, making the forest more susceptible to fire, especially during El Niño events (Nepstad et al. 1999, Cochrane et al. 1999).

What I am trying to argue here is that while I acknowledge the environmentally undesirable consequences of logging, the cutting of trees is not the best indicator of the type of deforestation that is of greatest concern. Rather, deforestation that is the most significant environmentally is the settling of land, with the conversion of that land to ranching or farming.[18] For that type of measure of deforestation, the agricultural census is well-suited. Furthermore, it has the advantage that clouds do not conceal land use—as is the case for satellite images—and it can be used to understand land conversion and use in *cerrado* areas, where satellites have difficulty detecting land conversion. A disadvantage is that the data available for analysis are aggregated, so that instead of knowing what is taking place on each 0.09 hectare gridcells—as we do with satellite images—we know what takes place on average in 558,000

---

[18] Even in the case of fires destroying forests, most destructive fires are started by farmers and ranchers who lose control of them.

hectare pieces of land. Nevertheless, we are interested in determinants of land use for agricultural purposes, and the 6,776 units of observation that are used in this study give us a significant amount of information. I will refer to the units of observation as census tracts, though for the purposes of this study I combined many of the extremely small census tracts with their neighbors. This combination was necessary because the agricultural census records the principal location of each farm, but many farms are located in multiple census tracts. The magnitude of this effect is reduced when the smallest census tracts are combined with neighbors. Furthermore, a number of these tiny census tracts represent village or town centers, and have nothing to do with agriculture. By merging, I reduce the census tracts from the original 10,032 to 6,776.

This chapter is related to the work I did with Ken Chomitz (Chomitz and Thomas 2001 and forthcoming). It differs in the following respects.

- I improved upon the river network, including rivers from the Tocantins and northern coast basins, and included rivers that are used for transportation, but are too small for ocean-going freighters.
- I accounted for the *varzeas*—seasonally flooded areas which have special significance for settling (a deterrent) and grazing (an advantage)—taking special care to distinguish the Marajó *varzea.*
- I tested the importance of nearness to coasts.

- I included the parts of Maranhão excluded by Chomitz and Thomas because the rainfall data did not extend past 45 degrees west.
- Because Chomitz and Thomas found high significance of previous deforestation, I also controlled for deforestation that might have been initiated by government settlement schemes.
- I tried the same for mineral deposits, trying to determine whether mining served as an initiator of deforestation due to agricultural settlement.
- I tested for the "Manaus" effect—that Manaus is truly a different kind of city, growing to its size because of the tax-free zone the government established there in 1967.
- I generated a new principal road network that tries to eliminate potentially endogenous roads included in our earlier study.
- I recomputed non-water areas, based on the new river network.
- I used the WWF ecozones instead of the IBGE / IMAZON vegetation zones.
- I used a weighted heteroscedastic tobit instead of the unweighted heteroscedastic tobit used previously, this time weighting by census tract area, rather than allowing census tract area to "choose" its own weighting based on the likelihood function.

- I dropped nearness to towns of greater than 25,000 people from the analysis, believing these to be endogenous.
- I updated the location and sizes of cities used in the analysis, based on the 1996 count of the population.
- Finally, I ran the regressions on the forest and *cerrado* ecozones of the Legal Amazon separately.

This chapter is also similar to the work done by Pfaff (1999), but differs in more significant ways than it does from the Chomitz and Thomas work. First of all, Pfaff only has 240 units of observation, and those in areas which are mostly forest (he even excludes the northern *cerrados*, which I include as part of the forest ecozone). This means that the size of his average unit of observation must be around 14 million hectares—or approximately 26 times the size of the unit of observation used in this chapter. Secondly, he uses aggregated satellite data. Third—and this I believe to be the main advantage of his data—he had a panel of observations. However, the explanatory data I use are much more detailed, and I have removed some of the most endogenous data.

There have been a number of papers that have used GIS data and deforestation data in the Brazilian Amazon, but most of these data focus on small portions of the region (Mertens et al.; Brondízio et al.; McCracken et al.; Walker et al.; Caldas et al.). While this work is invaluable for studying the processes involved in deforestation, it is not useful as a predictor for deforestation caused by infrastructure being built through portions of the Amazon not previously settled, because of sample selection bias: the study areas were chosen specifically because of some type of rapid, on-going deforestation process.

Data Used in Analysis

While most of the agricultural census data at the *município* level are available on the web—and that which is not available can be purchased from IBGE on CDROM—the census tract dataset that I use in this chapter and Chapters 3 and 5 is not publicly available. I am therefore extremely grateful to Sérgio Besserman Vianna of IBGE for granting me use of the census tract data, which included land use type, value of production, and gross farm income and expense. I also used the GIS files for the census tracts provided by IBGE.

Rainfall data was critical to this study. Jeffrey Richey and Miles Logsdon of the University of Washington EOS Amazon project provided the precipitation layer, based on interpolation of the ANEEL (the Brazilian meteorological and hydrological agency) rainfall gauge data. The dataset they kindly shared was in 0.2 degree gridcells, and covered most of the Amazon, except in the east where it stopped at 45 degrees West. I extrapolated the data using quadratic approximation in GAUSS. I also interpolated the set to 5 kilometer resolution, using double linear interpolation.

Hari Eswaran and Paul Reich of the USDA World Soil Resources kindly provided the soil limitations dataset, which is based on the FAO world soils dataset at 1:5,000,000 scale. The World Wildlife Federation ecozones presented in Chapter 1 are used in the regressions of this chapter.

I used river polygon and river arc files, and lake polygon files provided by ESRI from data they have readily available for their software users. I used a river file given to me by IMAZON to guide in selecting the level of river that is useful for cargo ships. All river arcs were converted to 228 meter gridcells before I computed the area of water in each census tract. To get the total area of a census tract, I used Arc View to compute the area of each census tract polygon in a sinusoidal projection, 54 degrees west, with a spherical datum. Then I subtracted the area of water contained in the census tract, computed using Arc View’s tabulate areas feature. I used IBGE census tract files together with ESRI water files to determine which indentations on the map were for rivers, and which were coastal inlets. I also used ESRI files for mineral locations.

I chose the primary roads by looking at road shapefiles provided to me by IMAZON, after they had processed files from IBGE’s *Diagnostico Ambiental* CDROM (IBGE 1997a). I chose roads based on whether they connected major cities, and by advice given in papers by Alves and by the team at Woods’ Hole LBA project. The settlement dataset was given me by IMAZON. The *antropísmo* dataset was also given to me by IMAZON, which was processed by them from the *Diagnostico Ambietal* CDROM (IBGE 1997a). I processed the varzea locations from *Diagnostico Ambiental* myself.

The location of cities came from many sources. The larger cities came from ESRI's dataset. Some of the other cities came from NIMA, and when all else failed, from a printed map.

Model of Deforestation

The simplest model of deforestation is a static model in which gross revenue is given by (expected) price of the agricultural output, p, times a production function, $f(\,)$, with cleared land, D, as the only input. Total land available to the farmer is S. Presumably the production function gives decreasing returns to scale, because household labor is fixed, and the labor market for hiring additional laborers is thin; or because of imperfect capital markets; or because land has varying characteristics of quality—slope being the most obvious and the one of perhaps greatest importance on a given farm—and the best land is cleared first. There is a cost to clearing land, and a cost to maintaining land for agriculture. In a static model, we would not model these as separate costs, so would simply posit a cost of clearing function, $c(\,)$, that has one input, D. In a simple model, it makes sense to assume a linear function for $c(D) = cD$. The farmer then chooses D to maximize profit, equal to $pf(D) - cD$, subject to the land constraint of $S - D > 0$. An interior solution is the well known result that the farmer clears the land until the marginal revenue from an additional unit of cleared land equals the marginal cost, c, of clearing that land.

The simplest dynamic model would assume the same gross revenue function for each period, but a slightly more complicated clearing function. Assume d_t is the quantity of land cleared during period t, with total cleared land at the beginning of period t, D_t, equal to D_{t-1} plus d_{t-1}. Clearing cost is now a function of d_t. In practice, we rarely observe farmers clearing all of their land during one period of time. This is probably because the clearing function has increasing marginal costs, caused either by imperfect labor markets[19] or imperfect capital markets. It would also be possible to postulate an alternative hypothesis: that the farmer is learning about the production process in a particular location, and clears more land as he or she learns how to optimize production on his or her farm. Research by Perz and Walker support the earlier point about imperfect markets. They look at land use changes as household composition changes. When children leave a farm to start their own household, then it is common for some land that has already been cleared on

---

[19] If a farmer relies on household labor alone, this devoting labor to clearing land may pull labor away from farming activities, reducing farm income. In a case where a farmer has unlimited capital, with thin labor markets, hiring additional units of labor drives up the cost of each additional unit of labor. When capital markets are thin, borrowing additional units of money becomes increasingly costly.

that farm to be put into fallow—pointing to labor market constraints or the non-substitutability of household labor and hired labor.

We can write the discrete current value Hamiltonian for this problem as

$$H = p_t\, f(D_t) - c(d_t) + m_{t+1}\, d_t + l_t\, (S - D_t)$$

where l_t is the Lagrange multiplier on the constraint limiting land available for clearing, and m_{t+1} is the co-state variable for D_t. In continuous time, we would write the Hamiltonian as

$$H = p(t)\, f(D(t)) - c(d(t)) + m(t)\, d(t) + l(t)\, (S - D(t))$$

Suppressing the time indices, the first order conditions are

$$\partial H / \partial d = 0 = -c'(d) + m \tag{1a}$$

$$\partial H / \partial D = rm - \dot{m} = pf'(D) - l \tag{1b}$$

$$\partial H / \partial m = \dot{D} = d \tag{1c}$$

$$l\, (S - D) = 0,\ \ l \geq 0 \text{, and } S - D \geq 0 \text{ for every t.} \tag{1d}$$

Solving (1a) gives $m = c'(d)$. Taking the derivative with respect to time gives $\dot{m} = c''(d)\dot{d} = c''(\dot{D})\ddot{D}$. Substituting into (1b) and assuming that this is not a corner solution gives

$$\dot{d} = -\left[pf'(D) - rc'(d)\right] / c''(d) \tag{2}$$

At steady state, neither d nor D will be changing. Furthermore, d will equal 0. This implies that at steady state D solves $pf'(D) = rc'(0)$; or, if the land constraint is binding, $D = S$.

Figure 12 shows a phase-plane diagram which uses equations (1c) and (2). Graphing $\dot{D} = 0$ is trivial using (1c). At $\dot{d} = 0$, (2) reduces to

$$0 = pf'(D) - rc'(d).$$

Totally differentiating and rearranging gives us

$$\frac{dd}{dD} = \frac{pf''(D)}{rc''(d)}$$

By second order necessary conditions, the denominator is non-negative (Kamien and Schwartz, p. 128), and I will assume it is strictly positive for $d > 0$ due to my earlier argument about my belief that marginal costs of clearing are increasing. So for $d > 0$, the slope is negative since the production function is assumed to have decreasing marginal returns. Kamien and Schwartz (p. 133) tell us that the Kuhn-Tucker conditions are also sufficient, as long as the single-period profit function is concave, and maximized over a closed convex region.

We could adapt the model to allow for land abandonment. Farmers might choose to abandon land if incentives for farming change, including changes in prices, taxes, subsidies, and soil fertility. This change requires a piecemeal clearing cost function: one value for non-negative *d*; another for negative *d*.[20]

---

[20] It would be piecemeal simply because we would normally expect the cost of abandonment to be 0, and no function is exactly equal to 0 for negative values, increasing marginal costs with positive values; and continuous and twice differentiable at 0.

In this model in which abandonment is costless and clearing is costly, we would never choose to simultaneously do both, because we could have a less expensive outcome by simply clearing the net of the two. However, when mechanisms for maintaining soil fertility are not available or perhaps not optimal to use (Barrett, Krautkraemer), farmers might allow natural regeneration to restore soil fertility (i.e., allow land to become fallow). In such an instance, simultaneous clearing and abandonment might occur on separate plots.

Alston, Libecap, and Mueller suggest that the perverse property right system in Brazil may lead to greater deforestation and clearing than is optimal. Brazil's laws allow for expropriation of unused private land. In order for a farmer to guard against expropriation, he or she needs to make the land claimed to appear productive, often by establishing low-value pasture with widely scattered cattle. This is at cross-purposes with other Brazilian legislation which limits deforestation to 20 percent of land owned in the forest zone of the Amazon, or 50 percent of the land owned in the *cerrado* part of the Amazon. The implication of deforesting to protect property rights is a larger than optimal up-front deforestation.[21] But since the model already shows a declining clearance through time, adding the cost of defending against ex-

---

[21] And possibly a larger than optimal steady state deforestation

propriation adds little to the outcome of the model. If we think that farmers actually worry that they will one day be penalized for exceeding the legal limits of clearance on their farmland, then we could also include this in the model.

Consider the implications of this simple micro-model at the farm level in the broader context of what causes on-going deforestation in the Amazon. Clearly some of the deforestation is taking place on already established farms as farmers incrementally clear land, as the model predicts. Furthermore, if prices shift, so that they increase at a point in time (whether expected or not), farmers who have already reached the steady-state level of clearance might decide to deforest additional land. These price shifts might occur through increases in the world or local price of a commodity, or they might simply occur because of decreased transportation costs resulting from cheap fuel or new infrastructure. The focus of this thesis is on the effect of the latter.

Of course, the already established farm is only one source of on-going clearance. The other source is the establishment of new farms by settlers. Prospective settlers include children of Amazon farmers who move out of their original household to establish a household of their own. New settlers could also come from other areas of Brazil—either with or without previous farming experience—or from urban areas in the Amazon. While it is beyond the scope of this dissertation to model the process which drives new settlement, it seems clear that the prospective farmer is one who is weighing the option of establishing a farm or ranch in the Amazon against one or more alternatives. Presumably prospective farmers would try to time their settlement to when the market is favorable to farming, especially as measured relative to their current situation and other alternatives. Unemployment and inflation rates are just two factors which might influence decisions.

The prospective farmer, perhaps after deciding to establish a ranch or perhaps simultaneously with that decision, decides where to locate his ranch. Presumably the economic factors include profitability measures, including farmgate prices, rainfall, and soil quality. They might also include availability of off-farm work, as a means of minimizing risk if the ranch is unproductive, or as a strategy to maximize total household income. Non-economic factors might include nearness to hospitals, schools, or neighbors.

Since the best land from an economic perspective is already taken (in the absence of changes to the *status quo*, such as road construction), and since economic values are geographically correlated (e.g., output prices paid at neighboring farms would be very close), apart from change in accessibility brought on by new infrastructure, we would expect settlers to move in near already established farms. Travel is costly, as is building access roads. If a settler establishes a farm in a “fishbone” settlement (see figures with satellite imagery in Chapter 4), he or she is likely to move only as far down one of the bones as is necessary to find a vacant plot. In a non-fishbone settlement where a settler has to build his own access road, he would like to build as short a road as possible, and that would be from the neighbor's access road to his own new farm location.

Returning once again to the basic model in translating its implications for applied analysis, we would like to estimate a reduced form equation of total deforestation. Recognizing, first of all, the on-going nature of on-farm land clearance, and second, the desire to minimize costs in settling a new farm, 1976 *antropísmo* is an important variable to include in the analysis because it establishes a baseline that is indicative of how far along farms are in the land clearance process, and provides a base from which we believe new settlements will have spread out.

While output price should be in the regression, we do not have a price measure. What we do have, however, are factors affecting price, including accessibility (roads and rivers) and nearness to demand centers (cities with population greater than 100,000).

Anything that increases the productivity of the land will likely affect both the rate and the steady state level of deforestation. Things such as soil type and rainfall fall in this category. Furthermore, anything that increases the cost of clearance will affect the rate and level of deforestation. This would include vegetation zones.

Now that we have motivated the explanatory variables of the analysis, we are ready to move onto the results.

Results

In the regressions reported in this chapter, I use census tract-level data from the 1996 agricultural census. The dependent variable is the proportion of the census tract cleared for agriculture. There are 613 census tracts with no deforestation (i.e., censored at 0) and 276 census tracts with clearing greater than or equal to the area of the census tract.[22] Because of the censoring on each end, I used a double tobit. Furthermore, I believe that there will be some heteroscedasticity present, not only because there is measurement error in the dependent variable that is likely to be correlated with the perimeter-to-area ratio of the census tract (i.e., “skinny” census tracts are more likely to have “spillover” than rounded ones, and smaller census tracts are also more likely to have “spillover” than larger ones), but also because the variance might be effected by levels of certain variables. One example is that I would expect that there would be lower variance at higher rainfall levels, since I expect that higher rainfall levels as a rule will result in lower clearance levels.

---

[22] Clearing is greater than census tract area for some census tracts with farms that “spillover” into neighboring census tracts, but are reported as being present in one census tract.

The first regression I ran is not reported in any table, because it simply regressed proportion of census tract land used for agriculture[23] on road buffer variables and a constant, with perimeter-to-area ratio (of the census tract) and its square in the heteroscedasticity term. The results were 0.227 in the 0 to 25 kilometer buffer; 0.137 in the 25 to 50 kilometer buffer; and 0.120 in the 50 to 100 kilometer buffer; all with t-statistics of 10 or larger. Each parameter in the heteroscedasticity term had a t-statistic of 27 or larger. Large perimeter-to-area ratios have large heteroscedasticity, as expected. The road buffer parameters in this regression using the 1996 agricultural census data are similar to the cross-tabulations from the 1992 TRFIC data in Table 13: 20.5, 11.9, and 6.3. Since the datasets were 4 years apart, and since the sources picked up different types of deforestation, the similarities are remarkable.

The results of more complex weighted heteroscedastic double tobits are reported in Table 15, and the results by biome are reported in Table 16. A graph of the regressions reported in Table 15 is shown in Figure 13. In Figure 13, we see that the *cerrado* curves cut off at 2,100 millimeters of rainfall, because rainfall averages simply are not observed in the *cerrado* above that limit.

---

[23] In the previous chapter, the cross-tabulations on cleared land excluded natural pasture. In this chapter, we consider natural pasture as agricultural land, and it is part of the numerator of the dependent variable, which includes all of the other components of cleared land (crops, perennials, tree plantations, fallow, and vacant but productive—i.e., long-term fallow).

Table 15. Tobit Regressions on Proportion of Census Tract in Agricultural Land

	All variables		Roads and natural regressors	
	param	**t-stat**	**param**	**t-stat**
Observations	6,693		6,693	
Log likelihood	4,156.4		3,242.6	
Rain, annual, mm	-2.13E-03	-10.97	-2.67E-03	-10.44
Rain squared	8.10E-07	10.94	1.03E-06	10.46
Rain cubed	-1.00E-10	-10.79	-1.30E-10	-10.39
Protected area	-0.0206	-8.23		
Road buffer, <25km	0.0650	10.09	0.1210	16.09
25-50km	-0.0083	-1.01	0.0164	1.59
50-100km	0.0048	1.00	0.0214	3.68
Settlement, <25km	0.0465	7.83		
25-50km	0.0096	1.53		
50-100km	-0.0014	-0.31		
Antropísmo, inside	0.4117	24.30		
<25km	0.1417	17.66		
25-50km	0.0337	4.26		
50-100km	-0.0007	-0.16		
City buffer, <25km	-0.0959	-1.35		
25-50km	-0.0435	-1.25		
50-100km	0.0442	3.89		
Manaus, <25km	-0.5245	-1.90		
25-50km	-0.1207	-1.01		
50-100km	-0.0711	-2.47		
Low organic mtr.	0.1520	6.43	0.1114	3.92
Seas. excess water	0.0096	0.08	-0.0577	-0.45
Minor root restr.	0.0283	3.34	0.0153	1.44
Impeded drainage	-0.0339	-3.59	-0.0478	-3.91
High aluminum	0.0679	3.03	0.0925	3.02
Excess nutr. leach	0.0059	0.79	-0.0100	-0.97
Low nutr. holding	-0.0097	-2.32	-0.0426	-8.18
High P, N, & organic retention	-0.0325	-1.61	-0.0199	-0.84
Low water holding	-0.0316	-3.39	-0.0739	-6.39
Salin. or alkalinity	0.1436	5.54	0.1411	4.34
Shallow soils	-0.0026	-0.40	-0.0625	-6.72
Cerrado	0.1941	20.81	0.2074	18.78
Northern savannas	0.0311	1.71	0.0313	1.47

	All variables		Roads and natural regressors	
	param	**t-stat**	**param**	**t-stat**
Varzea, Marajo	0.0143	1.13	0.0301	1.82
Varzea, other	-0.0188	-2.99	-0.0091	-1.07
Pantanal	0.2011	6.58	0.1819	5.25
Mangroves	-0.1793	-2.43	-0.1069	-1.57
Campinarana	0.0106	2.31	-0.0005	-0.06
Dry forests	0.0231	3.65	0.0718	9.43
Babaçu forests	0.0561	2.24	0.1377	4.55
Varzea, fluvial	-0.1488	-5.88	-0.1232	-3.76
Varzea fluvial *Varzea, Marajo	0.3546	8.65	0.2637	5.02
Varzea, marinha	-0.4641	-6.31	-0.4411	-6.38
River area	0.1549	3.78	0.1821	3.56
Constant	1.8357	10.91	2.2744	10.45
$ln(s_i)$				
Perimeter / area	642.523	43.24	606.164	39.92
Squared	-8,023.2	-37.98	-7,617.7	-35.54
Within 10 km of river	-0.1471	-3.92	-0.2394	-12.32
Rain, annual, mm	-1.26E-03	-49.82	-1.07E-03	-39.50
Constant	-0.4624	-9.09	-0.5822	-11.92

In fact, very few places in the *cerrado* have values above 2,000 millimeters. In the forest biome, rainfall is observed below 1,600 millimeters (also for the *cerrado*) and up to 3,600 millimeters, but both of these extensions cover such small portions of the land that it seemed to give an incorrect perception to extend these curves further.

In Figure 13, each biome has 3 curves. All curves show predicted deforestation based on the assumption of the dominant soil and vegetation type of the biome, but far from cities and rivers. The solid curves show deforestation expected at distances far from roads and *antropísmo*. The long-dashed curves

Table 16. Regressions on Proportion of Census Tract in Agricultural Land, by Biome

	Forest		***Cerrado***	
	param	**t-stat**	**param**	**t-stat**
Observations	5,412		1,281	
Log likelihood	4,598.2		71.7	
Rain, annual, mm	-1.64E-03	-11.54	-1.53E-03	-0.11
Rain squared	6.16E-07	11.51	1.79E-06	0.21
Rain cubed	-7.57E-11	-11.40	-5.33E-10	-0.32
Protected area	-0.0111	-6.91	-0.3512	-12.38
Road buffer, <25km	0.0531	9.67	0.0463	1.91
25-50km	-0.0102	-1.44	0.0298	1.12
50-100km	-0.0033	-0.95	0.0437	1.99
Settlement, <25km	0.0460	9.67	0.0153	0.50
25-50km	0.0163	3.25	0.1194	4.11
50-100km	-0.0011	-0.36	0.0670	2.44
Antropísmo, inside	0.4434	16.73	0.4734	6.88
<25km	0.1351	18.90	0.1336	4.81
25-50km	0.0193	2.87	0.0974	3.55
50-100km	-0.0065	-2.09	0.0707	2.78
City buffer, <25km	-0.0598	-0.72	-0.0018	-0.01
25-50km	-0.0630	-1.71	0.1425	2.03
50-100km	0.0307	2.99	0.0744	2.18
Manaus, <25km	-0.5728	-1.27		
25-50km	-0.0865	-0.67		
50-100km	-0.0425	-1.76		
Low organic mtr.	0.1429	7.20	-0.2702	-1.23
Seas. excess water			0.1411	1.33
Minor root restr.	0.0146	2.02	0.1012	4.09
Impeded drainage	-0.0244	-3.55		
High aluminum	0.0715	4.20	-0.2818	-4.28
Excess nutr. leach	0.0046	0.91	-0.2137	-2.58
Low nutr. holding	-0.0076	-2.53	-0.4987	-4.57
High P, N, & organic retention	0.2625	5.50	-0.0338	-1.03
Low water holding	-0.0232	-3.15	0.0351	1.60
Salin. or alkalinity	0.1298	6.35		
Shallow soils	-0.0065	-1.41	0.0593	1.06
Cerrado	0.1460	6.11	-0.0175	-0.06
Northern savannas	0.0054	0.38		
Varzea, Marajo	0.0356	3.34		

	Forest		***Cerrado***	
	param	**t-stat**	**param**	**t-stat**
Varzea, other	-0.0128	-2.86		
Pantanal			-0.0124	-0.04
Mangroves	-0.2564	-5.70		
Campinarana	0.0028	0.98		
Dry forests	0.0252	5.70	-0.0728	-0.26
Babaçu forests	0.0844	4.25	-0.3292	-0.90
Varzea, fluvial	-0.1147	-6.03	0.4052	0.98
Varzea fluvial *Varzea, Marajo	0.2792	8.62		
Varzea, marinha	-0.3958	-9.04		
River area	0.1673	5.41	-1.0323	-0.73
Constant	1.4272	11.52	0.2788	0.03
$ln(s_i)$				
Perimeter / area	958.127	48.28	192.342	7.70
Squared	-12,209.8	-44.53	-2,430.8	-3.88
Within 10 km of river	-0.2015	-4.89	0.0371	0.44
Rain, annual, mm	-1.34E-03	-43.10	-1.86E-04	-0.83
Constant	-0.9967	-14.81	-1.5990	-4.09

show predicted deforestation based on being within 25 kilometers of a principal road. The short-dashed curves show predicted deforestation based on being within 25 kilometers of a principal road and 25 kilometers of *antropísmo.* We might see this as a worst-case scenario in the very long run.

The curves show us a number of important points about the deforestation. First, we see that all other things being equal, deforestation (more precisely, converting land for agricultural uses) is higher by far in the *cerrado* biome than in the forest biome. Second, we see that increased rainfall leads to a tapering off of deforestation, then a flat response after around 2,300 millimeters of rain. Third, we see that road construction by itself leads to only a modest increase in probability of deforestation, but in conjunction with "starter deforestation"—a locus of deforestation from which future deforestation may spread—there is a dramatic increase in probability of deforestation. We noted this interaction earlier in the discussion of Table 11, but it is an important point that cannot be over-emphasized.

If policies can be put in place to limit the establishment of "starter deforestation" along principal roads, then the unintended consequence of deforestation resulting from building a road through undisturbed forest can be limited dramatically. This may involve policies that set aside most of the land along the road as protected area. When we examine the results in Tables 15 and 16, we see that protected areas seem to be effective in limiting agricultural conversion in the *cerrado*. In the forest biome and in the combined analysis, the parameter estimate for protected areas is quantitatively small. However, if most protected areas are in areas with low predicted deforestation even without a protected area, the small parameter might still represent a significant percentage reduction in deforestation.[24]

---

[24] Indigenous areas were chosen principally because of the location of the indigenous people. The location of national forests and other non-indigenous-area protected areas, however, may be endogenous (in a statistical sense of the word). If that is so, the parameter estimate likely overstates the true protection offered by protected areas.

A more radical second policy to be used in conjunction with the first might be to concede that there will be some development along the new road, and actually encourage the development to be concentrated in a specific location. This second policy option might backfire—it might create more deforestation by promoting the development in a specified location than it prevents by letting development take place where it would occur naturally, even given the first policy. On the other hand, it might make establishing protected areas along the rest of the road more politically feasible, and it might limit the number of "starter deforestation" patches, resulting in less deforestation. The model I ran suggests that one concentrated patch results in less deforestation, but the model is not sensitive to the nuánces of this policy suggestion.

Figure 13 suggests that deforestation would more readily approach the worst-case scenario in the *cerrado* biome than in the forest biome. That is because the probability of deforestation is so much higher in the *cerrado* than in the forest, that more "starter deforestation" patches are likely to occur there than in the forest.

The curves in Figure 13 lead us naturally into warning the reader about dangers in interpreting parameters from heteroscedastic tobits, compared to interpreting parameters from OLS regressions. Note that the solid forest curve never rises, even though the parameters tell us that it should rise for values greater than 2,300 millimeters of rain. The failure to rise is due to two effects. First, the variance of the error term decreases with increasing rainfall, and a small variance has a dampening effect relative to a large variance. Second, as other parameters cause the uncensored predictions to be below 0, any positive effects of rainfall (which the parameters tell us should increase after 2,300 millimeters) are muted by the censoring. Compare the solid forest curve to the short-dashed forest curve: the rainfall rise is clear in the latter. Also, note the "muting" effect of censoring for rainfall below 1,800 millimeters, by comparing the difference between the solid forest and long-dashed forest, to the difference between the solid *cerrado* and the long-dashed *cerrado*: they are both shifted by the same parameter, but the shift is much smaller in the forest curves.

To see why, look at the following formula for the expected value of a tobit.

$$EY_i = \Phi_a a + (\Phi_b - \Phi_a)(m_i + s_i l) + (1 - \Phi_b)b$$

where $m_i = X_i b$; $s_i = exp(Z_i a)$; $l \equiv \frac{f_a - f_b}{\Phi_b - \Phi_a}$; $\Phi_a = \Phi((a - m_i)/s_i)$ and F_b defined similarly; $f_a = f((a - m_i)/s_i)$ and f_b defined similarly; a is the lower censoring value; b is the upper censoring value; and F and f are the normal cdf and pdf. If m_i is negative, then F_a is greater than 0.5 but less than F_b. In such a case, the difference between F_a and F_b is less than 0.5, and l is positive, since its denominator is always positive, and since the pdf is decreasing after 0, the numerator is also positive. As m_i becomes more negative, Φ_a and Φ_b converge, producing the “muting” effect of censoring.

By taking the derivative of the expected value of y with respect to x, we see that the marginal effect of a change in an independent variable x on the conditional mean is $m'(\Phi_b - \Phi_a) + s'(f_a - f_b)$, where m' and s' are shorthand for partial derivatives with respect to x. If m_i is either a large negative number or a large positive number, F_a and F_b will be close to each other, as will f_a and f_b. In such a case, m' will be multiplied by a number close to zero (as will s'), meaning that any change in the level of x will have a small impact on y. What is also important is the fact that the marginal is affected by the variance as well as the mean (assuming the variable of interest influences the variance).

While we have looked at some of the most important variables in our discussion of Figure 13, we also need to shift our focus to some of the other variables reported in Table 15. The first set of results in the table is the main regression for this chapter. The second set of results show road variables and “natural” variables. Variables that reflect the intervention of people—apart from roads—were dropped from the first set of results, to show the effect of roads, controlling for agricultural and locational suitability. Earlier I reported that not controlling for agricultural and locational suitability, the parameters on the road buffers were 0.227 in the 0 to 25 kilometer buffer; 0.137 in the 25 to 50 kilometer buffer; and 0.120 in the 50 to 100 kilometer buffer. Now we see that they are 0.1459, 0.0483, and 0.0249. The difference is quite important for the debate about the impact of roads on deforestation, because the previously published results in *Science* (e.g., Laurance *et al* 2001) and elsewhere used cross-tabulations (equivalent to the regression results which did not control for agricultural suitability) to argue for the large adverse effects from building roads, and the second set of results in Table 15 show that the effects are quantitatively smaller: by one-third out to 25 kilometers; almost two-thirds in the 25- to 50-kilometer buffer; and almost four-fifths in the 50- to 100-kilometer buffer. These results might be looked upon as the combination of direct- and indirect-effects of road construction, controlling for agricultural suitability.

However, as I argued above when discussing the meaning of the results in Figure 13, the direct effect of roads is much smaller. Table 15 reports that they are 0.0623, 0.0091, and 0.0018, respectively—the latter two not statistically different than 0. The regression shows *antropísmo* (human disturbance) by 1976 is the variable that has high quantitative and statistical significance. However, when we regress 1976 *antropísmo* on proportion in various road buffer categories, we find that nearness to roads explains *antropísmo.* It is through the *antropísmo* variable that roads appear to exercise an indirect effect.

We also see that government settlements might possibly offer “starter deforestation”, but I am hesitant to conclude this, because the data only show a response out to 25 kilometers, and a large part of this response might be the settlement itself rather than deforestation that grew up around the settlement. This is not a serious blow to the hypothesis of disturbance being a “seed” for further deforestation, because we did not have the precise age of any of these settlements, and if they had been formed not long before 1996, then the “seed” effect would not have had a chance to spread beyond a 25-kilometer radius.

Testing the “starter deforestation” hypothesis even further, I tried seeing if mines and minerals led to deforestation. But there was no statisticially significant response at any distance, and so I dropped the variables from the regression. Because of uncertainty about what the minerals dataset actually represents, I am cautious to draw too strong an inference from this result as a blow against my “starter deforestation” hypothesis.

Buffers around centers of cities with populations greater than 100,000 in 1996 show negative parameters out to 50 kilometers. I believe this reflects the absence of farms in urban and suburban areas, not the absence of deforestation. Indeed, Table 13 shows that deforestation is high in these areas, not low. The positive response in the 50- to 100-kilometer buffer is small but significant, and probably indicates some of the demand effect of cities on agriculture. The actual effect could be even larger if suburban areas extend out into this buffer, as they would in a large metropolitan area like Belém, with a population well over one million. For cities with populations closer to 100,000 people, it is not clear that the non-farm influence would extend that far.

It has been suggested that development in and around Manaus is different from anywhere else in the Legal Amazon, because of the free trade zone declared for it in 1967. The regression results confirm that something is different about Manaus, with less agricultural activity in all buffers around it than would be expected from a city of its size. It is possible that the free trade has distorted prices enough that it is not optimal to have agricultural activities near the city to produce the type of von Thünen product that is costly to transport greater distances. An alternative hypothesis is that the demand by manufacturers for space in the free trade zone displaced farmers to even greater distances from the city. The latter hypothesis seems weak, since much of the 50- to 100-kilometer buffer around Manaus is outside of the 10,000-square kilometer trade zone.

In regression results not included here, I also explored the influence of cities with populations between 25,000 and 100,000. These tended to be positive in all buffers, with relatively large values and statistically significant. I decided in the end that this size city was probably endogenous (in the sense that a number of these cities grew to their size because of principle roads passing though them), or at least another avenue through which the indirect effects of major roads is expressed.

Protected areas provide a quantitatively modest but strongly statistically significant negative effect on agricultural clearing. This does not mean that it is not important for preventing settlement and deforestation. The weighted mean of proportion of land used for agricultural purposes, as calculated from the agricultural census, is 13.0 percent. However, it is only 8.2 percent for the forest biome, while it is 38.8 percent for the *cerrado* biome. But 77.6 percent of the land in our study area is in the forest biome[25]. Furthermore, for the forest biome, 26.2 percent of it is in a protected area, while that is true for only 10.8 percent of the *cerrado.* That means that the parameter on protected areas mostly reflects its value in limiting agricultural conversion in the forest zone. Still, a negative shift of 2.1 percent is not as large as expected when mean deforestation is 8.2 percent. As argued previously, if many of the protected areas are in remote places where the agricultural potential may also be quite low, it is likely that a parameter of –0.0201 reduces expected deforestation significantly (to a small percentage of what it would have been without the protected area).

---

[25] using the definition of land being in the *cerrado* biome if 50 percent of its area is in the Pantanal or southern *cerrado*, using the WWF definitions

As we look at the effects of WWF vegetation zones, we see that *cerrado* and Pantanal have almost equally high shifts, indicating that they are similar to each other in terms of land clearing, and that there is a large difference between agriculture in those areas relative to agriculture in moist forests. When comparing the *cerrado* parameter to the parameter on the northern savannas variable, we also note a significant difference. Unless the qualities of the northern savannas are significantly different than the southern savannas (*cerrado*)—and I have not found anything in the literature to suggest such a difference—then the difference must be due to a mixture of economic and non-economic advantages of being nearer to the major population centers of Brazil. For ranchers, this could include price effects from restrictions on interstate transport of beef for most of the Legal Amazon states, due to foot and mouth disease. The restrictions have not been as strict for Mato Grosso, most of which is in the *cerrado* zone.

Varzeas (floodplains) are considered important areas for ecology and agriculture. However, the WWF designation of *varzeas* differed substantially from that of IBGE. As an economist, it was difficult to determine which designation to use, so I tried to use both, and included an interaction term on Marajó island, to distinguish the two parts of Marajó. The part of Marajó that was included in both types of varzea showed an extremely large and statistically significant value, while the part of Marajó not counted as *varzea* by IBGE was not statistically different from zero. Both fluvial and marine varzeas were large and negative determinants of agricultural settlements. This makes sense for agricultural census figures, since these areas are often seasonally flooded, and may be undesirable for permanent settlement, though they may be good as common grazing areas in the drier parts of the year.

The primary limiting factors for soils are more difficult to interpret, probably due to an ignorance on my part of the technical obstacles each limiting factor presents. Presumably some limitations are relatively easy to overcome with inexpensive interventions, while others are more costly to deal with. Soil limitations can put constraints on agricultural activity. As a simple example, very steep slopes seem most suitable for trees; steep slopes can be suitable for livestock; land requiring large farm equipment needs to be reasonably flat. Steep land would eliminate the possibility of growing soybeans, which are most economically cultivated with large tractors.

Some of the soil limiting factors with the largest absolute values on the associated parameters are areas which make up a very small portion of the study area. Soils with salinity or alkalinity problems make up less than one percent. Soils with low organic matter make up 1.5 percent, and soils with high aluminum make up 1.3 percent. The main soils are those with seasonal moisture stress (the omitted category), which makes up almost a quarter of the study area, and almost 50 percent of the *cerrado* zone; and those with low nutrient holding capacity, which makes up almost 45 percent of the study area. While the parameters on soils are jointly significant, they seem to be useful in explaining the differences in only a small part of the Legal Amazon.

Having a river in a census tract resulted in a positive and statistically significant effect on agricultural clearing. This might indicate the role of rivers in accessibility and transporting people and products. In some regressions not reported in the tables, I included buffers around rivers. They were either not statistically significant, or only mildly significant, and so were excluded. I even tried separating major rivers from less important rivers, but this had no statistically significant effect. I also tested buffers from the coast. They were not statistically significant.

Table 16 records the results from regressions run separately on forest and *cerrado* biomes. We note the high significance of running them separately, by comparing the sum of the values of the log likelihood for the two regressions reported in Table 16, and comparing it to the log likelihood in the first regressions reported in Table 15. The difference is 513.5, and the likelihood ratio test statistic is twice this number. This statistic is distributed as a chi-square with 55 degrees of freedom. Needless to say, it is statistically significant at all meaningful standards of comparison. Note that a number of parameters were dropped from the *cerrado* regression. I normally did this because the variables were not represented in the dataset (e.g., no part of the biome was less than 100 kilometers from Manaus), but in some cases it was dropped due to being statistically insignificant.

The effect of rainfall on deforestation in the forest biome behaves almost identically to what is reported in Table 15 and Figure 13 for the whole sample: deforestation declines until 2,300 millimeters, then rises to 3,200 millimeters, then declines again, with most of the change occurring between 1,400 millimeters and 2,000 millimeters. The effect of rainfall on deforestation in the *cerrado* behaves differently: it rises from 1,400 millimeters to 1,630 millimeters, then declines afterward.

The magnitude on the parameters for protected areas differ markedly between the two zones: –0.011 to –0.352. As argued earlier, this probably represents the values of the parameters necessary to show protected areas as being quite effective in limiting agricultural encroachment in the respective biomes, given the difference in mean agricultural clearance between the two biomes, and taking into consideration the location of the protected areas in each biome.

The next set of parameters we come to in Table 16 are those for road buffers. In the forest biome, road buffers seem to have an effect out to 25 kilometers, and then nothing afterward. Whereas in the *cerrado* biome, we see an effect all the way to 100 kilometers (though the parameter for the 25- to 50-kilometer buffer is not statistically different from zero). The magnitude of the road effect out to 25 kilometers is larger in the forest biome. Perhaps this indicates the importance of roads in the forest biome for agricultural clearance, but the difficulty for farmers and ranchers to cut their own access through the forest is too great, and so the depth of cutting into the forest is much more limited in the forest biome.

Antropísmo continues to be an important factor in deforestation, with parameters similar across biomes. The impact of government settlement areas are different, however. In the forest biome, government settlement areas seem to only have the strongest effect out to 25 kilometers, and then a smaller but statistically significant effect from there out to 50 kilometers; while in the *cerrado* biome, they seem not to have an effect out to 25 kilometers, but a large and significant effect from 25 to 50 kilometers, and a moderate and significant effect from 50 to 100 kilometers. It is not clear why this would be. It may simply be an artifact of not having dates for each settlement scheme.

The last parameters I want to comment on are the city buffers. The really large cities—Belém, Manaus, and São Luís—each with a population larger than three-quarters of a million, are all located in the forest biome. The largest city in or near the *cerrado* is Cuiabá with around 400 thousand. It seems reasonable that under the effect of these large cities, agricultural settlements would be smaller than expected out to 50 kilometers in the forest biome; while that effect would only go out to 25 kilometers in the *cerrado* biome.

In an effort to control for possible spatial autocorrelation in the tobit, I tried a number of possible solutions. The first 3 used the same principle: grouping observations by geographical units (here, I tried 9 states, 30 macroregions, and 103 microregions), and running the tobit with dummy variables for each geographical unit (except one, to avoid multicollinearity with the constant term). The fourth attempt at compensating for spatial autocorrelation used the method that Pfaff used in his paper: including a weighted product of each of the independent variables of the neighbors. That is, I made a row-normalized weights matrix, W, of all neighbors of each census tract, and multiplied it by the matrix of independent variables, X. I then ran a tobit regression with X and WX.

I tested whether each solution sufficiently corrected for spatial autocorrelation using the Kelejian-Prucha generalized Moran I test. In each case, I had to reject the null hypothesis of no spatial autocorrelation. This means that while Pfaff's method—which was incorporated into the analysis by Andersen et al. as well—provides some interesting insight about how *municípios* influence deforestation in neighboring *municípios*, it ultimately does not correct for one of the main problems he was trying to address: spatial autocorrelation. While my experiment of trying to find ways to correct for spatial autocorrelation was disappointing in not accomplishing the main objective, it had the positive benefit of demonstrating that the critical parameters of interest were robust to all alternative specifications.

In the next chapter, we will repeat the main regression from Table 15, but use a new methodology which seeks to be able to disaggregate results geographically within the census tract, and improve efficiency by retaining the information contained in the more detailed distribution of each of the explanatory variables.

Chapter 3: Using GIS to Disaggregate Census Tract Data Spatially

Introduction

The second half of Chapter 2 presented the results of several tobit regressions, investigating how the proportion of each census tract cleared for agriculture is explained by infrastructure, agroclimatic suitability, and other variables. This chapter reproduces the main tobit regression from the previous chapter using a technique which will be outlined in the following section. This new technique subdivides each census tract using a gridded pattern, and predicts deforestation in each grid, based on the level of clearing in the census tract and explanatory variables in each grid. The second half of the chapter presents the results of this new regression, and compares them to those presented in the preceding chapter.

Wasting Valuable Information

It is common in studies to have information on different scales. For example, some variables might be known at a county level, while others are known only at the state level. Typically information is adjusted to fit the dependent variable. If an explanatory variable is coarser than the dependent variable—such as when the dependent variable is at the village-level and the explanatory variable is at the state-level—then most researchers feel comfortable saying that the coarser (state) variable is a crude indicator of the true (but unmeasured or unavailable) finer (village) variable. If an explanatory variable is finer than the dependent variable, then researchers typically sum or average (depending upon whether the variable is better aggregated as a sum or a mean) the finer explanatory variable. An example might be a case in which the dependent variable is agricultural output by state, with one of the explanatory variables being quantity of land cultivated, which might be known for each county in each state. Here, the researcher would normally sum all of the cultivated land in the state, and not be concerned about what the data perhaps could have told us about output in each county.

Geographical Information Systems (GIS) provide information about variables of economic interest across continuous surfaces. As we have already seen in the preceding chapters, types of information include location of and distances to roads, rivers, cities, etc.; climate variables such as precipitation and temperature; land survey data such as soils and native vegetation; and satellite based information such as current vegetation and fire location. Since a GIS dataset is continuous across surfaces, it tends to be fine relative to many typical dependent variables. Until now, researchers have not exploited the fineness of the data. In the last chapter, for example, since the dependent variable was proportion of land cleared in the census tract, I used explanatory variables such as rainfall at the centroid of the census tract and proportion of census tract in each type of vegetation zone. These variables disregarded the information that I had about the distribution of rainfall levels throughout the census tract, the locations of each vegetation zone inside the census tract, etc. Disregarding this information implies some loss in efficiency.

How can we keep from disregarding explanatory data that we have available to us that happen to be at a finer level than the dependent variable? The simple answer is that if we have a model that tells how the process under study combines the explanatory variables to generate the dependent variable, we can form expectations at the finer (disaggregated) level, and then sum (or average, as appropriate) these expectations to get the expectation of the dependent variable.

For the example of agricultural output at the state level and area cultivated at the county level, we need to add a model of output. Let us say we believe output is a function of land cultivated, labor, and capital. If our labor and capital data are only available at the state level, then we do not seem to gain any efficiency by first forming expectations at the county level. But if we have measures for capital and labor at the county level, we could potentially improve the efficiency of our analysis by forming expectations at the county level, and summing to get the state total. An important additional benefit is that we would also have predicted output for each county—something we could not have done (or at least not well) if we had simply summed our data for land, labor, and capital up to the state level before doing the analysis.

Since GIS data are continuous, at first pass one might think that we could disaggregate the data to infinitesimally fine points. However, while the boundary between two vegetation types appears infinitesimally small, the maps on which the GIS data are based have some type of scale—and by implication, are likely to have measurement error near the boundary between two types or regions. For example, if the map is drawn on a 1:1,000,000 scale, a mistake on the map of one millimeter represents a difference on the ground of 1 kilometer.

Map scale is just one reason for locational errors with GIS data. Even satellite data are subject to georeferencing errors. Before the ETM+ was put into service in 1999, satellite data could have errors of 2 or 3 kilometers. With ETM+, error is likely to be less than a few hundred meters. Another reason for locational error is poor data handling by GIS practitioners. It is not unusual for a GIS dataset to pass through many hands before it arrives for the researcher's use. Because practitioners do not always document the work they do on a dataset, it is not unusual to find a 5 kilometer error. Such an error might result from converting the dataset from one projection to another when there was confusion over whether the dataset was in a spherical datum or other datum, such as WGS84. Finally, some layers of data—such as precipitation—are “guestimates” based on models. With all the inherent error and uncertainty in GIS data, it does not pay to disaggregate too finely. For the deforestation study—using the data of the previous chapter—I chose to disaggregate to 5 kilometer squares, though it was tempting to go with the more conservative use of 10 kilometer squares. Before I present the study results, I will first give a brief overview of where this idea falls into the literature, then in the following section describe more explicitly the method of disaggregation that I have sketched above.

One of the first major works dealing with linear aggregation was written by Theil in 1955. His work was motivated by the earlier discussion presented in three articles in a single issue of Econometrica by Klein, May, and Pu. Later work by Lütjohann extended Theil’s work to include the type of analysis suggested here, which he called “semiaggregation”. A recent review of empirical approaches to aggregation over individuals was written by Stoker. Van Garderen, Lee, and Pesaran extend and apply aggregation to non-linear models. While my approach does deal with aggregation, it is an aggregation of artificial units of observation (gridcells) rather than true units of observation (farms).

When applied to geographic analysis, this idea is related to small area estimation[26] and to the modifiable areal unit problem—MAUP[27] (Openshaw 1984a; Paelinck; Holt et al.; Wong), which leads into the issue of ecological fallacies[28] (Openshaw 1984b; Steel and Holt). This idea is also close in spirit to the work being done with poverty mapping, except that it does not require the linking of a survey and census (Deichmann; Demombynes et al.; Henninger and Snel; Davis; Ravallion and Wodon; Jalan and Ravallion).

Disaggregation Methodology

We will begin with the simplest kind of aggregation, in which we simply sum an expectation that is modeled linearly. We assume that there is a process at the unit-level that generates each yij in group i, but we only observe the

---

[26] A good overview is contained in Bigman and Deichmann. Small area estimation techniques are ones which use related geographic areas and additional data sets to strengthen the geographic inference made by surveys which otherwise had too few observations from which to make predictions at such a fine resolution (Bigman and Deichmann).

[27] The modifiable areal unit problem arises when one aggregates individual units of observations. The units of observations can be aggregated in a number of different ways, because there are few logical constraints to place on the method of aggregation. Furthermore, different aggregations may lead researchers to different conclusions from the data, and this is the main point of the MAUP (Openshaw 1984b).

[28] The ecological fallacy "occurs when analyses based on area level... [statistics] give conclusions very different from those that would be obtained from an analysis of unit level data, if it were available" (Steel and Holt, p. 39).

sum of these values, y_i. Nevertheless, we observe the explanatory variables at the unit level. Let $y_{ij} = \boldsymbol{b_0} + X_{ij}\,\boldsymbol{b} + \boldsymbol{e}_{ij}$, where $\boldsymbol{e}_{ij} \sim \mathrm{N}(0,\ s_e^2)$. Then

$$y_i = \sum_{j=1}^{J_i} y_{ij} = \sum_{j=1}^{J_i}\left(\boldsymbol{b}_0 + X_{ij}\boldsymbol{b} + \boldsymbol{e}_{ij}\right) = J_i\boldsymbol{b}_0 + X_i\boldsymbol{b} + \boldsymbol{e}_i \tag{1}$$

where

$$X_i \equiv \sum_{j=1}^{J_i} X_{ij} \text{ and } \boldsymbol{e}_i \equiv \sum_{j=1}^{J_i} \boldsymbol{e}_{ij}\ .$$

Note that the constant term from the unit-level model sums to J_i at the group-level, which is the number of disaggregated units in group *i*. Care must be taken when estimating (1), because it is easy to run a regression accidentally with a constant rather than, or in addition to, the J_i. If the $\boldsymbol{e}_{ij}$ are uncorrelated, then $\boldsymbol{e}_i \sim \mathrm{N}(0, J_i\ s_e^2)$, which shows that unless $J_i = J_k$ for every *i* and *k*, the regression of y_i on J_i and X_i is heteroscedastic, and an appropriate adjustment needs to be made in the estimation procedure to avoid a bias to the standard error estimates of the parameters.

If the dependent variable is a ratio (as is the case in this chapter), then each ratio at the unit level has to be multiplied by a z_{ij} equal to the denominator of the ratio divided by the sum of the denominators for all units in a group. To make this more concrete, in the case of proportion of land cleared for agriculture, the multiplier for any unit is equal to the number of hectares in the unit divided by the number of hectares in the group. The regression equation for this case is

$$y_i = \sum_{j=1}^{J_i} z_{ij} y_{ij} = \sum_{j=1}^{J_i} z_{ij}\left(X_{ij}\boldsymbol{b} + \boldsymbol{e}_{ij}\right) = X_i\boldsymbol{b} + \boldsymbol{e}_i \tag{2}$$

where

$$X_i \equiv \sum_{j=1}^{J_i} z_{ij} X_{ij} \text{ and } \boldsymbol{e}_i \equiv \sum_{j=1}^{J_i} z_{ij}\boldsymbol{e}_{ij}\text{ , with } \boldsymbol{e}_i \sim \mathrm{N}\Big(0,\ s_e^2 \sum_{j=1}^{J_i} z_{ij}^2\Big).$$

Here, the constant term was included as part of the matrix of independent variables. Since the z_{ij} sum to 1, the constant term does not change when aggregated.

The log likelihood function for this is

$$\begin{aligned} \ln L = & -\tfrac{1}{2}\ln(2p) - \tfrac{1}{2}\ln\left(s_e^2\right) - \tfrac{1}{2}\sum_{i=1}^{I} \ln\left(\sum_{j=1}^{J_i} z_{ij}^2\right) \\ & - \frac{1}{2s_e^2}\sum_{i=1}^{I}\left[\left(y_i - \sum_{j=1}^{J_i} z_{ij}X_{ij}\boldsymbol{b}\right)^2 \Big/ \sum_{j=1}^{J_i} z_{ij}^2\right] \end{aligned} \tag{3}$$

This is also the log likelihood function for the case where the aggregation is a simple sum, if we let $z_{ij} = 1$ for all i and j. It can also be seen as the log likelihood function for any other kind of weighted sum or average. The advantage of the linearity is seen in that the $z_{ij}\ X_{ij}$ and z_{ij}^2 can be multiplied and summed before entering the optimization algorithm.

Suppose we make the description of the problem more general yet, by allowing the variance to be heteroscedastic, and both the expected value and the variance to be nonlinear in parameters. That is, let $y_{ij} = m(X_{ij},\boldsymbol{b}) + \boldsymbol{e}_{ij}$, where e_{ij} has mean 0 and variance $s_{ij}^2 = s^2(R_{ij},a)$. Then we may write the log likelihood function as

$$\begin{aligned} ln\, L = -{}^{I}\!/_{2}\, ln(2p) - {}^{1}\!/_{2} \sum_{i=1}^{I} ln\left(\sum_{j=1}^{J_i} z_{ij}^{2} s^{2}(R_{ij}, a) \right) \\ - {}^{1}\!/_{2} \sum_{i=1}^{I} \left[\left(y_i - \sum_{j=1}^{J_i} z_{ij} m(X_{ij}, b) \right)^{2} \Big/ \sum_{j=1}^{J_i} z_{ij}^{2} s^{2}(R_{ij}, a) \right] \end{aligned} \qquad (4)$$

Relative to (3), (4) is computationally more difficult, not only because of the nonlinearities, but also because in general we have to compute each unit-level expectation and unit-level variance every time a new parameter vector is proposed in the optimization algorithm.

In the preceding chapter, we estimated a doubly censored tobit. Is there a way to treat doubly censored data using this “adding up” methodology? Yes, there is. Feasible Weighted Nonlinear Least Squares (FWNLS) is an alternative to the tobit for estimating censored regressions. In order to see this, let us begin by assuming that there is an unobserved variable, y_{ij}^*, such that $y_{ij}^* = X_{ij} b + u_{ij}$, where $u_{ij} \sim N(0,(h(R_{ij} a))^2)$; that is, I allow the error term to be heteroscedastic. Recall that for the case of a doubly censored regression, instead of y_{ij}^* we observe a variable y_{ij} which is equal to y_{ij}^* for y_{ij}^* between a and b; equal to a for $y_{ij}^* < a$; and equal to b for $y_{ij}^* > b$. In the FWNLS for censored data, we regress y_{ij}—not y_{ij}^*—on the expected value of y_{ij} given X_{ij} and the censoring points. In order to use (4), we need to note that the $m(X_{ij},b)$ in (4) is the expected value of y_{ij}, given by

$$m(X_{ij}, b) = Ey_{ij} = \Phi_a a + (\Phi_b - \Phi_a) X_{ij} b + (f_a - f_b) h(R_{ij} a) + (1 - \Phi_b) b \qquad (5)$$

where Φ_k is the standard normal cdf and f_k is the standard normal pdf, both evaluated at $y_k = (k - X_{ij}b) \,/\, h(R_{ij}a)$, with k representing either a or b. We also note that $s^2(R_{ij},a)$ in (4) is given by

$$\begin{aligned}\Phi_a a^2 + (\Phi_b - \Phi_a)\left[(h(R_{ij}a))^2 + (X_{ij}b)^2\right] + 2(f_a - f_b)(X_{ij}b)h(R_{ij}a) \\ + (h(R_{ij}a))^2 (y_a f_a - y_b f_b) + (1 - \Phi_b) b^2 - (Ey_{ij})^2. \qquad (6)\end{aligned}$$

Note that (5) and (6) are meant to be the m and s^2 of (4), even though in (4) m is only a function of X_{ij} and b, and s^2 is only a function of R_{ij} and a.

To use the above formulas for single censoring, if we have a case of left censoring, b is implicitly equal to infinity; and in a case of right censoring, a is implicitly minus infinity. Any cdf evaluated at minus infinity is 0, and at plus infinity is 1. Any pdf is 0 at plus or minus infinity. Using L'Hospital's Rule, it is also possible to show that $y_k f_k$ is 0 at both plus and minus infinity.

For the specific data that I use in this analysis, I made two further adjustments to the likelihood function. First, I assumed that y_i was measured with measurement error, but that the error was uncorrelated with the residual.[29] I

---

[29] I made the assumption of measurement error because some farms were contained in more than one census tract, while all of the area and activity of that farm was tabulated as belonging to one census tract. Smaller census tracts and narrow census tracts are more likely to be affected by this error—

therefore was able to add the measurement error to the variance that was already modeled. Second, I weighted larger census tracts with higher weights than smaller census tracts, by dividing the variance by the weight equal to the area of the census tract, times a normalization factor given by the number of census tracts divided by the total area.

To compensate for a degree of correlation between the disaggregated units within the same census tract, but not with units from other census tracts—which I did not choose to do here—we would allow both m_{ij} and s_{ij}^2 to have variables and parameters measured at the aggregated level.

The likelihood function I maximize is

$$ln\,L = -\tfrac{1}{2} ln(2p) - \tfrac{1}{2}\sum_{i=1}^{I} ln\left\{\left[\left(\sum_{j=1}^{J_i} z_{ij}^2 s^2\left(R_{ij}, a\right)\right) + w^2\left(S_i g\right)\right]\Big/ w_i\right\}$$
$$-\tfrac{1}{2}\sum_{i=1}^{I}\left[\left(y_i - \sum_{j=1}^{J_i} z_{ij} m\left(X_{ij}, b\right)\right)^2 \Big/ \left\{\left[\left(\sum_{j=1}^{J_i} z_{ij}^2 s^2\left(R_{ij}, a\right)\right) + w^2\left(S_i g\right)\right]\Big/ w_i\right\}\right] \quad (7)$$

where w_i is the weight described previously and w^2 is the variance of the measurement error, which is a function of S_i, the vector of explanatory variables in unit i which effect the measurement error of the dependent variable and vector of parameters; and g, the vector of parameters effecting the meas-

---

both of which have higher perimeter-to-area ratios—the variable I use to control for this type of error.

urement error of the dependent variable. Here, S_i is the perimeter-to-area ratio and its square.

Data and Results

I made a GIS grid of 5 kilometer square cells. I intersected the grid with the census tract shapefile. There were 303,617 units of intersection, which I considered to be the “individual” units for the analysis. In the regression, I used the same constraint as in the previous chapter, which reduced the total number of units for the regression to 301,763. Some census tracts intersected with only one cell, while the largest census tract intersected with 2,799 gridcells. The median had 21 intersections.

The underlying data are the same as used in the previous chapter, so the reader should refer back for a complete description. Except for calculating the area of the individual unit, the data were collected by determining what the value of each variable was at the center of each gridcell, and treating those variables as the value for the entire gridcell. The value of the gridcell was then assigned to all related units of observation.[30]

The results are reported in Table 17. It took more than 70 hours for the regression to converge—a clear disadvantage to this method on highly nonlinear specifications like the FWNLS approach to censored regressions used here. The regression was run using Stata version 6.0.

Table 17. Regressions on Proportion of Census Tract in Agricultural Land

	param	**t-stat**
Disaggregate observations	301,763	
Aggregate observations	6,693	
Log likelihood	-39.0	
Rain, annual, mm	-0.0110	-13.56
Rain squared	1.82e-06	10.31
Protected area	-2.8869	-43.27
Road buffer, <25km	1.1477	16.30
25-50km	0.5634	7.58
50-100km	0.4737	7.31
Settlement, <25km	1.0503	13.82
25-50km	1.1867	15.73
50-100km	0.4155	5.81
Antropismo, inside	2.6674	24.08
<25km	1.5176	21.45
25-50km	1.1130	15.00
50-100km	0.8874	12.77
City buffer, <25km	-1.1821	-6.85
25-50km	-0.4390	-3.59
50-100km	-0.1983	-2.46
Manaus, <25km	-4.4490	-6.38
25-50km	-2.1716	-5.63
50-100km	0.7874	2.83
Low organic mtr.	1.8132	13.93
Seas. excess water	-2.0316	-5.95
Minor root restr.	-0.5658	-6.84
Impeded drainage	-1.5242	-12.89
High aluminum	0.6410	4.00
Excess nutr. leach	-0.3887	-3.22
Low nutr. holding	-1.0940	-16.48
High P, N, & organic retention	-0.3575	-2.54
Low water holding	-0.2393	-2.55
Salin. or alkalinity	0.1674	0.87
Shallow soils	-1.4906	-12.17
Cerrado	0.8342	9.94

---

[30] Some gridcells will be associated with more than one census tract.

	param	**t-stat**
Northern savannas	2.5423	16.68
Varzea, Marajo	-0.0168	-0.11
Varzea, other	-1.3078	-12.06
Pantanal	-0.7843	-4.05
Mangroves	6.1985	32.85
Campinarana	-2.6383	-4.54
Dry forests	0.6105	6.97
Babaçu forests	-0.8819	-5.24
Varzea, fluvial	-3.7184	-21.47
Varzea fluvial *Varzea, Marajo	8.0069	20.05
Varzea, marinha	-7.9751	-31.02
River area	2.3673	27.65
Constant	10.8341	11.90
ln(s_{ij})		
Constant	0.9336	> 100
Log measurement error variance		
Perimeter / area	1,005.7	34.37
Squared	-24,518.8	-21.29
Constant	-11.1660	-62.76

The parameters are quantitatively much larger in magnitude than those of the previous chapter. This is one of the quirks of working with censored data—the parameters can be large even when the marginal effects are small or moderate. The marginal effect of a change in an independent variable x on the conditional mean is $m'(\Phi_b - \Phi_a) + s'(f_a - f_b)$, where m' and s' are short-hand for partial derivatives with respect to x. So if the cdf evaluated at a has a value similar to that when evaluated at b, the multiplier on m' can be quite small. Furthermore, if s is also a function of x, the marginal effect of a change in x can be opposite in sign to that of m'. The multiplier at the weighted mean is 0.8139 for the main regression from the previous chapter, and is 0.0317 for the regression in this chapter. The ratio of the two is 25.68, which means that to compare marginal effects here to those of the previous chapter, we need to divide the parameters here by 25.68 to get the true marginal effects evaluated at the means. Since only rainfall is non-linear in variables, we can simply multiply (or divide) the parameter estimates by these values to get the true marginal effects.

Upon doing this, we find that the marginal effects for most variables are smaller in the disaggregated method that in the standard method of the previous chapter. We recall, however, that the variables are not entirely comparable. In the previous chapter, most of the variables (i.e., those related to soil, vegetation, distance classes) were the proportion of each type found in the census tract. In the disaggregated approach they are still proportions, but because the disaggregated units are so small, the proportion is either a one or a zero—each gridcell consists of exactly one type of soil, one type of vegetation, etc.

The multiplicative factor in this disaggregated regression is highly responsive to rainfall. The weighted mean for rainfall is 2,174 millimeters per year—the value at which the multiplicative factor of 0.0317 was computed. Yet in the *cerrado* portion of the Legal Amazon, the normal mean annual rainfall is just about 1,600 millimeters. At 1,600 millimeters, the multiplicative factor is approximately 3 times higher, which means the marginal effect of all of the other parameters is 3 times higher at 1,600 millimeters than at 2,174 millimeters. The multiplicative factor in the aggregated model of the previous chapter is not very responsive to rainfall, and in fact is actually smaller at 1,600 millimeters than at 2,174 millimeters (though it is larger at 3,000 millimeters).

Figures 14 and 15 are similar to Figure 13 from the previous chapter. These figures show the effect of rainfall on clearing at a "typical" site (i.e., at the most frequently found soil type, vegetation type, etc.) Figure 14 is for the aggregated regression of the previous chapter, and Figure 15 is for the disaggregated regression of this chapter. The lower curve in each shows the situation far from roads and earlier settlement (i.e., *antropísmo*). The middle curve shows the effect of being within 25 kilometers of a primary road, and the top curve shows the effect of also being within 25 kilometers of previous settlement. We see in Figure 14 that the curves are relatively evenly spaced regardless of whether the rainfall is low or high. This shows that rainfall has very little effect on the marginal effect (the slopes of the middle and top curves) through the multiplicative factor. But in Figure 15, we see that the distance between either the middle or top curve and the bottom curve is very high at 1,600 millimeters, but very small after 2,100 millimeters, and that there is a large magnitude of slope at low rainfall, but a very flat curve at high rainfall. This shows that rainfall has a high multiplier effect on marginal values (in the figure, the marginal values shown are the parameters for being less than 25 kilometers from a road and less than 25 kilometers from *antropismo*).

This emphasizes the importance of rainfall in clearing for agricultural purposes. It says very strongly that roads or settlements are much less likely to cause clearing if the rainfall levels are too high to be hospitable to agriculture. It also says that the marginal benefits of protected areas are smaller in preserving forest in high rainfall areas, because these areas are not very endangered to start with. This is the result I had expected to find in the previous chapter but did not. Perhaps this shows the importance of trying to disaggregate the data as much as possible.

In order to see how well the aggregated FWNLS fit the data, I compared the absolute mean square error and the root mean square error to those of the ones found in the previous chapter. The aggregated method seemed to perform slightly better, but not by much. The mean absolute deviation was 10 percent lower in the gridded method (0.202 vs. 0.181), but the mean squared deviation was less than 2 percent lower (0.069 vs. 0.068). This was disappointing, but not entirely unexpected. The FWNLS is not as efficient as the tobit (maximum likelihood), as Ruud points out. Furthermore, censored data often have thick tails. Ruud suggests using a distribution other than the normal for censored analysis, such as the t-distribution with a small number of degrees of freedom, the Laplace distribution, or the logistic distribution.

Censored data analysis is not a fair test of the potential of this method to increase efficiency in estimates. Gains in efficiency from using the disaggregated explanatory variables can be lost in the lower efficiency of FWNLS relative to the tobit.

However, Figure 16 shows an advantage of using this method. It is a map of predicted deforestation by the more than 300,000 individual units used in the analysis. Compare this to Figure 17, which are the predictions from the previous chapter, and to Figure 18, which are the actual levels of deforestation by census tract observed in the 1996 Agricultural Census.

If we use the actual levels of clearing by census tract to scale the predictions made by the disaggregated method, we get Figure 19, which is a map showing prediction in and near the state of Amapá. Figure 20 shows the equivalent census tract measures of true clearing from the agricultural census. The “disaggregation” of the prediction would be useful for anyone interested in a more refined picture of deforestation than a census can give.

There was less refinement in the smaller census tracts than in the larger ones, because the smaller census tracts sometimes only had one or two grid cells to be disaggregated to, while the larger ones may have had hundreds. Paradoxically, the larger census tracts are in many ways the ones of least interest: the reason they are large is that there is little or no economic (and demographic) activity in them. What might be useful in future research using this method is instead of overlaying a regular grid, use an algorithm that subdivides each aggregate unit into *n* different polygons of similar size, where *n* might be in the range of 4 to 16 (though there is no specific restriction). This would let us peer into both the large and small aggregate units. How to subdivide the aggregate units is a programmer's challenge: it is too time-consuming to do the division by hand, and it is not an easy task to program, either. One possible route would be to find the containing rectangle for each unit polygon (easy for GIS software to do), then divide the containing rectangle into subrectangles, and finally intersecting the subrectangles with the unit polygon. This can result in tiny (and not very useful) sliver polygons, and the challenge then is to modify the algorithm to get rid of the slivers by somehow merging with neighboring subpolygons.

The values in Figure 19 were calculated by multiplying each $\mathrm{E}y_{ij}$ by a scale factor. If $\mathrm{E}y_i$ was greater than y_i, then I simply multiplied each $\mathrm{E}y_{ij}$ by the ratio y_i / $\mathrm{E}y_i$. If $\mathrm{E}y_i$ was less than y_i, then I multiplied each $(1 - \mathrm{E}y_{ij})$ by the ratio $(1 - y_i)$ / $(1 - \mathrm{E}y_i)$, then subtracted the result from 1. This second method simply reverses the direction in which the scaling is done. As a result, expected values are not scaled to numbers greater than 1. I acknowledge that this method of scaling is somewhat *ad hoc* from the perspective of the model. A more satisfactory method from a modeling perspective would have been to assume that each gridcell in the census tract had a "fixed effect" common to the census tract. The problem with implementing this is that it would require a complex algorithm, using either an optimization routine or a gridsearch, solving for the fixed effect one at a time for each aggregated unit. The solution would have had identical qualitative effects (i.e., the values would have been shifted in the same direction) and similar quantitative effects (i.e., the magnitude of the shift would have been similar) to the much simpler method I used. Since I had no stong *a priori* belief in the “fixed effect”, I chose the much simpler proportional scaling method to predict disaggregated values adjusted for known aggregate levels.

Chapter 4: A Bird's-Eye View of Deforestation

IBAMA's Satellite Data

Except for summary data provided by INPE and my own tabulations of data from TRFIC, the previous chapters have relied on census and survey data to study Amazonian deforestation. The time dimension was also largely ignored, because the main data on which I have relied—the 1996 Brazilian agricultural census—covered only the one period of time.[31] Reis and colleagues (notably Andersen et al.; Reis and Guzmán; Reis and Margulis) have used data from the five Brazilian agricultural censuses from 1970 to 1996 to examine the dynamic aspects of deforestation. However, this panel was not available to me.

This chapter differs from the earlier ones in that it is based on satellite data; it has an explicit time dimension and dynamic component; and it (unfortunately) focuses on a subset of the Brazilian Amazon, because of data limitations. The deforestation data are from the *Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis* (IBAMA), the Brazilian agency re-

---

[31] The time dimension was partially treated, in that I used nearness to 1976 *antropísmo* and pre-1995 settlements. However, it was not a true dynamic analysis.

sponsible for enforcing environmental regulations. Since their focus is on enforcement of deforestation regulations, their data necessarily concentrate on areas where violations are thought to be the highest. IBAMA kindly provided digital maps of 1996 forest cover and incremental deforestation for 1996-97, 1997-98, and 1998-99. These data were constructed through visual interpretation of 30-meter resolution Landsat satellite images.

I combined the latter 3 datasets into one that described deforestation that took place between 1996 and 1999. Together with the first IBAMA dataset listed, these two datasets describe the rate of new deforestation over the three-year period, and the level of deforestation up to 1996. These two topics—rate and level of deforestation—are the focus of this chapter. I also attempt to distinguish deforestation on already established farms from deforestation caused by new settlement.

The chapter begins with observations concerning what the two deforestation datasets show concerning patterns of deforestation. It then describes how the farmgate prices of beef and milk variables were generated, and areas which were intentionally excluded from the analysis. Tobit regressions are used to analyze deforestation rates and levels in areas with previous deforestation, and a probit regression is used to examine determinants of new settlement areas. It concludes with a summary of lessons learned from the analysis in this chapter.

The “Arc of Deforestation” is an approximately semi-circular area covering the “frontier” of deforestation in the Brazilian Amazon—areas currently experiencing high deforestation rates, spanning from Rondonia in the south and west, eastward through Mato Grosso, then northward through Tocantins and southern and eastern Pará. The IBAMA data that I have available for this chapter cover the southern portion of the “Arc of Deforestation”. I restrict my attention to areas corresponding to Landsat scenes for which deforestation data were reported in all three years (1997 through 1999), and for which I have baseline data for 1996. I decided to create a series of artificial units of observation (since I did not have information on boundaries of farms and ranches), 5-kilometer square gridcells. In Figure 21, they appear as tiny blue, green, orange and pink squares. Each Landsat scene (the overlapping, tilted, square, see-through windows in the map in Figure 21) is approximately 180 kilometers square.

Figure 22 is a closer look at deforestation, and is the area enclosed by the red rectangle in Figure 21. In Figure 22, the pink, green, orange, and blue areas are 5 kilometer squares. These are overlaid with the smaller gold and red squares, which are 30 meter squares. In both Figure 21 and Figure 22, we see that new deforestation tends to be located very close to previous deforestation. First, in Figure 21, the green squares show that most of the deforestation that occurred between 1996 and 1999 occurred in the same 5 kilometer squares where pre-1996 deforestation had occurred. Second, we note that the blue squares are generally surrounded by green squares. This means that areas that have had previous deforestation but no new deforestation are at the "core" of new deforestation, suggesting a pattern of an expanding frontier, where "old" areas have finished their deforestation, and newer, adjacent areas are in the process of finishing their deforestation. This pattern is more obvious in Rondonia than in Mato Grosso. Rondonia is noted for the organized colonization schemes designed for small farms on plots of around 100 hectares in size.

The third thing we note is that most of the orange is just outside and adjacent to the green. That is, "frontier" deforestation tends to simply extend the previous deforestation. This last observation is confirmed in Figure 22, where we see the red new deforestation is almost always adjacent to the gold pre-1996 deforestation.

These figures suggest that deforestation starts in a core, and then expands outward. In order to examine this hypothesis further, I created rings of 5-kilometer gridcells. The frontier or outer edge of 1996 deforestation appears in orange in Figure 23, and is given the designation of "Neighbor Level 0", which I abbreviate as "nbr0", "nbrs = 0", or "the '0' ring". These are the 5-kilometer gridcells which have some pre-1996 deforestation, and have neighbors that have no pre-1996 deforestation. The 5-kilometer cells immediately adjacent to this ring and part of the interior (i.e., they have pre-1996 deforestation) is designated at the "–1" ring; adjacent and interior to that is the "–2" ring; and adjacent and interior to that is the "–3" ring. Everything more interior to that is designated as "–4". Adjacent and outside the "0" ring is the "1" ring; "2" and "3" are analogous. Ring "4" is everything not deforested by 1996 and external to ring "3".

Figure 24 shows the indigenous and protected areas in two shades of green, based on Arc View shapefiles from IMAZON and *Instituto Socioambiental*. Comparing Figure 21 to Figure 24, we see a high correlation between indigenous and protected areas on the one hand, and deforestation prevention, on the other. The multivariate analysis should at least partially distinguish the influence of the protected and indigenous areas from the desirability of the land for deforestation, as indicated by farmgate price, soil, and rainfall.

I am concerned that either the measurement of deforestation (i.e., omissions or misinterpretations), or the actual nature of deforestation might be different in forest and non-forest (i.e., *cerrado*) portions of the study area. Figure 25 shows the non-forest areas (based on *Ministerio da Agricultura* et al. 1988). The visual comparison of Figure 21 and Figure 25 does not give the impression that non-forest areas have different patterns of deforestation than strictly forest areas, though this figure is not fine enough to show all of the potential differences between these two types. In the statistical analysis which follows, I check for differences.

The model for our analysis was already presented in Chapter 2. From this model and the basic principles of optimal control, we know that in steady state, d, the rate of clearing (i.e., annual clearing) equals 0; and D, the total clearing, depends on prices and agricultural suitability measures. However, along the path to steady state, the optimal action, d^*, depends as well on D. The analysis we are about to do should shed some light on the optimal path.

This chapter is based on an unpublished manuscript by Thomas, Chomitz, and Arima, but does not include the simulation of tax policies that was included in that paper, and adds new material not reported there. Other studies that have analyzed the effect of previous deforestation on current deforestation rates have shown that proximity to past deforestation is a strong predictor of current deforestation (Liu et al; Mertens and Lambin; Alves). Chomitz and Thomas—and my work already presented in Chapter 2—find that agricultural conversion as of 1996 is strongly influenced by proximity to pre-1976 clearing, controlling for agroclimatic conditions and current road access.

Data

I have already discussed the deforestation data from IBAMA, as well as protected and indigenous areas, and non-forest areas. In this section I will briefly discuss the other variables used in the analysis.

Farmgate price derivation

Farmgate prices of beef and milk were imputed based on surveys of Amazonian slaughterhouses and dairies conducted by IMAZON (2000d, e, f, g) The surveys determined purchase prices for cattle and milk at these processing facilities, together with typical per-kilometer transport costs. After adjusting transport costs for road quality, and allowing for the possibility of river and off-road transport, I constructed a friction grid representing the cost of transportation of cattle across each grid cell in the sample. I used GIS methods to compute for each grid cell the maximum farmgate price of beef, where the farmgate price is the price at the slaughterhouse, less transport costs to the slaughterhouse, less fixed costs of loading. I used a similar procedure for milk, but imposed a maximum radius of seven hours transport time to the dairy. The resulting farmgate price map is in Figure 26.

Excluded areas

Some additional exclusions were necessary for applying the model. Because I wanted to focus on the decision to convert to pasture, I excluded relatively small areas where soybeans were predicted to be profitable (based on Costa). I also excluded small areas where cloud cover made data unavailable.

Results

Table 18 shows the number of 5-kilometer gridcells in this study, by state. Rondonia has the most cells, followed closely by Mato Grosso. Especially once areas are excluded, the contributions of Amazonas and Pará are minimal. We see that almost a third of the cells contain more than 100 hectares (out of 2,500 hectares total) of either protected or indigenous areas. We also see that more than 40 percent of the gridcells would be excluded if we studied only the cells that are not within 10 kilometers of a non-forest area. If both exclusions are in effect, only 42 percent of the study area remains.

Table 18. Number of 5-kilometer square gridcells by state and various restrictions

State	Minimal exclusions	Excludes protected areas	Excludes non-forest areas	Excludes non-forest areas and protected areas
Rondonia	8,354	6,003	5,720	4,168
Amazonas	1,388	1,025	422	265
Pará	784	103	207	0
Mato Grosso	7,751	5,543	4,106	3,302
All	18,277	12,674	10,455	7,735

Notes:
1) A cell is excluded because of protected areas when the protected area (including indigenous areas) contained in the 2,500 hectare cell is greater than 100 hectares.
2) Non-forest areas and a 10-kilometer buffer around non-forest areas are excluded in columns claiming “excludes non-forest areas”.

I compared total deforestation and rate of deforestation for the study area and the area excluding non-forest areas for the nine rings mentioned earlier. The values for the two comparison groups were almost identical in every ring. Since excluding non-forest areas made no difference in the results (nor in any other cross-tabulations I do in this study), from this point forward, I do not report their results. Table 19 shows total deforestation, rate of deforestation, and number of 5-kilometer gridcells by ring for the study area and the area excluding gridcells with more than 100 hectares in protected areas.

Table 19. Deforestation statistics by ring, per 2,500 hectare gridcell

	Entire study area			**Excluding protected areas**		
Ring	**Mean hectares deforested, 1996-1999**	**Total area deforested before 1996**	**# of 5-km cells**	**Mean hectares deforested, 1996-1999**	**Total area deforested before 1996**	**# of 5-km cells**
–4	63	1,449	1,694	63	1,449	1,686
–3	99	1,105	548	95	1,118	533
–2	108	951	918	109	961	867
–1	101	769	1,857	104	772	1,667
0	86	331	3,810	95	350	3,125
1	24	0	3,465	33	0	2,348
2	6	0	1,688	10	0	846
3	3	0	1,153	8	0	457
4	0.5	0	3,144	0.8	0	1,145
All	*48*	*362*	*18,277*	*64*	*493*	*12,674*

Notes:
1) A cell is excluded because of protected areas when the protected area (including indigenous areas) contained in the 2,500 hectare cell is greater than 100 hectares.

The first thing we note from this table is that pre-1996 deforestation increases as we move from the edge or frontier (ring 0) to the core (ring –4). This is supportive of the idea that deforestation expands outward from a core area. The second thing we note is that the rate of deforestation is lower in the core than on the interior but near the frontier (though perhaps not right on the frontier). This suggests that in the core, some cells (and by implication, farms and ranches) have reached steady state (i.e., the desired level of deforestation), and deforestation has stopped. Recall the blue cells in Figure 21. It also suggests that deforestation rates slow as ranchers approach their desired level of deforestation.

The third thing we note is that most deforestation that takes place outside the frontier is located very close to the frontier. This again suggests expansion of deforestation from a core. This does not imply that nearness to deforestation is causing deforestation. It is equally plausible at this point in the analysis that farmgate prices are driving this phenomenon. That is, if settlers claim land with the highest profitability (high farmgate prices and good agricultural conditions), the next group of settlers would naturally settle nearby, since farmgate prices nearby are almost the same, and the agricultural conditions are likely to be similar. Furthermore, some infrastructure or public services may have been established as a result of the first cohort of settlers, making the area even more attractive.

Finally, in Table 19 we note that very little of the core and rings –2 and –3 are in protected areas, but as we expand outward, an increasing proportion of gridcells are in protected areas, rising to 64 percent in ring 4. The rise in rings 0 and –1 does not mean that the deforestation is taking place in protected areas, but in gridcells that have at least 100 hectares of protected areas. That is, the deforestation frontier is at least right up against protected or indigenous areas (which I treat the same in the regression).

Table 20 shows the impact of farmgate price of beef on total deforestation and rate of deforestation. Most statistics are broken down into “inside boundary” (rings –4 to 0) and “outside boundary” (rings 1 to 4). From this table we note, first of all, that except for the highest price category and a flattening in the middle categories, the general trend is that the higher the expected farmgate price, the higher the level of deforestation observed before 1996. Assuming that relative prices have been stable in the nineties and late eighties, this is what we would expect, with the areas receiving the highest price for beef exhibiting higher levels of deforestation.

The next thing we note is that rates of deforestation inside the boundary are lower in the top 3 price categories than in the second to fifth price categories. This is consistent with the idea that deforestation rates slow as ranchers approach their desired level of deforestation. We also see that deforestation rates, for the most part, rise with prices outside the boundary.

The next set of columns in the table show us how many gridcells are in each category, which tells us how reliable the levels and rates of deforestation are for that category. For example, there are only 10 gridcells inside the boundary with farmgate price of beef less than R$436 per MT. We would not want to put too much stock in the deforestation rate of 10 hectares or 64 hectares total deforestation before 1996.

Table 20. Deforestation statistics by farmgate price of beef, per 2,500 hectare gridcell

	Pre-1996 deforest ation, hectares	Deforestation, 1996-1999 (hectares)		# of 5-km cells		% of cells with some new deforestation		Rate (has. in 3 yrs) for cells with some deforestation 1996-1999	
Farmgate price (R$/MT)	**Inside boundary**	**Inside boundary**	**Outside boundary**	**Inside boundary**	**Outside boundary**	**Inside boundary**	**Outside boundary**	**Inside boundary**	**Outside boundary**
385-436	64	10	0	10	68	40.0	0.0	26	NA
437-486	294	104	6	181	403	79.0	11.2	132	50
487-559	406	114	17	991	1,163	65.1	12.2	174	138
560-606	649	107	9	1,354	1,229	68.8	8.7	156	105
607-642	578	100	17	1,228	1,194	67.4	16.3	148	106
643-674	1,039	76	39	2,122	404	70.6	33.7	108	116
675-712	1,058	89	62	1,489	279	61.9	29.7	144	208
711-759	845	60	47	503	56	52.9	25.0	113	189
All	*794*	*92*	*19*	*7,878*	*4,796*	*66.5*	*15.1*	*138*	*124*

Notes:

1) Data are for gridcells in which the protected area (including indigenous areas) contained in the 2,500 hectare cell is less than 100 hectares.

The next columns tell us the proportion of gridcells with new deforestation between 1996 and 1999. Inside the boundary, there is not much variation, except for the lowest category—which we have suggested is not reliable because of the limited number of cells—and the highest category, which shows a slight tapering. Outside the boundary, the trend is that the higher price categories have a higher percentage of cells with new deforestation than the lower price categories.

The last set of columns in the table show the rate of deforestation for cells with any deforestation at all. On average, the rates of deforestation in cells where deforestation is taking place is not very different inside and outside the boundary: 138 hectares to 124 hectares. There are too few observations in the price categories outside the boundary to make solid inferences, but inside the boundary, it appears that except for an unexplained deviation in the R$643 to R$674 category, the rates of deforestation across the non-extreme price categories do not vary much from each other, and certainly not in a way which reveals a trend. This seems to imply that prices have a weak or non-existent effect on rates of deforestation (but not on levels of deforestation) inside the deforestation frontier.

In the discussion for Table 19, I suggested that part of the tapering off in the rate of deforestation in the core was due to some ranchers stopping deforesting altogether, while others were slowing down. If the highest farmgate prices tend to be in the core—and in fact, they are—then Table 20 has shown how much of the slowing of deforestation is due to ranchers stopping deforestation, and how much is due to ranchers slowing their rates, with both explanations contributing to the modest slowing noted in the core.

Table 21 is similar to Tables 19 and 20, except that the focus is on the effect of deforestation levels on deforestation rates. The category that has no pre-1996 deforestation represents rings 1 to 4. We see from this table that the rate of deforestation rises to the peak that occurs when a gridcell of 2,500 hectares has between 400 and 800 hectares already cleared in it, and then falls to close to zero for the category of 1,800 to 2,500 hectares deforested[32]. We note that all of the categories seem to have a sufficient number of gridcells to make the results not subject to a high standard error. When we consider the rate of deforestation in cells which have at least some new deforestation, we once again observe an inverted-U relationship of the effect of deforestation level on deforestation rate, but the "U" is flatter on the left or lower end. We also see a surprisingly constant percentage of gridcells with new deforestation for categories between 100 and 1,800 hectares. Even above 1,800 hectares, more than 40 percent of the cells have new deforestation, which implies that the steady state level of deforestation for the type of soils and precipitation observed in the study area might well be above 80 or 90 percent of total land. Because IBAMA ignores reforestation, this conclusion may be premature, because reforestation over the same 3 years could conceivably exceed the new deforesta-

---

[32] By law, 80 percent of a farm establishment in the forest biome of the Amazon is to be kept in forest.

tion of 24 hectares (less than 1 percent of all land in the grid) observed for this category.

Table 21. Deforestation rate, by deforestation level (rates and levels are per 2,500 hectare gridcell)

Pre-1996 deforestation, hectares	Hectares deforested, 1996-1999, all cells	5-km cells	Hectares deforested, 1996-1999, cells with any new deforestation	% of cells with new deforestation
0	19	4,796	124	15.1
1-100	77	1,477	145	52.9
101-400	116	1,626	162	71.3
401-800	131	1,452	166	78.8
801-1,200	108	1,110	140	77.2
1,200-1,800	70	1,280	98	71.7
1,800-2,500	24	933	58	40.6
All	*64*	*12,674*	*136*	*47.0*

Notes:
1) Data are for gridcells in which the protected area (including indigenous areas) contained in the 2,500 hectare cell is less than 100 hectares.

Multivariate analysis of deforestation rates

While the cross-tabulations have shed light on the effects of price, level of deforestation, and nearness to the core of deforestation on the rate of deforestation, they have not answered some of the most important questions. This is because the proposed explanatory variables are correlated, and two-way cross-tabulations cannot sort out the competing effects. For this, we must rely on multivariate regressions. Since the effect of prices inside the deforestation frontier appears to differ from the effect outside the frontier, I will analyze the areas independently.

Table 22 presents a tobit regression of the annual deforestation rate inside the frontier on farmgate prices, controlling for other agronomic determinants of profitability (rainfall and soils), protected area status, and prior deforestation. The second set of results in the table also controls for the nearness to the core; that is, in which "ring" the cell is located. Because the tobit is a nonlinear functional form, Figure 27 plots the partial effect of some of the most important variables.

In the top graph in Figure 27, we see that farmgate prices have little effect on rate of deforestation. In fact, their effect is opposite what we might have anticipated. That is, all other things being equal, we might expect that higher prices result in faster deforestation, due to the higher incentives to clear the land more quickly. But we observe a slight negative trend, with the rates falling just over 25 percent over the feasible price range of R$450 / MT to R$750 / MT (approximately representing the first and ninety-ninth percentiles of observed prices). Indeed, over the interquartile range of observed prices, the drop is less than one percent.

I believe that the reason we observe a decrease in rate of deforestation as prices increase is because the cubic specification for the level of deforestation imperfectly captured all of the effect of the level, and the parameters on price compensated. Other variables—like soil type or rainfall—may have also imperfectly captured their own effects (recall that this is a linearized specification of a highly non-linear function reviewed in Chapter 2). Alternatively, the assumption that the price variables were relatively constant through time may have been in error, and instead of observing the optimal path from no deforestation to steady state deforestation, we may instead be observing movement from one equilibrium to another as the critical parameters change.

Table 22. Tobit showing the effect of prices on annual rate of deforestation inside deforestation frontier

Variable	**param**	**t-stat**	**param**	**t-stat**
Log likelihood	*-34,618.67*		*-34,593.19*	
Rain (mm)	3.3643	1.87	3.6547	2.03
Rain, squared	-0.0018	-2.02	-0.0019	-2.15
Rain, cubed	3.19e-07	2.16	3.34e-07	2.26
Farmgate price of beef (R$/MT)	-4.1255	-1.92	-4.1060	-1.91
Farmgate price, squared	0.0071	1.96	0.0070	1.93
Farmgate price, cubed	-4.01e-06	-2.01	-3.94e-06	-1.97
Within 7 hours of a dairy	-7.6656	-3.09	-9.9622	-3.95
Protected area (hectares)	-0.0155	-10.50	-0.0142	-9.58
Hectares deforested by 1996	0.1330	14.84	0.1190	12.88
Deforestation, squared	-1.15e-04	-11.42	-1.08e-04	-10.61
Deforestation, cubed	2.18e-08	7.26	2.08e-08	6.82
Soil: minor root restricting layer	-19.900	-3.66	-18.565	-3.42
Soil: seasonal moisture stress	11.497	4.64	10.845	4.37
Soil: excessive nutrient leaching	5.6966	1.67	5.1186	1.49
Soil: low water holding capacity	4.9394	1.46	4.7710	1.41
Ring "–3"			-0.7333	-0.19
Ring "–2"			-1.5581	-0.47
Ring "–1"			-4.1239	-1.39
Ring "0"			-17.160	-5.57
Intercept	-1,256.2	-0.93	-1,462.0	-1.09

Notes:
1) There were 8,827 observations.

The other thing to note from the top graph in Figure 27 is that protected areas seem to reduce deforestation inside the frontier by around 40 percent.

The bottom graph in Figure 27 shows that the level of deforestation has a major effect on the rate of deforestation. This graph shows the rate of deforestation increasing until around 30 percent of the gridcell has been deforested, and then declining until all land has been deforested. This graph and Table 21 help us see that deforestation is a relatively slow process in most cases. The peak rate of deforestation in the bottom graph of Figure 27 is just over 60 hectares per year. This compares reasonably well with the cross-tabulations of Table 21, where we observe 131 hectares deforested over 3 years (about 44 per year) in the 401 to 800 hectare category—and that category includes some cells with protected areas and poorer agroclimatic conditions than we assumed for the graph. Now, if we assume ranchers deforest at a rate of 60 hectares per year regardless of the level of deforestation, it would still take 42 years to completely deforest a ranch the size of a gridcell.[33]

Table 22 shows a small negative effect of being near enough to a dairy to make selling milk feasible. Location within this radius slightly moderates the impact of beef profitability on deforestation, lowering the peak by between 7 and 10 hectares per year. It is possible that presence of a dairy industry promotes intensification, reducing pressures for pasture expansion.

---

[33] The average rate for this graph, using trapezoidal approximation, is 39.4 hectares per year.

The effect of rain on deforestation levels appears to be U-shaped. The curve has a value of almost 68 hectares deforested per year at 1,650 millimeters of rain per year; dips to around 56 hectares deforested per year at 2,100 millimeters; and rises to near 68 hectares deforested per year at 2,350 millimeters per year. The work in the earlier chapters led me to expect a relatively flat effect over this rainfall range—and certainly not a U-shape. However, it may simply reflect particular variations that are due to unmeasured variables. In any case, the effect of rainfall on deforestation rates—in the relatively narrow range observed in this study area—is similar in variation to that observed for farmgate prices, and is not close to the degree observed in the effect of deforestation level on deforestation rate or the effect of protected areas on deforestation reduction.

Soil type appeared to have a significant quantitative effect on rate of deforestation, with the "least preferred" soil (in terms of estimated parameters) —soil with a minor root restricting layer—leading to 30 hectares per year less deforestation than the most preferred soil—that with seasonal moisture stress.

Adding dummy variables for rings had little effect on the other parameter estimates or their standard errors. As it turned out, with the core (ring "–4") the omitted category, dummies for rings "–1" through "–3" had no statistical significance either individually or jointly.[34] However, the frontier ring "0" was highly statistically significant. Its parameter indicated that the deforestation rate on the frontier was 17 hectares per year lower than in the core. This clearly reflects the fact that the true decision-making unit—the farm or ranch management—does not fit neatly into the artificial units, the gridcells. That is, on the frontier, many of the gridcells are only partially claimed by ranches, whereas in the more interior rings, most or all of the gridcell is likely to be claimed by farms or ranches. The slower rate on the frontier may also possibly reflect a slower rate of deforestation for ranches just starting out. This latter hypothesis is only partially supported by cross-tabulations in Tables 19 and 21.

Given the preceding comments about the frontier cells (ring "0") being only partially settled, I re-estimated the tobit, excluding frontier cells. I was especially concerned that the rising portion of the bottom graph in Figure 27 was caused by including these cells. I will not show the tobit results or graph here, but the result was that the rising portion was still present, but the difference in deforestation rates at 0 and at the peak was only 17 hectares per year instead of 29 hectares per year.

---

[34] Because there is likely spatial autocorrelation here which has not been controlled for, and because these results are from a tobit regression, the parameter estimates are likely to be inconsistent. Therefore, discussions about statistical significance should not even be our primary concern.

I was also concerned that pioneers—settlers who try to get ahead of the frontier (here, I mean the “true” boundary of productive, non-speculative agriculture), so that they can sell their land when the frontier arrives—might be causing the rise. They tend to have smaller farms, and so their pattern of deforestation could easily be different than those who are more permanently settled and who are not relying on capital gains in land value for their main profit. In Figure 22, the settlers would likely be in the blue gridcells at the bottom and just left of center. These patches of deforestation are 10 or more kilometers away from the fishbone pattern of settled deforestation. These pioneers would likely appear in the 1 to 100 hectare category of Table 21, and could explain why the percentage of cells with new deforestation was smaller for this category than for the four categories below it in the table with more pre-1996 deforestation.

Trying to eliminate pioneers from the analysis, I ran another tobit (equivalently, a simple OLS, since there were no censored observations), excluding cells with no new deforestation, as well as cells on ring “0”. It turns out that the rise in the curve from 0 to the peak—and the fall in the curve from the peak to 2,500 hectares—were both shortened, but not eliminated. The rise was only 12 hectares, while the fall was 30 hectares. Eliminating all cells with no new deforestation, I removed those cells which truly had reached steady state levels of deforestation. So I ran a similar tobit, but this time only eliminated gridcells not on ring “0” and that had stopped deforestation in a cell with less than 500 hectares of total deforestation when the deforestation ceased. This time, the rising portion was eliminated, suggesting that those interested principally in agriculture and not primarily in capital gains start with a high rate of deforestation, rather than starting slow and building to a higher rate of deforestation. This is exactly what was predicted in the phase diagram presented in Chapter 2.

However, a possible explanation for the rising portion at the beginning of the curve (if the reader chooses not to accept the justification I gave for eliminating the cells that I did in the above exercise) is that when a household settles a new plot of land, their resources of capital and labor are also devoted to getting settled, which might include building a house and other start-up tasks. Once these other tasks are significantly completed, then all resources can be focused on clearing land and planting food crops. The point is, however, that other objectives not incorporated in the simple dynamic model of Chapter 2 could explain deforestation starting off slowly, increasing in speed, then finally tapering, as the dynamic model predicted.

For areas outside the frontier, Table 23 presents a probit regression for the likelihood of new deforestation taking place inside the gridcell. Figure 28 illustrates the effects of greatest interest to us from Table 23—those of the farmgate price of beef on probability of a cell having deforestation in it during the three-year period from 1996 to 1999. The graph shows three different lines. The top green curve shows the probability of deforestation in ring "1" for a typical soil (seasonal moisture stress) and rainfall (2,000 millimeters), outside the area where dairy production is feasible, and with no protected area. The middle blue curve shows the effect of a gridcell being completely within a protected area, but still in ring "1". The bottom red curve shows the effect of being in ring "4" but not in a protected area. The soil, rainfall, and dairy feasibility assumptions are the same for all 3 curves.

Table 23. Probit showing the effect of prices on the probability of clearing a gridcell outside the deforestation frontier

Variable	**param**	**t-stat**	**marginal change**	**variable mean**
Log likelihood	*-2,429.7*			
Rain (mm)	-0.0114	-3.86	-0.0013	2,084
Rain, squared	2.84e-06	4.02	3.15e-07	4.4e+06
Farmgate price of beef (R$/MT)	-0.0090	-2.20	-9.94e-04	575.49
Farmgate price, squared	9.25e-06	2.59	1.03e-06	336,441
Within 7 hours of a dairy	0.0122	0.20	0.0014	0.1304
Protected area (hectares)	-2.90e-04	-13.63	-3.22e-05	1,134
Soil: minor root restricting layer	-0.1828	-1.95	-0.0180	0.0897
Soil: seasonal moisture stress	0.0768	1.28	0.0087	0.3503
Soil: excessive nutrient leaching	-0.1991	-1.86	-0.0191	0.0516
Soil: low water holding capacity	0.1211	1.60	0.0143	0.1643
Ring "1"	0.9784	14.64	0.1384	0.3667
Ring "2"	0.4798	6.32	0.0680	0.1786
Ring "3"	0.1684	1.78	0.0207	0.1220
Intercept	11.58	-3.86	-0.0013	1

Note: There were 9,450 observations.

The curves all show that the probability of deforestation doubles when comparing farmgate prices on the low-end of those observed to the high-end of those observed. Furthermore, we see that the probability of deforestation does not change very much between R$400 per hectare and R$560 per hectare, and then begins to rise to R$700 per hectare. We note the effectiveness of protected areas in keeping deforestation from starting, reducing the probability by a third or more. This complements the results from Figure 27, where we noted that even if deforestations starts, it is slowed dramatically by protected areas. Finally, we see that the probability of settling and clearing land more than 15 kilometers from the frontier is very small relative to land less than 5 kilometers from the frontier. The rainfall effect, while statistically significant—is quantitatively small over the range 1,650 to 2,350 millimeters. The influence of being near a dairy is both statistically and quantitatively insignificant. The soil type, however, does have a noticeable quantitative effect, with a moderate measure of statistical significance.

One question of great interest is how much land ranchers and farmers ultimately intend to deforest. It is not clear that we can answer that question with the data available here, but we can perhaps try to shed some light on this matter. If there is no deforestation between 1996 and 1999 but there was some pre-1996 deforestation, it seems reasonable to ask whether the reason for no new deforestation is because the cells have reached their steady state level. It turns out that 35 percent of the cells on the interior (including frontier cells) have no new deforestation.

If it is not true that these cells have reached steady state, then the alternative conclusion is that it is common to have long pauses between periods of deforestation, which means that a mean annual deforestation rate of 40 hectares per year might really mean that 200 hectares is cleared every fifth year, or 400 hectares every tenth year—a kind of “pulse” deforestation rather than a steady process. This “pulse” deforesting is consistent with the finding of Brondízio et al.

Table 24 shows the results of a tobit, regressing pre-1996 deforestation on the same variables we have been using throughout this chapter. The tobit was restricted so that it only considered cells with pre-1996 deforestation and which had no new deforestation after that, and eliminated cells on the fron-

Table 24. Tobit showing the effect of prices on the ultimate desired level of clearing at steady state

Variable	**param**	**t-stat**
Log likelihood	*-8,483.7*	
Rain (mm)	154.27	2.239
Rain, squared	-0.0706	-2.103
Rain, cubed	1.06e-05	1.952
Farmgate price of beef (R$/MT)	19.455	2.684
Farmgate price, squared	-0.0163	-2.902
Within 7 hours of a dairy	-3,310.4	-2.306
Farmgate price of milk	7.7204	2.379
Protected area (hectares)	-0.0672	-1.052
Soil: minor root restricting layer	-502.37	-4.020
Soil: seasonal moisture stress	-331.05	-6.109
Soil: excessive nutrient leaching	-137.40	-2.475
Soil: low water holding capacity	-809.52	-7.763
Intercept	-114,942	-2.460

Note: There were 1,128 observations.

tier (i.e., ring “0”). Because of the existence of pioneer areas that have not reached their long-run steady state, I further restricted this regression to be for cells with more than 500 hectares of pre-1996 deforestation.

Key results are graphed in Figure 29. In the top panel, we see that above R$600 / MT, increased beef prices result in a lower desired level of deforestation. This result is robust to dropping all other explanatory variables, one group at a time (e.g., all soils), or all at once[35]. It is not clear why this would be, but seems to point to intensification of land use in high farmgate price areas[36]. This result is consistent with the cross-tabulations in Table 20.

In the top panel of Figure 29, we also see that in the long-run, the impact of protected areas—once deforestation has started—might be limited. In the cells which were used in the regression, only 4.5 percent had any deforestation. However, these are more than 92 percent indigenous areas only. That is, the parameter is not a good measure of the true impact of protected areas once deforestation has started, because these are not truly protected areas such as national forests, but instead are indigenous areas. Deforestation is permitted in indigenous areas. Furthermore, the part of Rondônia in this

---

[35] In some of these alternative specifications, the value of beef prices at which maximum deforestation occurred sometimes rose as high as R$650 per hectare.

study area is the exact area where there was much concern expressed by the World Bank in the late 1970s and early 1980s that the rights of the indigenous peoples were not being protected (Redwood; Lele et al.).

In the second panel of Figure 29, we see that increases in the farmgate price of milk leads to higher levels of desired deforestation. This contrasts with the slower rate of deforestation for dairy-feasible areas, reported in Table 22. It is possible that relatively new dairies have caused a shift from beef herds to dairy herds, and that during the shift, new deforestation is limited. But for the longer-term, it is more reasonable to expect higher prices to lead to higher deforestation.

The third panel of Figure 29 shows that the highest level of deforestation is found at around 1,950 millimeters per year, with a considerably lower level of deforestation noted with higher rainfall levels. The drop is steep: for a 350 millimeter change in annual rainfall, desired level of deforestation drops 43 percent. The fall is consistent with the results of Chapters 2 and 3, but the rise noted at lower levels of precipitation is not. Nevertheless, it is very believable that over a short range in the low-end of Amazonian rainfall (values < 1,950 millimeters per year), deforestation would increase as rainfall increases.

---

[36] The effect appears robust even when dropping restrictions on pioneers, boundary cells, and even allowing for deforestation to be on-going in the 1996 to 1999 period.

Using the results of Table 24, I predicted the desired level of deforestation for all cells. Statistics showing the goodness of fit for the predictions are presented in Table 25. The within sample predictions—those with at least 500 hectares of pre-1996 deforestation but no new deforestation—had a predicted mean less than 1 percent away from the actual mean, and in 45.5 percent of the cases, the prediction was larger than the actual. These figures simply confirm that the regression did a good job in fitting the model to the data.

Pioneers—those with less than 500 hectares of pre-1996 deforestation and no new deforestation—had vastly less deforestation than predicted, and 100 percent of the predictions were larger than the observed levels of deforestation.

Table 25. Summary statistics for desired level of deforestation predictions based

Variable	**# of cells**	**mean of pred**	**mean of actual**	**% of pred > actual**
Within sample	1,128	1,630	1,640	45.5
Pioneers	285	1,452	185	100.0
Others w/ pre-1996 deforestation	3,603	1,515	960	81.7

Note: There were 1,128 observations.

Finally, cells that had some pre-1996 deforestation and additional post-1996 deforestation had a predicted mean of 1,505 hectares (out of 2,500), while having an observed mean of 960 hectares. This is as it should be if the original hypothesis is true, that the cells used for the regression are at steady state.

Since some of the cells in this group are likely to be near steady state, while others are likely to be far, the fact that 81.7 percent of the predictions were greater than the actual level of deforestation is a favorable statistic. Because of the unmeasured residual, we would expect that some of the gridcells that are near to steady state would have real values greater than predicted values, just as the "within sample" predictions showed slightly more than half of the actual values being greater than the predicted values.

While these statistics are not proof that the tobit of Table 24 is a reasonable estimate of the desired level of deforestation at steady state, their behavior is consistent with this hypothesis.

Conclusion

In this chapter, we have used satellite data for 1996 through 1999 to study the dynamic nature of deforestation. There are several key points to take away from this study. First is that deforestation patches seem to lead to expansion of deforestation outward, even after we have controlled for prices and agroclimatic conditions. Second, protected areas seem effective in limiting the start of patches, slowing deforestation once it starts, and putting boundaries on patches that are growing. While it is true that many protected areas are on land that has less agricultural potential, the effectiveness of protected areas is still significant after controlling for agricultural suitability. The first two points lead to the third point: pro-active use of protected areas might be an important tool, especially in the case of changing economic conditions, brought about by many things, but especially road construction, which raises farmgate prices for agricultural products.

The fourth point to take away is that in general, deforestation is incremental, averaging around 40 hectares per year in cells of 2,500 hectares that have been brought under settlement (as indicated by having some pre-1996 deforestation). This is most likely an indicator of both labor and capital shortages. It should give policy makers desiring to slow deforestation some hope, because at the micro-level (i.e., gridcell), total deforestation could easily take 60 years. Even if deforestation-control policies take several years to design and implement, the damage incurred by implementation delays seems like it would be limited.

Fifth, as we have seen in Chapters 2 and 3, agroclimatic conditions have a statistically and quantitatively significant impact on deforestation level and rate. In this chapter, we saw how soil type caused large differences in both the level and rate of deforestation; and rainfall had a particularly large effect on the desired level of deforestation.

Finally, while rates of deforestation on land that is settled seem not to be affected by farmgate price, or possibly negatively affected, the choice of which land to settle is positively affected by farmgate price. This implies that policies attempting to modify incentives—such as taxes or subsidies—might be limited in their effectiveness against farmers and ranchers once they have settled, but could have good effectiveness against those considering establishing a farm or ranch.

Chapter 5: *Avança Brasil* and Deforestation

Introduction

This chapter seeks to answer the basic question asked in this dissertation: How much deforestation will result from the planned road paving under *Avança Brasil*? It begins by reviewing four published predictions, pointing to the strengths and weaknesses of each. It then uses an almost identical regression to the main one in Chapter 2 to make my own predictions. Instead of relying on a single estimate, this chapter makes multiple estimates, restricting the analysis in various ways, looking at both gross and net deforestation, and using both agricultural census and satellite data. It then briefly discusses the estimates in terms of possible policy avenues to pursue.

The Brazilian government's current plan for the economic development of the Amazon is called *Avança Brasil*. It is an ambitious plan to invest $45 billion (Cattaneo) to construct and improve infrastructure, so that products produced in the Amazon can be transported to international markets more cheaply. The plan includes construction of railroads, waterways, hydroelectric dams, and paving several major roads. It has come under attack by noted environmentalists.

In its defense, Mr. José Paulo Silveira, who at the time of his letter to *Science* in 2001 was Secretary of Planning and the *Avança Brasil* Plan Coordinator, pointed out that the infrastructure interventions have been carefully planned to avoid a rush toward new deforestation. Evidence of such care include paving only existing roads rather than creating new roads, use of waterways instead of highways, using natural gas to replace oil and to limit the need for new hydroelectric dams, and dams designed to "minimize reservoir size and impact."

Even with care taken to minimize the impact of the infrastructure development on deforestation, we would expect some impact on the Amazonian forest. The question is, "How much?" There have been three well-known attempts to answer this question, and one lesser-known attempt. The first two were published just one week apart in January 2001: Carvalho et al. in *Nature*, and Laurance et al. (2001a) in *Science*. Carvalho et al. was less than a one-page summary of a report by the *Instituto de Pesquisa Ambiental da Amazônia* (IPAM), the *Instituto Socio-Ambiental* (ISA), and The Woods Hole Research Center (WHRC).[37] Laurance et al. did not seem to have an earlier report, but they did provide extensive supplementary material on the *Science* webpage. It was the report of Laurance et al. to which Silveira was respond-

---

[37] Andersen et al. attribute this report to Nepstad et al. (2001).

ing. The third well-known estimate was published at the end of 2002 as a key part of a book reporting on deforestation in the Brazilian Amazon by Andersen et al. The fourth, lesser known estimate by Cattaneo was published by IFPRI, also at the end of 2002. One of the main purposes of this thesis is to add another estimate to the debate.

Environmentalists worry that paving major roads will give year-round access to areas that are not highly developed, allowing loggers, miners, and settlers to move in and clear trees at record rates. IPAM et al. believe that logging increases the flammability of forests, which become particularly dry during *El Niño* events[38]. Furthermore, they are concerned that increased deforestation and wildfires will lead to decreased rainfall, which would provide undesired feedback that might increase fire threat and lead to even more damage.

Those in favor of development see ranches, farms, and agricultural businesses hindered by the high cost of getting products to overseas markets. While the immediate impact of infrastructure development is to help wealthy agro-business people,[39] we would expect some spillover to all sectors because

---

[38] See also Nepstad et al. (1999) and Cochrane et al. (1999).

[39] This was a point raised by Laurance et al. (2001b) in their reply to Silveira. The main item for export is likely to be soybeans. There appear to be economies of scale in land in soybean production, which means that soybeans are naturally produced on large farms. Beef might also become an export prod-

of supporting industries and because of demand for inputs and consumer products.

The Four Predictions

Laurance et al.

Laurance et al. model deforestation rates with and without the *Avança Brasil* projects. To understand their model requires reference to the technical details posted on the *Science* website.[40] They overlay the 1995 road network on the 1992 deforestation data from the Tropical Rain Forest Information Center (TRFIC) of the Basic Science Remote Sensing Institute (BSRSI) at Michigan State University, where one of the coauthors (Cochrane) works. Their Supplemental Figure 1 shows the road network, including the present and future road surfaces (paved and unpaved). Their Supplemental Figure 2 shows a graph of the tabulations of deforestation for various categories measuring distance to paved and unpaved roads. The graph shows deforestation starting

---

uct, and this can be produced on small- and medium-size farms. However, foot and mouth disease has regularly kept beef from most of the Amazon from being exported—even across state lines. Also, Kaimowitz and Angelsen (2001) report that Belém slaughterhouses are demanding a higher quality animal for the local consumer, and that small-ranchers are having difficulty upgrading their livestock, and therefore being squeezed out of business. It seems sensible that the overseas consumer might be as demanding as urban Amazonians, and that these small farmers might have difficulty taking advantage of the potential export market, even if the problem with foot and mouth disease is eradicated.

[40] http://www.sciencemag.org/cgi/content/full/291/5503/438/DC1, which I downloaded on May 17, 2001.

as equal for the two road types in the 0 to 10 kilometer category, with deforestation further out from unpaved roads fading more rapidly than that for paved roads.[41] Using their Figure 2 as a guideline, they compute possible changes in deforestation.

Without *Avança Brasil*, they believe deforestation will be 269,000 to 506,000 hectares less per year. Their Supplemental Table 4 outlines their baseline deforestation prediction, and in their prediction for deforestation with *Avança Brasil* they predict annual deforestation will increase from 1.89 million hectares per year (the average they calculated for the 1995 to 1999 period[42] using INPE published data) to between 2.16 million and 2.40 million hectares per year, depending upon whether one looks at their optimistic or pessimistic scenario.

Laurance et al. believe that instead of the infrastructure development proposed by *Avança Brasil*, investments should be made which would encourage high-value agroforestry and perennial crops, so that the land would develop intensively rather than extensively.

---

[41] Distance categories are 0-10, 11-25, 26-50, 51-75, and 76-100—not too dissimilar to my model.

[42] This figure may be a little high, since INPE reports that from August 1995 to August 2001, gross deforestation has averaged 1.67 million hectares per year.

IPAM et al.

While the tone and recommendations are similar to Laurance et al., there are some differences found in IPAM et al. While they are against most of the paved roads, they did seem to be fairly positive about the prospects of paving the stretch of the Transamazon Highway between Marabá and Rurópolis, as long as it was accompanied by investments in schools, health care, and technical assistance. This road project received good reviews because it was already densely populated, and the population is stable. Their hope is that investment will lead to intensification and a reduction of slash-and-burn, with farmers adopting other types of agriculture or techniques for cultivation and fertility maintenance.

IPAM et al. have strong reservations about the other road projects. While they acknowledge that the roads are likely to reduce transport costs substantially for soybeans and other exported agricultural commodities and products, they are concerned that opening up year-round accessibility will lead to rampant logging and settling, and will stretch the government's ability to monitor deforestation and provide public services. They estimate that over the next 25 to 35 years, projected deforestation within 50 kilometers of the proposed *Avança Brasil* roads will average between 3,429 and 10,800 square kilometers per year. Their lower figure adjusts for deforestation that took place in the presence of unpaved roads; their upper figure does not. Using my own calculations based on their description of their methodology, they must have estimated that 41,000 square kilometers of deforestation (7.1 percent of the area) would take place along these roads if they remained unpaved, meaning that the upper range of their estimates for the increase in deforestation with paving is 229,000 square kilometers total, which implies an upper bound on the increased rate of deforestation equal to 9,160 square kilometers per year.

Andersen et al.

Andersen et al. take a highly complex regression approach which eliminates statistically insignificant variables randomly one at a time, then repeats the process of running the regression and removing another insignificant variable until only statistically significant variables remain. They repeat the entire variable elimination process 100 times to see which variables are robust.

Instead of satellite data, used by both Laurance et al. and IPAM et al., Andersen et al. use agricultural census data. Instead of one time period, which the other two teams used, Andersen et al. have a panel spanning the period from 1970 to 1995. However, they only have 257 cross-sectional units, one of them an entire state.

Andersen et al. do not use a model to guide their estimates, which makes projecting the results into the future a much bolder action. Nevertheless, they suggest that their regression shows that the road paving proposed in *Avança Brasil* will actually reduce deforestation by 15,580 square kilometers after 10 years. They state that their starting point was the roads in place in 1995. However, their starting level of clearing must have been the 1985 level of clearing (389,169 square kilometers, from p. 67), because their 1995 level of deforestation (485,809 square kilometers) exceeds their baseline 2005 level of deforestation (441,550 square kilometers)[43]. Using this interpretation, we see that Andersen et al. seem to predict only 2,619 square kilometers per year being cleared, in contrast to the just under 10,000 square kilometers per year actually cleared between 1985 and 1995 (Andersen et al. p. 67).

Cattaneo

Cattaneo uses a computable general equilibrium (CGE) model for Brazil to focus on Amazonian deforestation. In his study, he investigates a number of policies: devaluation, *Avança Brasil,* regulating and monitoring land titles more closely, agricultural technology changes, logging taxes, deforestation taxes, and non-timber forest extraction subsidies. The work he did on *Avança Brasil* shows that in the short run, deforestation would increase by 15 percent, and in the long run, by 40 percent. Cattaneo also finds that if Brazil continues

---

[43] The alternative explanation for their numbers is that under their baseline scenario (even without roads being paved), they predict a 4,426 square kilometer decline in deforestation between 1995 and 2005, and that road paving increases the decline to 5,984 square kilometers per year.

enforcing changes in titling laws passed in 1991, which strengthen discovery and prosecution of holders of fraudulent titles, deforestation in the Amazon could be reduced by 23 percent.

Cattaneo's prediction of an increase in deforestation of 40 percent from the implementation of *Avança Brasil* provides an upper bound on his estimate. Since he also predicts a 23 percent reduction in deforestation if the Brazilian government continues with its enforcement of recently passed laws designed to root out fraudulent land titles, I will treat the difference between his long term estimate of the effect of *Avança Brasil* and the impact of Brazil's new titling enforcement as the lower bound of his deforestation prediction. That is, it seems reasonable to argue that the policy environment in Brazil has changed from the 1970s and 1980s, and as a result of one example of that change, deforestation rates in the future should not be as steep as if the old policy regime prevailed.

Table 26 summarizes the four predictions. I use the baseline assumption of Laurance et al.—1.89 million hectares of annual deforestation—to compute percentage increase and decrease in deforestation for the IPAM et al. estimate, and I use the observed mean annual clearing (on farms) between 1985 and 1995—966,400 hectares—to compute the annual percentage decrease in deforestation from road paving from the Andersen et al. prediction.

Table 26. Deforestation Estimates under *Avança Brasil* by Three Different Teams of Researchers

Team	Annualized change in Deforestation
Laurance et al.	14.2% - 26.8%
IPAM et al.	18.1% - 48.5%
Andersen et al.	*-16.1*%
Cattaneo	17% - 40%

Notes:
1) Laurance et al. is based on their prediction that *Avança Brasil* will lead to an additional 269,000 to 506,000 hectares of deforestation per year, and their assertion that current deforestation averages 1,890,000 hectares per year, with this being the baseline for future deforestation without *Avança Brasil*.
2) IPAM et al. predicted net increase in deforestation along the roads to be paved by *Avança Brasil* will be between 120,000 and 229,000 square kilometers in the next 25 to 35 years. The upper value was adjusted by an assumed 41,000 square kilometers of deforestation along the unpaved roads. This means that the annual rate is projected to be between 3,429 and 9,160 square kilometers per year. Using Laurance et al. baseline gives the annual percentage increase.
3) Andersen et al. predict a reduction in deforestation of 15,580 square kilometers over 10 years. The observed mean annual clearing from 1985 to 1995 of 966,400 hectares gives the baseline.

Some comments on methods used

The Laurance et al. methodology is very similar to that of IPAM et al., and their results are in good agreement. However, they are vastly different from the results of Andersen et al.—resulting in completely opposite policy prescriptions! In terms of changes in deforestation, it is the Andersen et al. results that are the most shocking—completely opposite of what almost any model would predict. But the magnitude of increase in deforestation in Laurance et al., IPAM et al., and Cattaneo predictions is also surprising. My prior expectations for my own estimates were for them to show increased deforestation, though not as much as those predicted by Laurance-IPAM.

My first criticism of Andersen et al. is that their lack of a model leaves us with curious multi-variate correlations, but little to base predictions on. Their most intriguing result—decreased deforestation with increased paving—is only supported by their regression results from the 1985 to 1995 time period, but not at all for the 1980 to 1985 time period. Unfortunately, they did not publish means and interquartile ranges of their data, so it is difficult to delve more deeply into their work.

Kaimowitz and Angelsen (2001) point out that "economic theory suggests that cattle ranchers that adopt more profitable livestock technologies will be inclined to expand their pasture areas unless one of two conditions applies: 1) The new technologies depress livestock product prices by greatly increasing aggregate beef and / or dairy production... or 2) The new technologies require more capital, labor, or managerial efforts per hectare of pasture and cattle ranchers have limited access to those resources... Where neither of these two conditions hold, one would expect technological improvements in livestock production to have no affect on deforestation or to encourage it." While paving is not a shift in agricultural technology, it does represent a shift in potential profitability of ranching and farming, and the same principles apply in predicting its impact on deforestation. The paving of roads should actually lead to a broader market for beef and other agricultural products, and there is not an implied shift in the required inputs of labor and agriculture—so since neither of the conditions are met, we must conclude that there ought to be an expansion of clearing with paving.

If I had to point to one variable in the Andersen et al. analysis that would most likely explain their counter-intuitive results, it is land prices. We would expect paved roads to drive up land prices much more than unpaved roads, yet they do not appear to model how land prices evolve with the addition of roads. Since this parameter is positive and highly significant, it could offset the negative parameter on the interaction of clearing and unpaved roads.

Another item that could be driving their perverse conclusion is that they are using agricultural census data. More paved roads implies more urban and suburban area, which means that there is less room in each minimum comparable area (MCA)[44] for farms. Their "proportion cleared" variable uses the area of the entire MCA as the denominator, while the numerator is restricted to clearing done on farms. So, more paved roads means more urban areas, which in turn means less farms and therefore less clearing. Neverthe-

---

[44] "Minimum comparable area" is their term for their cross-sectional unit, and essentially it is either a 1970 *município,* or a joining together of two or more 1970 *municípios.*

less, whether forest is cleared for agricultural or residential purposes, it is still cleared.

The final suggestion as to what might have gone wrong in their estimation procedure is that it appears that they did not weight the regressions to adjust for land area. Here is an extreme example of why it is important to weight by land area. Let there be ninety-nine MCAs that total ten percent of the area of the Amazon and which are not growing at all. And let there be one MCA representing ninety percent of the land area of the Amazon growing at a rate of five percent per year. Weighting the observations by land area would let this one large MCA tell us that the land cleared in the Amazon is growing at a rate close to five percent; while an unweighted regression would tell us that land clearing in the Amazon is hardly growing at all. In their data, the whole state of Rondonia—which has grown at a phenomenal rate since the seventies—is only one MCA, out of a total of 259 used in their analysis. At the same time, geographically small MCAs tend to be those which were highly developed in 1970, and therefore would tend not to be growing rapidly in later periods. They would also tend to be the ones with the highest density of paved roads.

As for the Laurance-IPAM methodology, one wonders if the simple cross-tabulations are merely reflecting the correlation of paved roads and deforestation with some other variable that is really the pertinent one for the regression. As shown in Chapter 1, there is a very high correlation between principal roads and *antropísmo.* Whether *antropísmo* should even be an explanatory variable is somewhat of a philosophical question which I will address later. We have already seen in Chapter 2 that there is a vast difference in the effect of roads on deforestation if *antropísmo* is included in the regression.

In addition to *antropísmo,* if paved roads currently happen to be located in areas with good agroclimatic conditions, while unpaved roads are in areas with poor agroclimatic conditions, we would expect that simple cross-tabulations would show a higher expected deforestation rate than would materialize if *Avança Brasil* road paving were to take place.

One problem with their analysis, and mine also, is that to extrapolate deforestation rates, we need to assume that the policies and conditions of the present and future are the same as those of the past (or if not, control for how they have changed). Clearly, policies in the 1970s and 1980s which encouraged deforestation (e.g., tax and subsidy policies; incentives for farmers in other parts of Brazil to relocate to the Amazon) have been changed to limit the incentives for deforestation. Perz argues that in the 1990s, deforestation was no longer driven by in-migration as it was in the 1970s and 1980s.

INPE's data supports the idea of some type of shift in deforestation patterns from the 1980s to the 1990s: the average rate of deforestation between January 1978 and April 1988 was 2.20 million hectares per year, while between April 1988 and August 2001, deforestation averaged 1.70 million hectares per year. Note, however, that this is gross deforestation, which does not adjust for reforestation, which is probably sizeable. The main regressions in Andersen et al. showed that there was probably some type of shift between the early eighties and the late eighties and early nineties, as a number of their parameters changed. Nevertheless, when they made predictions for *Avança Brasil*, they assumed that the latter conditions would prevail into the future, therefore committing the same error as the other teams and I have.

Regression Revisited

The main regression in Chapter 2 suggests that the direct effect of roads on deforestation is modest. However, this regression did not distinguish between paved roads and unpaved roads. It is conceivable that their impact on deforestation could be vastly different. To examine whether there is a difference—and if so, its magnitude—I re-ran the regression, but this time distinguishing paved and unpaved roads. I used the map of Laurance et al. as the primary reference, but also compared it with the map of Nepstad et al. and the digitized road map I used earlier—which already had information on road surfaces—to compile two new sets of principal roads: one paved and one unpaved. In order for the experiment to work, I redesignated some roads as "principal" that in the earlier regression I had excluded. These upgraded roads were state roads that were either already paved, or were designated to be paved under *Avança Brasil*. I had already designated all federal roads as "principal". The resulting regression is found in Table 27.

Table 27. Differentiating the Effect of Paved and Unpaved Roads on Deforestation

	param	t-stat
Observations	6,693	
Log likelihood	4323.13	
Rain, annual, mm	-0.0018	-10.918
Rain squared	6.77e-07	10.826
Rain cubed	-8.26e-11	-10.642
Protected area	-0.0167	-8.053
Road buffer, <10km, paved	0.153	9.743
<10km, unpaved	0.1262	7.103
10-25, paved	0.0323	1.756
10-25, unpaved	-0.0037	-0.168
25-50km, paved	0.0376	3.475
25-50km, unpaved	-0.0192	-1.677
50-100km, paved	0.0085	1.390
50-100km, unpaved	0.0117	2.210
Settlement, <25km	0.0214	3.849
25-50km	0.0030	0.518
50-100km	-0.0070	-1.715
Antropismo, inside	0.3637	20.828
<25km	0.1360	18.333
25-50km	0.0278	3.924
50-100km	7.40e-04	0.198
City buffer, <25km	-0.1034	-1.504
25-50km	-0.0785	-2.408
50-100km	0.0480	4.637
Manaus, <25km	-0.6022	-2.039
25-50km	-0.1021	-0.916
50-100km	-0.1248	-4.820
Low organic mtr.	0.1702	7.709
Seas. excess water	0.0302	0.269
Minor root restr.	0.0356	4.742
Impeded drainage	-0.0293	-3.544
High aluminum	0.0596	3.152
Excess nutr. leach	0.0076	1.252

	param	**t-stat**
Low nutr. holding	-0.0056	-1.592
High P, N, & organic retention	-0.0276	-1.476
Low water holding	-0.0320	-3.864
Salin. or alkalinity	0.1171	5.188
Shallow soils	-5.69e-04	-0.106
Cerrado	0.1900	22.290
Northern savannas	0.0025	0.152
Varzea, Marajo	0.0133	1.167
Varzea, other	-0.0170	-3.185
Pantanal	0.2262	7.829
Mangroves	-0.1456	-2.152
Campinarana	0.0055	1.582
Dry forests	0.0328	5.923
Babaçu forests	0.0327	1.442
Varzea, fluvial	-0.1790	-7.815
Varzea fluvial *Varzea, Marajo	0.3802	10.317
Varzea, marinha	-0.4370	-6.626
River area	0.1761	4.756
Constant	1.5800	10.928
$ln(s_i)$		
Perimeter / area	692.45	47.569
Squared	-8,607.6	-41.196
Within 10 km of river	-0.0071	-0.195
Rain, annual, mm	-0.0014	-56.175
Constant	-0.4276	-8.946

I added a finer distance measure than I used in the regression of Chapter 2. I split the 0- to 25-kilometer category into a 0- to 10-kilometer category and a 10- to 25-kilometer category. The results show that paved roads lead to only slightly higher deforestation than unpaved roads. The parameter for the 0- to 10-kilometer category on paved roads is 0.153, while the one on unpaved roads is 0.126. The deforestation falls off rapidly in the 10- to 25-kilometer category, dropping to 0.032 and –0.004 (not statistically different than 0).

Breaking the 0- to 25-kilometer category into two groups revealed something quite important: almost all of the predictable deforestation takes place in the first 10 kilometers away from a road, and very little in the next 15 kilometers.

In the 25 to 50 kilometer categories, we find that paved roads have a small but statistically significant[45] parameter of 0.038 (not statistically different than the measure observed for the 11 to 15 kilometer category), and that unpaved roads have a measure of –0.019—a measure which is just barely significant at the 10 percent level—but for purposes of interpretation, probably should be considered equal to zero. Finally, in the 50 to 100 kilometer category, the paved roads have a parameter of 0.008, while the unpaved roads have a parameter of 0.011. The latter is statistically significant, though quantitavely small.

With the parameter estimates of paved and unpaved roads being so similar, one would not expect a paving project—even a massive paving project—to have a strong impact on overall deforestation. To test this for the case of *Avança Brasil* road paving, I predicted the level of deforestation for the *status*

---

[45] The statistical significance is based on the standard errors reported in the analysis. Since there is likely to be spatial autocorrelation, the parameters are likely to be inconsistent, and therefore we should not conclude too much from their statistical significance. I did not control for spatial autocorrelation, because the data is censored, and some of the techniques to deal with spatial autocorrelation in censored regressions rely on distributional assumptions which I felt were not appropriate for this analysis (chapter 5 of LeSage 1998; chapter 5 of LeSage 1999a; and chapter 7 of LeSage 1999b).

quo, and then predicted the deforestation if the paving projects had been completed. Actual deforestation (to be precise, “clearing”, since some of the area is *cerrado* and might not have been truly in forest), according to the 1996 agricultural census which I used in the regression, is 12.85 percent. Predicted deforestation before paving is 12.88 percent. Predicted deforestation after paving is 13.21 percent.[46] This is a 2.5 percent increase in the area deforested, or new deforestation representing 0.33 percent of the area of the Brazilian Amazon.

Table 26 shows that Laurance et al. estimate an increase of between 14.2 percent and 26.8 percent; IPAM et al. estimate an increase of between 18.1 percent and 48.5 percent; and Andersen et al. estimate a decrease of 16.1 percent. My estimate of an increase of 2.5 percent seems to be quite a reasonable intermediate estimate between the two camps (considering Laurance et al., IPAM et al., and Cattaneo to be in one camp). However, I decided to investigate further to see if I could explain the reason for the significant gaps between my estimate and those of Laurance et al., IPAM et al., and Cattaneo. Table 28 summarizes my investigation.

First, I decided to exclude the *antropísmo* and settlement variables from the regression. This leads to a substantial increase in predicted deforestation,

---

[46] These are in sample predictions. A few small census tracts near rivers were excluded from the regression.

Table 28. Deforestation Estimates for Now and under *Avança Brasil* or Various Datasets and Restrictions

					Predicted Deforestation[4,5]		
Type	**Biome**	**Dependent Variable**	**Variables[1]**	**Obs**	**Now**	***Avança***	**% change**
Agric. Census	All	Defor & regrow	All	6693	12.88%	13.21%	2.5%
	All	Defor & regrow	No antro	6693	12.66%	13.66%	7.9%
	Forest	Defor & regrow	All	5412	8.91%	9.12%	2.3%
	Forest	Defor & regrow	No antro	5412	8.40%	9.37%	11.5%
	Forest	Modified defor[8]	All	5412	7.47%	7.81%	4.5%
	Forest	Modified defor[8]	All[7]	5412	7.52%	7.86%	4.5%
Satellite (TRFIC, 1992)	All[2,3]	Defor & regrow	All	6283	10.26%	11.67%	13.7%
		Defor & regrow	No antro	6283	10.11%	12.51%	23.7%
		Current defor	All	6283	7.32%	8.03%	9.8%
		Current defor	No antro	6283	6.91%	8.41%	21.7%
	Forest[2,3]	Defor & regrow	All	5126	9.78%	11.19%	14.4%
		Defor & regrow	No antro	5126	9.63%	12.07%	25.3%
		Current defor	All	5126	6.97%	7.69%	10.3%
		Current defor	No antro	5126	6.53%	8.06%	23.4%
	No cities[6]	Defor & regrow	No antro	3972	8.01%	9.76%	21.9%
		Current defor	No antro	3972	5.34%	6.37%	19.3%
	Mills[9]	Defor & regrow	No antro	5126	10.13%	12.27%	21.2%
	Mills & cities	Defor & regrow	No antro	3972	8.35%	9.97%	19.5%
	Drop Rondonia	Defor & regrow	No antro	4317	9.07%	11.71%	29.1%

Notes:

1) “No antro” means “no *antropísmo*”. Settlements were also omitted when *antropísmo* was omitted.

2) With satellite data, non-omitted area is forest, deforested, and regrowth. Omitted area is cerrado, water, clouds, shadows, and missing scenes.

3) When using satellite data, excluded census tracts which had less than 100 hectares of non-omitted area.

4) Predicted deforestation is the mean of all predictions weighted by census tract area (census data) or non-omitted area (satellite).

5) Predicted deforestation is only for in-sample census tracts.

6) Excluded census tracts with a portion within 25 km of a city with a population of at least 25,000.

7) With an additional variable allowing for road interaction with *antropísmo*.

8) In this variable, I use the dependent variable of Andersen et al. who do not

consider natural pasture as having been deforested.
9) “Mills” is short for “timbermills”.

from 2.5 percent to 7.9 percent. The question is, “*Should* those variables be excluded”? The answer depends in part on how far into the future we want to predict deforestation occurring, and how much deforestation we believe occurred as a result of the long history of settlement in the Amazon near Belém and the region of Bragantina.

Table 29 is based on the agricultural census for states with at least some of their area in the Amazon. It shows that a reasonable amount of deforestation had occurred by the time roads were being built in the Amazon in the sixties and seventies. Both Laurance et al. and IPAM et al. want to attribute all deforestation that is observed in the 1992 satellite images to clearing which took place in the preceding 15 to 35 years.[47] Since my objective was to project ahead around 20 years, using the 1976 *antropísmo* as a baseline seemed to be the correct thing to do. Laurance-IPAM data were from 1992 (instead of 1996 for the agricultural census data), and so if they used *antropísmo* as a baseline, they would have only been able to project 16 years into the future. For them to make a projection of deforestation without using a baseline implicitly assumes that there was no deforestation prior to their starting point. In fact, they make arguments to the effect that if there were any deforestation when

---

[47] Laurance et al. suggest 15 to 25 years; IPAM et al. suggest 25 to 35 years.

the roads were originally constructed in the sixties and seventies, then the deforestation was so small as to be negligible.

Table 29. Historical Farm Area in Amazonian States

			Hectares		
	1950	**1960**	**1970**	**1980**	**1996**
Para	6,593,399	5,253,272	10,754,828	20,448,421	22,520,229
Mato Grosso	7,037,269	7,806,303	17,274,745	34,554,548	49,839,631
Amazonas	5,592,863	6,398,804	4,475,941	7,009,594	3,322,566
Rondonia	693,775	303,316	1,631,640	5,223,630	8,890,440
Acre	8,897,883	9,386,075	4,122,084	5,679,532	3,183,065
Roraima	595,795	869,582	1,594,398	2,463,106	2,976,817
Amapá	734,192	1,242,037	603,441	735,128	700,047
Subtotal	30,145,176	31,259,389	40,457,077	76,113,959	91,432,795
Tocantins	NA	NA	11,450,373	18,667,649	16,765,716
Maranhao	9,538,144	8,215,613	10,794,912	15,134,236	12,560,692
Total	39,683,320	39,475,002	62,702,362	109,915,844	120,759,203

Source: Agricultural census (IBGE 1998)

INPE's earliest measurement of deforestation in the Amazon was January 1978, with 15.2 million hectares of deforestation. By August 1992, it was 44.0 million hectares of gross deforestation. The 1978 figure was both gross and net, since there had not been any previous measurements. We know from agricultural censuses that at about the same time as INPE's first measurement (based on a linear interpolation of the 1975 and 1980 agricultural censuses), there were between 82 million and 96 million hectares claimed by farm operations in the Legal Amazon, including those in the *cerrado*. We also see in Table 29 that in 1950, there were almost 40 million hectares claimed (excluding Tocantins). Indeed, whether or not we omit Maranhão and Tocantins, land in farms in 1950 was about one-third of what it was in the latest agricultural census in 1996. So land has been farmed for a long time, and many of the processes of deforestation were set in motion before Laurance et al. assume they began. Therefore, they have probably overstated the impact of road construction by not controlling for the starting value of deforestation. Nevertheless, for the benefit of those who would vehemently disagree, I report values in Table 28 for calculations resulting from regressions that in some cases included the *antropísmo* buffer variables, and some cases which excluded them.

As I considered the Laurance-IPAM work, I realized that they had excluded most of the *cerrado* portion of the Amazon, so I also tried excluding that portion from the regressions. As we see in Table 28, using *antropísmo* as an explanatory variable, the effect of road paving on deforestation yielded a slightly smaller increase when we considered only the forest biome instead of the entire Amazon. However, ignoring *antropísmo* resulted in a sizable increase in the effect of road paving on deforestation. This latter result implies that road paving is more important in the forest biome than in the *cerrado* biome. If we believe that *antropísmo* should be included in the analysis, the result seems to imply that there are no differences in clearing between the forest and the *cerrado* portions of the Amazon when a road is converted from unpaved to paved.

The next two rows in Table 28 attempt to investigate the results of Andersen et al. in a crude way. Andersen et al. use a different dependent variable than I do. Their land clearing excludes natural pasture. I include it, because I believe that farm land that is used as pasture is degraded and is no longer "natural". The two rows show the results of using their dependent variable. The first prediction shows that there is a larger percentage change in deforestation when roads are paved using their definition of deforestation, which is the opposite of what I was expecting, since they predicted a decrease in deforestation with road paving. In the second prediction, I tried to imitate their interaction of roads and land clearing by interacting proportion of census tract within 10 kilometers of a paved road with proportion of census tract within the 1976 *antropísmo* areas. I did the same for unpaved roads and *antropísmo*. The outcome yielded highly significant parameters, both with t-statistics greater than 3. The parameter for unpaved roads interacted with *antropísmo* was more than 4 times larger than its paved counterpart, but the predictions for clearing if *Avança Brasil* were implemented were almost identical to those from the regression without these two interaction parameters. So my attempt to duplicate the results of Andersen et al. failed, meaning that I can offer no additional insight into their results.

Since I have the 1992 dataset used by both Laurance et al. and IPAM et al., I decided to tabulate the data by census tract. Their dataset classifies land as either forest, deforested, regrowth, *cerrado*, water, cloud, shadow, or missing (they were missing a small number of Landsat scenes). The rest of the rows in Table 28 show the results of my analysis using this data.[48]

The first two rows of this second section of Table 28 use a dependent variable equal to the sum of deforestation and regrowth divided by the sum of deforestation, regrowth, and forest. I call the denominator the "non-omitted" area. Census tracts with less than 100 hectares of non-omitted area were dropped from the regression. To my surprise, even when *antropísmo* was included in the regression, the increase in deforestation from paving roads was quite large. I had expected that the Laurance-IPAM results were driven by their use of cross-tabulations and their failure to take correlation with other variables into account. This appears not to be the case. The Laurance et al. range of estimates for increased deforestation are very close to the estimates I derived for the full regression (on the low end) and the regression without *antropísmo* and settlements (on the upper end).

My next thought was that the difference between their results and mine was due to the fact that they use gross deforestation instead of net deforestation. While it is true that in the agricultural census I use gross deforestation

---

[48] I ignore the measurement error in the data resulting from the satellite data file not being well geo-referenced to the shapefile with census tract boundaries, particularly in the western Amazon.

on farms—since I include land in fallow (not in use for less than 4 years) and vacant land (not in use for more than 4 years)—I naturally do not (and in fact, cannot, because the data are not available) include farms that have been totally abandoned (i.e., I cannot account for the portion of gross deforestation due to farms that no longer exist), but for which my discussion in Chapter 1 indicates might represent quite a large quantity of land. So I next report on the results using net deforestation (i.e., I exclude regrowth in the numerator of the dependent variable).

We see that a large proportion of the land was in regrowth, meaning that there is quite a gap between net and gross deforestation. We also see that the effect of paving on net deforestation is considerably less than on gross deforestation when *antropísmo* is included as an explanatory variable, but the predictions in percentage change in the two deforestation measures are not terribly different when *antropísmo* is omitted. Nevertheless, when all explanatory variables are used, there is still a large gap between the estimate of the effect of paving on deforestation between the results from satellite data and the results from census data.

To try to explain the troubling gap, my first thought was to try to make the regression using census data and the regression using satellite data as comparable as possible. So I re-ran the four satellite regressions, restricting the data to the forest biome. It was already partially restricted, because I excluded census tracts with less than 100 hectares of non-omitted area (i.e., forest or previously forested) from the satellite data analysis. But several census tracts in the *cerrado* biome had sufficient forest to get past this restriction, so I explicitly restricted the analysis to the forest biome. The next four rows in Table 28 show the results. They are all consistently larger than their unrestricted counterparts by 5 to 8 percent.

The next step was to consider the difference in what census and satellite data measure. Agricultural census data are necessarily restricted to active farms, which omits deforestation on abandoned farms, in urban and suburban areas, and in areas logged but not settled. As already mentioned, I do not have data on farm abandonment (except by state, which is too large an area to be useful). I do, however, have information about the location of towns and cities of various sizes. So to test whether the difference in results was driven by urban and suburban areas being missed in the agricultural census data, I re-ran the analysis. This time I omitted any census tract with a portion within 25 kilometers of the center of a city with at least 25,000 people. I found a reduction in the deforestation response to paving, but not of a magnitude that would explain all of the difference between satellite data and census tract data.

While I do not have data on logging operations, I do know the location of timbermills, thanks to my colleagues at IMAZON (2000h). I created buffers around the timber mills: 0-10 kilometers, 10-25 kilometers, 25-50 kilometers, and 50-100 kilometers. I computed the proportion of each buffer in each census tract, and used the proportion in the regression. The parameter estimates were highly jointly significant, the parameters tapered from 0.14 to 0.01, as we would expect, and two of the four parameters were individually statistically significant at the 1 percent level. Just as for the case of cities, nearness to timbermills—that is, logging operations—cannot explain all of the difference between the results for satellite data and census data, but it does seem to explain some of the difference. When cities and timbermills are used together in the same regression, the combined effect drops to 19.5 percent—not as much as one might expect if the two were completely independent. This is likely to be because a number of timbermills are in or near places with populations greater than 25,000.

It is important to consider whether the difference that controlling for cities and timbermills make might have nothing to do with logging outside of farms or areas of cities and suburbs. In other words, do the two experiments just conducted really explain some of the difference between the results from satellite data and those from census data? It may simply be that the timbermills locate where farms are expanding their areas, and that farms locate close to cities because of markets for crops and availability of inputs and consumer goods. But as we argued in Chapter 1, there is a great deal of logging that takes place outside of farms, and so even if timbermills get some of their timber from farms, most of it is coming from loggers working outside of the farms. We also saw in the regressions in Chapter 2 and in Table 27 that deforestation near cities with at least 100,000 is actually lower than otherwise would be expected out to at least 50 kilometers. That is, urban areas prevent farms from being located there, so analysis using agricultural census data reports that deforestation from farms is less near cities. So, the effort to control for logging and urban areas probably did explain real differences between the results from the two types of data.

The last thing I tested is whether all of the government-sponsored settlements in Rondônia might be driving some of the results. As the last row of Table 28 shows, the deforestation response to paving actually increases when Rondonia is omitted from the regressions.

So what have we learned from the results in Table 18? First of all, there are substantial differences in predictions, depending on whether the data are from an agricultural census or from satellite interpretation. We learned that some of the difference can be explained by urban and suburban areas, and some of the difference can be explained by logging. We can only guess at what explains the remaining portion of the difference. From my analysis in Chapter 1, it appears that at least one of the explanations for the difference is the number of abandoned farms. Perhaps forest fires account for another part of the unexplained difference (Cochrane et al. report that fires can be misclassified as true deforestation).

We also learned that there is a significant difference in response to paving roads between net deforestation and gross deforestation, with the former having significantly less response. This leads us into a philosophical issue of which we should be concerned with: net or gross? Those for whom fire hazard is an issue would argue that gross deforestation is what we should worry about, because a regenerating forest has less resistance to fire than an intact forest. Apart from this fire hazard argument, those who are concerned about carbon sequestration may not care.

What is my final prediction about the effect of *Avança Brasil* road paving on deforestation? Since it is more important to be concerned with deforestation from any source—not just farms—I choose results from satellite data over those of census data. So as not to choose between net and gross deforestation, I will choose two measures, and predict that *Avança Brasil* will lead to a 10.3 percent increase in net deforestation, and a 14.4 percent increase in gross deforestation, within the forest biome of the Legal Amazon. Since this is based on 1992 data, with a baseline deforestation measure from 1976, I predict that these changes will take place in the 16 years after paving is completed.

Note that these predictions are for changes in *levels* of deforestation. To make my predictions comparable to the predictions of the four teams of researchers, I need to convert these results to changes in rates of deforestation. Upon doing this, I find that the rate of net deforestation should increase by 17.5 percent,[49] while the rate of gross deforestation should increase by 19.7 percent.[50]

It would only be fair to point out that the percentages I cite using my methodology answer the question, "If the *Avança Brasil* roads had been paved in 1976, how much higher would the levels and rates of net and gross deforestation have been in 1992?" This is the same type of question answered in IPAM et al., though they would have chosen the base year as somewhere between 1957 and 1972, based on their assertion in their Table 1 that the frontier age was 20 to 35 years old along the roads of interest, and their data were from 1992. It is not as clear what the Laurance et al. estimates reflect, but it appears that they answer the question "what would the deforestation levels be if the roads in question had been paved 15 to 25 years earlier?"[51]

---

[49] Based on TRFIC's report of net deforestation of 286,714 square kilometers in 1992, scaled up by the ratio of INPE's 1992 gross deforestation (440,000 square kilometers) to TRFIC's 1992 gross deforestation (275,995 square kilometers), giving 286,714 square kilometers. Deforestation in 1976, extrapolated from INPE's 1978 deforestation, was computed to be 118,000 square kilometers.

[50] Based on INPE's 1992 gross deforestation level of 440,000 square kilometers.

be if the roads in question had been paved 15 to 25 years earlier?"[51] Andersen et al. actually use their model to predict the future, which they are better able to do than any of the other teams because their regression predicts deforestation rates, which is made possible by their use of panel data.

Compared to the question that the Laurance-IPAM teams and I seem to have answered, it is a much more difficult thing to predict the future. The reason that it is difficult for Laurance-IPAM is because, as my work in Chapter 4 showed, at some point, rates of deforestation begin slowing as a settled area ages. By now, many of the roads that are proposed to be paved under *Avança Brasil* have been in place for 30 or more years. It is probable that annual changes in deforestation are occurring slowly, if at all, along major portions of these roads.

In my case, I have established a baseline date of 1976 based on the location of pre-1976 *antropísmo*, and have distance measures based on that variable in Table 27. But for the "inside *antropísmo*" variable in that table, it is not clear how long it took for such an infilling to take place. If we forget for a moment that this was estimated with a tobit and treat the parameter as if it had been estimated with OLS, the implication is that inside areas that had been affected by human settlement, 36 percent will be cleared. The question we are left

---

[51] Endnote 16 in their supplemental article implies this.

with is how much had been cleared in 1976 in these areas designated as disturbed by human settlement? Our concern about future deforestation differs depending upon whether the answer is 10 percent or 36 percent. The analysis is less confusing for the distance categories outside of the original *antropísmo,* because we may assume that in 1976, the level of clearing in these areas was 0. My predictions for deforestation in 1992, based on the assumption that the roads to be paved by *Avança Brasil* were paved in 1976 did not need to deal with changing the values of *antropísmo,* but if I were to project from 1992 into the future, I would have to make a decision about how to treat deforestation that occurred before 1992. One sensible possibility would be to treat everything ever deforested by 1992 as new *antropísmo,* which would imply that my estimates for new deforestation by 2008 would be much larger.

An even more expedient estimate would be to take the values from Table 28 that did not use *antropísmo* in the regression. This would be a crude predictor into the future, and would result in my estimates ranging between 23.4 and 25.3 percent.

My estimates and those of Laurance et al. and IPAM et al. are partial equilibrium estimates. That is, we have been treating the Brazilian Amazon as if it were a region that was closed and independent of the rest of the country. We have failed to consider how Amazonian deforestation fits into the issue of Brazil as a whole. Almost all writings on the economic history of Brazil point out how shifts in Brazilian agriculture and opportunities for unskilled labor led to a massive in-migration to rural areas (see, for example, Faminow or Perz). The annual growth rate of the rural population (natural increase and net in-migration combined, since I do not have data on in-migration alone) was 3.6 percent in the fifties and 2.8 percent in the seventies (while at the same time the rural population in the rest of Brazil was falling at a rate of 1.2 percent annually). Between 1991 and 1996, however, rural population in the Brazilian Amazon fell at a rate of 1.4 percent per year, and between 1996 and 2000, fell at a rate of 2.2 percent per year (both of the latter two rates indicating faster drops in the rural population than experienced in the rest of Brazil for the same time periods). Clearly paving might change some incentives and cause a new wave of in-migration, but the surplus farmers from outside the Amazon that were available in the earlier mass in-migrations are just not there any more, as indicated by declining rural populations through all of Brazil.

Cattaneo's approach overcomes my objections to looking at the Amazon in isolation from the rest of Brazil. However, I am skeptical about his prediction that under *Avança Brasil*, agricultural production would increase in the Amazon, but net Brazilian production would be mostly unchanged. I am doubtful of such a small response in net agricultural production for Brazil, in light of the fact that the roads were designed to get products to export markets instead of domestic markets, and therefore prices in non-Amazon areas for agricultural products should not be very adversely affected.

Testing the effect of several policy prescriptions, Cattaneo finds that a logging tax would not lead to a decrease in the deforestation rate, but would be devastating to loggers. A logging tax would shift resources from logging into agricultural clearing. On the other hand, a R$50 per hectare deforestation tax could cut back deforestation by 9,000 square kilometers per year (which I believe means a 45 percent reduction in deforestation). He does not distinguish reduction in deforestation from its source; that is, whether the reduction comes from fewer new settlements, or from less deforestation on currently settled farms. However, his results would contradict my finding in Chapter 4 if he said most of the reduction came from the latter, since I found very little pricing response on rates of deforestation—though I did find that deforestation *levels* respond to price incentives. Cattaneo notes that under a deforestation tax, logging would only be minimally affected.

As an alternative policy, Cattaneo considers a subsidy on non-timber products, proposed by Vosti, Witcover, and Carpentier. He finds that a subsidy of R$240 per hectare of non-timber forest products could reduce deforestation by 30 percent. It seems to me that this policy might have perverse effects, in that it might encourage new settlers to move to the Amazon, because of increased incentives to have a farm there, especially during the early years.

It also seems that if farmers are cash constrained, this would provide money to them in the early years of settlement with which to deforest faster!

Recommendations to Policymakers

Noting the wide range of deforestation predictions presented in this chapter, I come away with the implicit message to the policymaker that there is a high degree of uncertainty as to what would really happen if all of the road paving proposed under *Avança Brasil* is completed. Even before I did my own analysis, I cited estimates of well-known researchers making predictions that ranged from a decrease in deforestation of 16.1 percent—made by Andersen et al.—to the IPAM et al. estimate in their worst case scenario of an increase in deforestation of 48.5 percent! It is interesting that my best estimate lies almost exactly in the middle of the two extremes, with an increase of 19.7 percent in the rate of gross deforestation.

Both Laurance et al. and IPAM et al. advised abandoning *Avança Brasil*, for the most part. But if the predicted economic benefits are high, perhaps proceeding cautiously would be better advice. In such a risky environment, where the penalty for choosing to pave the *Avança Brasil* roads might be high, it would be prudent for policymakers to try to shield the forest from the possibly disastrous consequences of the development project by passing proactive policies to curb rampant deforestation that might arise.

Which policies would be best? In this study—especially in the work using satellite data in Chapter 4—I argued how protected areas seemed to work well in deterring most deforestation. Adding new national parks and other forms of protected areas along the roads to be paved would be a solid step in the right direction.

While Cattaneo's predicted impact of a deforestation tax appears encouraging, one wonders what administrative capacity a deforestation tax would require. IBAMA already monitors deforestation and enforces the current laws using satellite data. But out of necessity, this enforcement focuses on larger patches of deforestation. It seems to me that attempting to tax every hectare of deforestation might be too difficult. On the other hand, if the tax were only applied to farms of a given size—500 hectares, say—then the task might be more manageable. If the enforcement were limited to farms of 500 hectares and above, Chapter 1 tabulations tell us that 48.7 percent of the land area of farms in the forest biome would be supervised, and this represents only 0.6 percent of the farms in the forest biome. Since such a large portion of on-farm deforestation comes from these mega-farms, perhaps the deforestation tax on large properties has potential to curb a significant amount of deforestation that might arise under *Avança Brasil.*

The suggestion to limit enforcement to large properties was so that the capacity of IBAMA and other agencies enforcing environmental laws would not be overwhelmed. An alternative policy which would also limit the scope of their task would be to put the deforestation taxes only on deforestation in *municípios* located along the roads to be paved by *Avança Brasil*. This would also solve one problem with limiting enforcement to farms greater than a certain size: in-migrants and others establishing new farms might simply choose to establish farms smaller than the minimum bound of such a policy.

In the case of uncertainty about the impact of development projects—and perhaps to make them more politically feasible, considering that Amazonian deforestation is an important concern to many people inside and outside Brazil—the government could institute rigid policies, such as declaring no new settlement along newly paved roads, or a steep deforestation tax for new clearing in such places. Then, after the roads are in place, relax the regulations along a portion, and see what new deforestation results. This relaxation of regulations can be done incrementally, both temporally and spatially. If loosening the regulations leads to high levels of new settlement, then maybe some of those who fear massive destruction of forest will be proved correct—and no further loosening of restrictions need be done. But if there is only a small response to the loosening of restrictions, then the optimists will be proved right, and more loosening of restrictions can be done. For a review of a much broader range of policy options, see Kaimowitz, Byron, and Sunderlin (1998).

Chapter 6: Summary and Conclusion

A Brief Summary of the Preceding Chapters

Instead of presenting a detailed summary of the points in the preceding chapters, I decided to begin this final chapter with a brief review of the preceding chapters, followed by a more in-depth review of the main themes in this dissertation. The ultimate goal of this dissertation was to estimate the impact of the road paving proposed in *Avança Brasil*. To reach that end it was necessary to examine some key issues in the study of deforestation, and provide some indicators to help us interpret the results. Sometimes what you learn in the journey is as important as the answer to the original question.

Chapter 1 began with a brief description of the importance of the Brazilian Amazon in global perspective. I then suggested that the Legal Amazon is really better thought of as two separate ecozones: the forest and the *cerrado*. Next came a review of satellite and survey data useful for studying deforestation in the Brazilian Amazon. Then I presented simple calculations and cross-tabulations which addressed two issues of importance throughout this study: the issue of forest regrowth, and the issue of "seed" deforestation, also possibly thought of as a baseline level of deforestation.

Chapter 2 reviewed the explanatory variables used throughout the rest of the study, then presented a simple dynamic model of land clearing on farms that formed the underlying framework for the regressions of Chapters 2 through 5, and especially the analysis done in Chapter 4. Chapter 2 concluded with several regressions using agricultural census data. These regressions explored the effects on farm area cleared of roads, protected areas, “seed” deforestation, and agroclimatic suitability.

Chapter 3 repeated the main regression from Chapter 2, using a spatial disaggregation technique I have been developing. While none of the results of Chapter 2 were reversed, the technique allowed us to map with finer detail the likely distribution of farm activity inside the cross-sectional units. Unfortunately, the efficiency gains of the technique were counteracted by the loss of efficiency in having to resort to Feasible Weighted Nonlinear Least Squares on the censored data, instead of the efficient maximum likelihood procedure. However, the groundwork was laid for future applications on non-censored data, where efficiency gains can be realized.

Chapter 4 used satellite data for 1996 to 1999 for portions of Mato Grosso and Rondônia, to gain a truly dynamic perspective of deforestation. The main analysis in the chapter regressed the rate of deforestation on farmgate prices for beef and milk, level of deforestation, agroclimatic variables, and protected areas. Deforestation rates on land already settled were not related to price levels; but new settlements were highly responsive to price levels.

Chapter 5 reviewed three well-known predictions and a fourth less well-known prediction for the effects of road paving as planned by *Avança Brasil* on deforestation. I pointed out some potential weaknesses in each of the models, particularly the Andersen et al. model which predicted a decrease in deforestation with road paving. After estimating the "obvious" model based on the agricultural census data—and finding that it predicted a very small increase in deforestation—I estimated additional models, trying to explain the reason for the difference between my model and those of two of the other research teams (Laurance et al. and IPAM et al.). From these additional regressions, I learned many things, which I will review in the following section.

Points to Take Away from This Study

The Legal Amazon: Two Very Different Biomes

Chapter 1 showed most clearly that agricultural and demographic dynamics are very different, depending on whether one is considering the forest biome or the *cerrado* biome of the Legal Amazon. In terms of levels of clearing and principal uses of agricultural land, Table 1 outlines the differences. *Municípios* in the forest biome in 1996 had approximately one-sixth of their land claimed for farms, while *municípios* in the *cerrado* biome had more than half of their land claimed for farms. Furthermore, farms in the forest biome have more than half of their land in native forest, while farms in the *cerrado* had less than 30 percent in forest. Livestock operations are 50 percent more intense in the forest biome, as measured by stocking density on pasture, or by value of livestock production per unit of pasture.

Table 3 shows some of the differences between biomes in terms of changes in land use between 1985 and 1996. Both have less farms in the latter period, but farms in the forest states have less total farm area. Both biomes, however, show a trend toward more cleared area (and therefore less forest); reducing natural pasture in favor of planted pasture; and reducing land in short- and long-term fallow. Table 3 also shows that cattle are expanding at a rate of 7.6 percent per year in forest states, and at a slower rate of 5.1 percent in *cerrado* states.

Table 4 shows the difference between the two biomes in farming operations. Cattle appear to be equally important in the two biomes in terms of gross value of production, yet as already noted, the *cerrado* favors more land-extensive cattle systems. Annual crops provide almost 60 percent of the agricultural revenue in the *cerrado*, and that is dominated by soybeans, which provide more value than all of the other annual crops combined. Annual crops are important in the forest biome, but not nearly as much as in the *cerrado*. Annual crops in the forest biome are also dominated by one crop: manioc, which provides more than half of all value of annual crops. The forest biome has significant portions of its total agricultural value coming from perennial crops (11.5 percent) and from extracted forest products (12.1 percent). This compares with 2.0 percent and 1.0 percent in the *cerrado* portion of the Legal Amazon.

Table 8 shows recent demographic trends in the Legal Amazon. While we see that both biomes share the trends toward moderate overall growth, declining rural populations, and increasing urban populations, the rural population decline is proceeding at a much faster rate in the *cerrado* portion, while the urban growth is proceeding at a much faster rate in the forest portion. The two preceding facts imply that the forest portion is growing faster—by around 60 percent—than the *cerrado* portion.

Surprisingly, the analysis in Chapter 4 showed no significant differences between deforestation in the forest and *cerrado* biomes. This is in part because the geographic area in Chapter 4 is a small subset of the entire Amazon. But it is also likely showing that it does not matter whether the trees are in the forest biome or the *cerrado* biome: ranchers and farmers treat them the same.

Stated a different way, this says that one of the most significant differences between the forest biome and the *cerrado* biome is that the former naturally has a higher proportion of dense forest than the latter, and farmers and ranchers naturally prefer less trees, because clearing land is costly. The other difference between the two biomes is that the mean rainfall in the forest biome is significantly higher than in the *cerrado* biome, which differentiates the agricultural technologies and crop choices that are feasible.

Population and cattle trends

While highlighting some of the differences between the forest and *cerrado* biomes in the previous subsection, I addressed issues related to demographics and cattle. While risking being repetitive, this section will also address those two topics, though this time not emphasizing the differences between biomes.

In terms of the human population, the main point to take away from this study is the trends over the past decade toward rural de-population and urban growth. While the urban growth rate in the Amazon has been higher than urban growth in the rest of Brazil since the 1960s, the rural depopulation in the Amazon is a new phenomenon.

Until the population census of 1991, rural Amazonia had a positive population growth. From the 1950 through 1991, each census showed the rural population in Amazonian states growing at a faster rate than in non-Amazonian states. In fact, non-Amazonian states exhibited a negative rural population growth rate beginning in the 1970s. However, as first revealed in the 1996 population count and confirmed in the 2000 census, after 1991 rural Amazonia not only experienced its first negative population growth rates, but these rates were even lower (i.e., depopulating faster) than the negative rates in non-Amazonian states. At the very least, these changes in Brazilian population demographics indicate that the patterns that accompanied the high deforestation rates in the 1970s and 1980s are no longer in operation in the 1990s to present.

The other point to take away from this study is that cattle dominate agricultural land use. Table 3 shows that pasture represented 67 percent of land converted for agricultural purposes on active farms in 1985, and 77 percent in 1996, and the proportion of planted pasture in land converted for agricultural purposes on active farms grew from 29 percent to 50 percent in the same time period. The cattle population grew at a rate of 6.0 percent per year between the two agricultural censuses, while pasture area expanded at a rate of 1.8 percent per year. Even when we acknowledge that the area of planted pastures grew at a rate of 5.2 percent per year, we still note that using either pasture growth figure leads us to find an intensification in pasture use.

Even though the soybean area has expanded rapidly in the *cerrado* zone of the Legal Amazon, crop area in 1996 was still less than one-ninth of pasture area, and was less than the land devoted to crops in 1985.

Effect of farmgate prices on various agents of deforestation

Because so much of cleared farm land is in pasture, policies which make cattle operations more costly seem like they should have a large impact on the rate of deforestation. Road construction and paving ought to have the opposite effect, because they essentially raise farmgate prices by making it less costly to get agricultural products to market. However, the analysis in Chapter 4 showed that there is a mixed effect of output price changes on forest clearing. In the analysis of Chapter 4, we considered two types of agents that are of interest to policymakers seeking to reduce on-farm clearing: the farm household which has already established residence, and the household which is thinking of moving to a new farmlot in the Amazon. We saw during the three-year period studied in Chapter 4, the former group was not likely to have their rates of deforestation easily influenced by price incentives. However, Brondízio et al., in their long panel micro-study which studied deforestation from 1970 to 1996 near Altamira, found that the timing and speed of deforestation on settled farms when measured over longer periods of time seem to be impacted by economic conditions. Turning to the second agent of deforestation, Chapter 4 shows that the rate of new settlement is strongly influenced by profitability signals, which policies can help modify. Brondízio et al. seem to confirm this as they noted that the rate of influx of settlers in their study area is highest in times of high economic stability and potential.

The observation about the two types of farm clearing agents partly answers the question raised by von Braun in the Foreword to Cattaneo’s study, and by Perz. They both observe that deforestation rates in the Amazon have not dropped much in recent years, despite many of the tax-breaks for homesteading and deforesting in the Amazon having been removed. The reason for the low response could be explained by the findings of Chapter 4: there is an inertia in deforestation that comes with those who have already settled land, and are deforesting that land at a steady rate. This is probably not the only factor causing the rate to hold steady despite net farm abandonment between 1985 and 1996, but it does explain quite a bit.

I discussed a number of policy options that pertained to prices throughout this study, particularly in Chapter 5. I argued for policymakers and researchers to consider whether the government realistically has the capacity to enforce some otherwise sensible policy prescriptions, such as a deforestation tax. In that case, I suggested limiting enforcement to either large farms which are quantitatively few but account for a large proportion of on-farm clearing; or enforcing the tax in specifically defined geographic regions, most sensibly in *municípios* located along the roads to be paved by *Avança Brasil*.

Another option might be a beef tax. Just as with a deforestation tax, it may not slow deforestation on already established farms, but it would perhaps limit the number of new farms established. Furthermore, it would probably be an easier tax to collect, because it could potentially be collected at slaughterhouses, and the number of slaughterhouses is small relative to the number of ranches. However, the effects of a beef tax need to be studied more carefully before considering it a viable option. Questions which need to be answered include how both demand and supply curves might shift in response to such a tax? On the demand side, consumers might potentially shift to other meats, and what they might be, and the resulting impact on nutrition and the environment need to be considered. On the supply side, the impact on ranchers needs to be more carefully weighed.

Data source and deforestation source

In Chapter 5 we discovered how much difference the type of data makes in determining the effect of road paving on deforestation. When using the agricultural census, we found that the effect of paving roads was quite moderate. However, when we used satellite data, we found the effect to be substantial. This led into a discussion about the source of deforestation. As we already noted in Chapter 1, when we sum gross deforestation reported on farms in 1996 in the forest biome, we only account for about half of the gross deforestation as measured by INPE. Some of the difference is because there is some forest in the *cerrado* biome, but the agricultural census does not tell us how much of the land that has been converted for agricultural use was once forest. Both Chapter 1 and Chapter 5 tried to account for a large portion of the difference. Some of it can be explained by abandoned farms. Another part can be explained by urban and nearby residential areas, which are noted as deforested in satellite data, but which do not appear in agricultural census data.

The analysis in Chapter 5 confirmed that another part of the difference between the two data sources is that satellite data pick up off-farm logging, which we saw in Chapter 1 is probably greater than on-farm logging by a factor of 10. Finally, I cited Cochrane et al. as arguing that forests burned by wildfires are sometimes incorrectly categorized as deforested during satellite image interpretation.

Forest regrowth

Chapter 1 discusses the difference between gross deforestation and net deforestation, which is forest regrowth. The question, “Which type of deforestation should we be most interested in?” is an important one. Those most concerned about carbon sequestration would probably favor net deforestation, because carbon growth rates may be similar in secondary succession and primary forest. The answer is not quite as clear for those interested in biodiversity. On the one hand, forest regrowth provides a good habitat for many animal species. But if the period of clearing on the land was too long or too geographically large, some animal species might have been adversely effected, as might some plant species. For those concerned with wildfires (see Nepstad et al. 1999 or Cochrane et al. 1999), gross deforestation is probably the more important measure, since forest regrowth is more flammable.

The main point that I want to summarize here is that there is a large gap between gross deforestation and net deforestation. According to my tabulations in Chapter 1 using TRFIC satellite data from 1992, roughly one-third of gross deforestation in 1992 was regrowing. Both INPE and IBAMA keep track of gross deforestation, which means that the gap between gross and net will be ever widening. Since INPE started keeping track of deforestation in 1978, some of the areas declared deforested then might have been regrowing for twenty-five years now, which means that from a satellite image, it would be indistinguishable from native forest and, even on the ground, it might be very difficult to tell apart.

"Seeds" and rates of deforestation

I presented a simple dynamic model in Chapter 2. The phase-plane diagram clearly showed that along the optimal path, deforestation rates would slow as deforestation levels increased. Chapter 4 confirmed that at least after an initial period of establishing a farm, rates of deforestation decline with increases in deforestation levels, and that this effect is quantitatively large. Rates of deforestation computed in Chapter 4 show that clearing of agricultural land takes place over a long period of time, perhaps up to 60 years, until steady state levels of deforestation are reached. The incremental nature of deforestation was also noted in the micro-study done by Brondízio et al. It is important for researchers studying deforestation to keep in mind that the agents are not likely to have reached equilibrium, and that failure to control for the impact of the level of deforestation on the rate of deforestation will likely bias the estimates.

In Chapter 5, I presented a similar critique of Laurance et al. and IPAM et al. for not properly establishing a baseline level of deforestation at some point in time before computing the level of deforestation. There are actually two dynamic issues involved, both treated in Chapter 4. First of all, as just argued, rates of deforestation slow as levels of deforestation rise. Second, even after controlling for prices and agroclimatic suitability, new farm establishments are more likely to locate close to already established farms than far from them. This was seen in the probit model of new settlement in Chapter 4, where the rings near the deforestation frontier were more likely to be settled, and at a decreasing probability as we moved further away from the frontier where deforestation meets undisturbed forest. We also saw this in all of the regressions of Chapters 2, 3, and 5, where nearness to areas showing signs of human disturbance (*antropísmo*) in 1976 had a large quantitative and statistically significant effect on levels of deforestation.

The spreading out from a deforestation “seed” or “core” may simply reflect accessibility issues when settling new land (i.e., it is easier to settle near someone who has established or assured that they have access to their land, then it is to cut a path far away from others). Or it may reflect social factors or the increased likelihood of tapping into public services when living near others. Regardless of the cause, the policymaker trying to channel or control the location, rate, and level of deforestation would be wise to try to limit the beginning of “seeds” or “cores” of deforestation.

Throughout the study, I have tried to note the effectiveness of protected areas and indigenous areas in preventing deforestation. While the ones in the forest biome tend to be located in areas that have low levels of expected deforestation, the ones in the *cerrado* biome are in high-risk areas. The *cerrado*-only regression in Chapter 2 showed their effectiveness, as did the regressions in Chapter 4. Schneider et al. argue in their logging study for more national forests to control deforestation.

Agroclimatic suitability matters

Analysis of deforestation levels or rates that do not take into account agroclimatic suitability (i.e., rainfall, soils, and native vegetation) are likely to bias the results. The importance of measures of agroclimatic suitability, particularly rainfall, was a key point in Chomitz and Thomas, and was confirmed in Chapters 2 through 5 of this study.

How much deforestation will result from paving roads under Avança Brasil?

Table 26 summarizes the predictions of the other authors. There is a large variation in the predictions, not just between teams of researchers, but also in the wide range of predictions some of the researchers put forward. In Chapter 5, I suggested that since there was such a wide diversity in predictions, and since the environmental cost is high if some of the pessimistic scenarios prove right, policymakers should choose to err on the side of caution by instituting deforestation limiting policies in conjunction with implementing *Avança Brasil.*

Table 28 presents the full range of my predictions based on different assumptions. Focusing on the scenarios with the most realistic assumptions, I believe that the increase in the level of deforestation on farms is likely to be only 2.5 percent more than it would be apart from paving, and possibly in the very long run, 7.9 percent in the entire Amazon, or 11.5 percent, when restricting the analysis to the forest biome. For on-farm and off-farm deforestation together, and restricting our attention to the forest biome, I predict an increase of 10.3 percent in net deforestation, and an increase of 14.4 percent in gross deforestation, with respective increases in rates of deforestation over a 16 year period being 17.5 percent and 19.7 percent. In the very long run, I predict that the level of deforestation might increase as much as 25.3 percent.

These values seem to be in very good agreement with those of Laurance et al., even though our methodologies are quite different. They are also in reasonable agreement with those of IPAM et al. and Cattaneo. They are different from those of Andersen et al., who predicted counterintuitively that there would be a decrease in deforestation.

Finally, it is important to emphasize that the predictions are based not on what will happen if *Avança Brasil* is implemented, but on what will happen if *Avança Brasil* is implemented *without any accompanying policy change to limit deforestation*. Since parts of *Avança Brasil* have already been implemented, the time for implementing proactive policies is now, before new deforestation "seeds" begin, and the problem arises concerning what to do about settlers who moved in before new legislation could be enacted.

GLOSSARY OF BRAZILIAN PORTUGUESE WORDS

antropísmo	Used in this study to mean an area that shows signs (in satellite images) of human settlement or disturbance.
Avança Brasil	$40 billion dollar program to develop infrastructure in the Amazon region, starting in 2000.
babaçu	A type of tree that is the natural secondary vegetation to spring up in large portions of Maranhão. It is a palm tree that produces a kernel from which an oil is extracted, which is useful for soap and foods. The nut shells are used for charcoal for cooking (Porro).
cerrado	Savanna.
município	County.
varzea	Seasonally flooded areas.

REFERENCES

Alston, Lee J., Gary D. Libecap, and Bernardo Mueller. 1999. *Titles, Conflict, and Land Use: The Development of Property Rights and Land Reform on the Brazilian Amazon Frontier*. Ann Arbor, MI: University of Michigan.

Alves, Diógenes S. 1999. “An Analysis of the Geographical Patterns of Deforestation in Brazilian Amazônia in the 1991-1996 Period.” Presented at the 48th Annual Conference on Patterns and Processes of Land Use and Forest Change in the Amazon, University of Florida, March 23 to 26.

Alves, Diógenes S. 2002. “An Analysis of the Geographical Patterns of Deforestation in Brazilian Amazonia in the 1991-1996 Period”, Ch. 3 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 95-106.

Andersen, Lykke, Clive W. J. Granger, Eustaquio J. Reis, Diana Weinhold, and Sven Wunder. 2002. *The Dynamics of Deforestation and Economic Growth in the Brazilian Amazon*. Cambridge, UK: Cambridge University.

Barrett, Scott. 1991. “Optimal Soil Conservation and the Reform of Agricultural Pricing Policies”, *Journal of Development Economics* 36:167-187.

Bigman, David and Uwe Deichmann. 2000. “Geographical Targeting: A Review of Different Methods and Approaches”, Ch. 1 in *Geographical Targeting for Poverty Alleviation: Methodology and Applications*, edited by David Bigman and Hippolyte Fofack. Washington: The World Bank, pp 43-73.

Bliss, Norman B. 1992. *Vegetation Map of Brazil*. Sioux Falls, SD: EROS Data Center, USGS. Digitized version of *Mapa de Vegetacao do Brasil*, produced by IBGE, 1988, at 1:5,000,000. Available from http://grid.cr.usgs.gov/clearinghouse/datalist.html.

BSRSI (Basic Science and Remote Sensing Initiative). n.d. *Digital Data on Deforestation in the Brazilian Amazon in 1992*. Ann Arbor, MI: Michigan State University. Available at http://www.bsrsi.msu.edu/trfic/.

Brondízio, Eduardo S., Stephen D. McCracken, Emilio F. Moran, Andrea D. Siqueira, Donald R. Nelson, and Carlos Rodriguez-Pedraza. 2002. "The Colonist Footprint: Toward a Conceptual Framework of Land Use and Deforestation Trajectories among Small Farmers in the Amazonian Frontier", Ch. 5 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 133-161.

Browder, John O. and Brian J. Godfrey. 1997. *Rainforest Cities: Urbanization, Development, and Globalization of the Brazilian Amazon*. New York: Columbia University.

Caldas, Marcellus, Robert Walker, and Stephen Perz. 2002. "Small Producer Deforestation in the Brazilian Amazon: Integrating Household Stucture and Economic Circumstance in Behavioral Explanation", CID Working Paper No. 96 (October), Harvard University, Boston, Massachusetts.

Carvalho, Georgia, Ana Cristina Barros, Paulo Moutinho, and Daniel Nepstad. 2001. "Sensitive Development Could Protect Amazonia Instead of Destroying It" (letter to editor), *Nature* 409(January 11):131.

Cattaneo, Andrea. 2002. "Balancing Agricultural Development and Deforestation in the Brazilian Amazon", IFPRI Research Report 129. Washington, D.C.: IFPRI.

Chomitz, Kenneth M. and David Gray. 1996. "Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize", *World Bank Economic Review* 10(September):487-512.

Chomitz, Kenneth M. and Timothy S. Thomas. 2001. "Geographic Patterns of Land Use Intensity in the Brazilian Amazon", World Bank Policy Research Working Paper no. 2687 (October), World Bank, Washington, D.C.

Chomitz, Kenneth M. and Timothy S. Thomas. Forthcoming. "Determinants of Land Use In Amazônia: A Fine-Scale Spatial Analysis", *American Journal of Agricultural Economics*.

Cochrane, Mark A., Ane Alencar, Mark D. Schulze, Carlos M. Souza, Jr., Daniel C. Nepstad, Paul Lefebvre, and Eric A. Davidson. 1999. "Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests", *Science* 284(June 11):1832-1835.

Cochrane, Mark A., Ane Alencar, Mark D. Schulze, Carlos M. Souza, Jr., Paul Lefebvre, and Daniel C. Nepstad. 2002. "Investigating Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests", Ch. 10 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 285-298.

Costa, Fabiano G. 2000. *Avaliação Do Potencial De Expansão Da Soja Na Amazônia Legal: Uma Aplicação do Modelo de Von Thünen*. Master's Thesis in *Economia Aplicada* at *Universidade de São Paulo.*

Davis, Benjamin. 2001. "Is it Possible to Avoid a Lemon? Reflections on Choosing a Poverty Mapping Method", paper presented at World Bank Seminar on Poverty Mapping, October 25.

Deichmann, Uwe. 1999. "Geographic Aspects of Inequality and Poverty", World Bank's Website on Inequality, Poverty, and Socioeconomic Performance. Available at http://www.worldbank.org/poverty/inequal/index.htm.

Demombynes, Gabriel, Chris Elbers, Jenny Lanjouw, Peter Lanjouw, Johan Mistiaen, and Berk Ozler. 2002. "Producing a Better Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", WIDER Discussion Paper No. 2002/39, The United Nations, Helsinki.

EOS Amazon Project. 1999a. *Digital Map of Amazon Basin Boundary*. Seattle: University of Washington.

EOS Amazon Project. 1999b. *Digital Map of Amazon Mean Precipitation*. Seattle: University of Washington.

ESRI. 1993. *Digital Chart of the World.* Digitized data, based on Defense Mapping Agency maps at 1:1,000,000. Available at many sites, including http://www.maproom.psu.edu/dcw.

ESRI. 1992. *Arc World.* Digitized data which is shipped with ESRI's *Arc View* software, and much of which is available at http://www.esri.com.

Faminow, Merle D. 1998. *Cattle, Deforestation, and Development in the Amazon: An Economic, Agronomic and Environmental Perspective.* New York: CAB.

FAO (Food and Agriculture Organization). 1970-1978. *Soil Map of the World*, scale 1:5,000,000, volumes I- X. Paris: United Nations Educational, Scientific, and Cultural Organization.

Henninger, Norbert and Mathilde Snel. 2002. *Where are the Poor? Experiences with the Development and Use of Poverty Maps.* Washington: World Resources Institute and UNEP/Grid-Arendal. Available at http://www.povertymap.net/pub.htm.

Holt, D., D. G. Steel, M. Tranmer, and N. Wrigley. "Aggregation and Ecological Effects in Geographically Based Data", *Geographical Analysis* 28(3, July):60-77.

IBAMA (*Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis*). 2000. Digital and Tabular Data on Deforestation in the Amazon. Brasília: IBAMA. Tabular data available at http://www2.ibama.gov.br/desmata/

IBGE (*Instituto Brasileiro de Geografia e Estatística*). 1996. Census tract-level Arc Info Export Files received on CDROM.

IBGE. 1997a. *Diagnostico Ambiental da Amazônia Legal*. Rio de Janeiro: IBGE. Digitized maps on CDROM.

IBGE. 1997b. *Mapa Da Série Brasil - Geográfico* 1:5,000,000. Digitized maps on CDROM.

IBGE. 1998a. *Censo Agropecuário 1995-1996: número 2 (Rondônia), número 3 (Acre, Roraima e Amapá), número 4 (Amazonas), número 5 (Pará), número 6 (Tocantins), número 7 (Maranhão), and número 24 (Mato Grosso).*

IBGE. 1998b. *Censo Agropecuário 1995-1996*. Electronic data received by email.

IBGE. 2002a. *Produção Agrícola Municipal* (PAM). Available at http://www.sidra.ibge.gov.br/.

IBGE. 2002b. *Pesquisa Pequária Municipal* (PPM). Available at http://www.sidra.ibge.gov.br/.

IBGE. 2002c. *Produção Extrativa Vegetal* (PEV). Available at http://www.sidra.ibge.gov.br/.

IMAZON. 2000a. *Digital Vegetation Map*. Unpublished, IMAZON, Belém, Brazil.

IMAZON. 2000b. *Digital Roads Map*. Unpublished, IMAZON, Belém, Brazil.

IMAZON. 2000c. *Digital Rivers Map*. Unpublished, IMAZON, Belém, Brazil.

IMAZON. 2000d. *Digital Map of Slaughterhouses in the Amazon*. Unpublished, IMAZON, Belém, Brazil.

IMAZON. 2000e. *Digital Map of Dairies in the Amazon*. Unpublished, IMAZON, Belém, Brazil.

IMAZON. 2000f. "Survey of Prices in Slaughterhouses in the Amazon", mimeo.

IMAZON. 2000g. "Survey of Prices in Dairies in the Amazon", mimeo.

IMAZON. 2000h. *Digital Map of Timbermills in the Amazon*. Unpublished, IMAZON, Belém, Brazil.

INPE (*Instituto Nacional de Pesquisas Espaciais*). 2002. *Monitoring of the Brazilian Amazonian Forest by Satellite, 2000 - 2001*. São José dos Campos, São Paolo, Brazil: INPE. Downloaded at http://www.inpe.br/Informacoes_Eventos/.

International Travel Maps. 1999. *Kevin Healey's Travel Map of Brasil* Scale 1:4,000,000. Vancouver: International Travel Maps.

IPAM (*Instituto de Pesquisa Ambiental da Amazônia*), ISA (*Instituto Socio-Ambiental*), and WHRC (The Woods Hole Research Center). 2000. "Report on the Scenarios Project", June 1. Downloaded May 19, 2003 from http://www.whrc.org/science/tropfor/setresearch.htm.

IPCC (Intergovernmental Panel on Climate Change). 2001. *Climate Change 2001: Impacts, Adaptation, and Vulnerability (Summary for Policymakers).* Geneva: IPCC.

Jalan, Jyotsna and Martin Ravallion. 1997. "Spatial Poverty Traps?", World Bank Policy Research Working Paper No. 1862, Washington, D. C.

Kaimowitz, David. 2002. "Amazon Deforestation Revisited", *Latin American Research Review* 37(2):221-235.

Kaimowitz, David and Arild Angelsen. 2001. "Will Livestock Intensification Help Save Latin America's Forests?", mimeo.

Kamien, Morton I. and Nancy L. Schwartz. 1991. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2nd ed. Amsterdam: North-Holland.

Kelejian, Harry H. and Ingmar R. Prucha. 2001. “On the Asymptotic Distribution of the Moran I Test Statistic with Application”, *Journal of Econometrics* 104:219-257.

Klein, Lawrence R. 1946. “Remarks on the Theory of Aggregation”, *Econometrica* 14(4, October):303-312.

Krautkraemer, Jeffrey A. 1994. “Population Growth, Soil Fertility, and Agricultural Intensification”, *Journal of Development Economics* 44:403-428.

Laurance, William F., Mark A. Cochrane, Scott Bergen, Philip M. Fearnside, Patricia Delamônica, Christopher Barber, Sammya D’Angelo, and Tito Fernandes. 2001a. “The Future of the Brazilian Amazon”, *Science* 291(January 19):438-440.

Laurance, William F., Philip M. Fearnside, Mark A. Cochrane, Sammya D’Angelo, Scott Bergen, and Patricia Delamônica. 2001b. “The Future of the Brazilian Amazon” (letter to the editor), *Science* 292(June 1):1652-1654.

Lele, Uma, Virgilio Viana, Adalberto Verissimo, Stephen Vosti, Karin Perkins, and Syed Arif Husain. 2000. *Brazil. Forests in the Balance: Challenges of Conservation with Development*, Evaluation Country Case Study Series. Washington: World Bank.

LeSage, James. 1998. “Spatial Econometrics”, mimeo. Downloaded July 2003 from http://www.spatial-econometrics.com/html/wbook.pdf.

LeSage, James. 1999a. “The Theory and Practice of Spatial Econometrics”, mimeo. Downloaded July 2003 from http://www.spatial-econometrics.com/html/sbook.pdf.

LeSage, James. 1999b. “Applied Econometrics Using MATLAB”, mimeo. Downloaded July 2003 from http://www.spatial-econometrics.com/html/mbook.pdf.

Liu, Dawning S., Louis R. Iverson, and Sandra Brown. 1993. “Rates and Patterns of Deforestation in the Philippines: Application of Geographic Information System Analysis.” *Forest Ecology and Management* 57:1–16.

Lütjohann, Harry. 1974. *Linear Aggregation in Linear Regression*. Stockholm: Institute of Statistics, University of Stockholm.

May, Kenneth. 1946. “The Aggregation Problem for a One-Industry Model”, *Econometrica* 14(4, October):285-298.

McCracken, Stephen D., Andrea D. Siqueira, Emilio F. Moran, and Eduardo S., Brondízio. 2002. "Land Use Patterns on an Agricultural Frontier in Brazil: Insights and Examples from a Demographic Perspective", Ch. 6 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 162-192.

Mertens, Benoît and Eric F. Lambin. 2000. "Land-Cover-Change Trajectories in Southern Cameroon", *Annals of the Association of American Geographers* 90(3):467-494.

Mertens, B., R. Poccard-Chapuis, M.-G. Piketty, A.-E. Lacques, and A. Venturieri. 2002. "Crossing Spatial Analyses and Livestock Economics to Understand Deforestation Processes in the Brazilian Amazon: The Case of São Félix do Xingú in South Pará", *Agricultural Economics* 27:269-294.

Ministerio da Agricultura, Secretaria de Planejamento e Coordenacao da Presidencia da Republica, and *Fundacao Instituto Brasileiro de Geografia e Estatistica* (IBGE). 1998. *Mapa de Vegetação do Brasil.* Digitized at the U.S. Geological Survey's EROS Data Center, Sioux Falls, South Dakota. Available at ftp://grid2.cr.usgs.gov/pub/data/datafile_descriptions/brveg11d.html

Mittermeier, Russell A.., Patricio Robles Gil, and Cristina Goettsch Mittermeier. 1998. *Megadiversity: Earth's Biologically Wealthiest Nations.* Mexico City: CEMEX.

National Space Development Agency (NASDA), Earth Observation Research Center (EORC), Japan. n.d. *Digital Data of the Amazon from JERS-1.* Pasadena, CA: Jet Propulsion Laboratory. Available at http://www.eorc.nasda.go.jp/amazon/.

Nepstad, Daniel C., Adalberto Veríssimo, Ane Alencar, Carlos Nobre, Eirivelthon Lima, Paul Lefebvre, Peter Schlesinger, Christopher Potter, Paulo Moutinho, Elsa Mendoza, Mark Cochrane, and Vanessa Brooks. 1999. "Large-scale Impoverishment of Amazonian Forests by Logging and Fire", *Nature* 398(April 8):505-508.

Nepstad, D., J. P. Capobainco, A. C. Barros, G. Carvalho, P. Moutinho, U. Lopes, and P. Lefebvre. 2001. "*Avança Brasil*: The Environmental Costs for Amazonia", draft, *Instituto de Pesquisa Ambiental da Amazônia* (IPAM).

Nerlove, Marc. 1958. *The Dynamics of Supply : Estimation of Farmers' Response to Price.* Baltimore : The Johns Hopkins Press.

NIMA (National Imagery and Mapping Agency). 2000. *GEOnet Names Server (GNS) Database*. An electronic database of place names and locations in geographic coordinates, available at http://164.214.2.59/gns/bin/vali-data_query.

Openshaw, S., 1984a. *The Modifiable Areal Unit Problem*. Norwich, U. K.: Geo Books.

Openshaw, S., 1984b. "Ecological Fallacies and the Analysis of Areal Census Data", *Environment and Planning A* 16:17-31.

Paelinck, Jean H. P. 2000. "On Aggregation in Spatial Econometric Modelling", *Journal of Geographical Systems* 2:157-165.

Perz, Stephen G. 2002. "Population Growth and Net Migration in the Brazilian Legal Amazon, 1970-1996", Ch. 4 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 95-106

Perz, Stephen G. and Robert T. Walker. 2002. "Household Life Cycles and Secondary Forest Cover Among Small Farm Colonists in the Amazon," *World Development* 30(6):1009-1027.

Pfaff, Alexander S. P. 1997. "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data", World Bank Policy Research Working Paper no. 1772 (March), World Bank, Washington, D.C.

Pfaff, Alexander S. P. 1999. "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data", *Journal of Environmental Economics and Management* 37(January):26-43.

Porro, Roberto. 2002. "Land Use, Cattle Ranching, and the Concentration of Landownership in Maranhão, Brazil", Ch. 12 in *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 315-337.

Pu, Shou Shan. 1946. "A Note on Macroeconomics", *Econometrica* 14(4, October):299-302.

Ravallion, Martin and Quentin Wodon. 1997. "Poor Areas, or Only Poor People?", World Bank Policy Research Working Paper No. 1862, Washington, D. C.

Redwood, John, III. 2002. "World Bank Approaches to the Brazilian Amazon: The Bumpy Road Toward Sustainable Development", Latin America and Caribbean Region Sustainable Development Working Paper No. 13 (November), The World Bank, Washington, D. C.

Reis, Eustáquio and Rolando Guzmán. 1994. "An Econometric Model of Amazon Deforestation," Ch. 12 in *The Causes of Tropical Deforestation: The Economic and Statistical Analysis of Factors Giving Rise to the Loss of the Tropical Forests*, ed. by Katrina Brown and David Pearce. London: UCL Press, pp 172-191.

Reis, Eustáquio and Sérgio Margulis. 1991. "Options for Slowing Amazon Jungle Clearing," Ch. 9 in *Global Warming: Economic Policy Responses*, ed. by Rudiger Dornbusch and James M. Poterba. Cambridge, MA: MIT, pp 335-380.

Ruud, Paul A. 2000. *An Introduction to Classical Econometric Theory*. New York: Oxford University Press.

Schneider, Robert R., Eugênio Arima, Adalberto Veríssimo, Paulo Barreto and Carlos Souza Júnior. 2001. *Sustainable Amazon: Limitations and Opportunities for Rural Development*. Washington: World Bank.

Silveira, José Paulo. 2001. "Development of the Brazilian Amazon" (letter to the editor), *Science* 292(June 1):1651-1652.

Skole, D. and C. Tucker. 1993. "Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988", *Science* 44(May 5):314-322.

Sombroek, Wim. 2001. "Spatial and Temporal Patterns of Amazon Rainfall: Consequences for the Planning of Agricultural Occupation and the Protection of Primary Forests", *Ambio* 30(November):388-396.

Steel, D. G. and D. Holt. 1996. "Analysing and Adjusting Aggregation Effects: The Ecological Fallacy Revisited", *International Statistical Review* 64(1):39-60.

Stoker, Thomas M. 1993. "Empirical Approaches to the Problem of Aggregation Over Individuals", *Journal of Economic Literature* 31(4, December):1827-1874.

Theil, Henri. 1955. *Linear Aggregation of Economic Relations*. Amsterdam: North Holland.

Thomas, Timothy S., Kenneth M. Chomitz, and Eugénio Arima. 2001. “Price and Profitability of Cattle Products and Amazonian Deforestation”, mimeo.

TRFIC (Tropical Rain Forest Information Center). n.d. *Digital Data on Deforestation in the Brazilian Amazon in 1992.* Ann Arbor, MI: Michigan State University. Available at http://www.bsrsi.msu.edu/trfic/.

Uhl, Christopher, Paulo Barreto, Adalberto Veríssimo, Edson Vidal, Paulo Amaral, Ana Cristina Barros, Carlos Souza Jr., Jennifer Johns, and Jeffrey Gerwing. 1997. “Natural Resource Management in the Brazilian Amazon: An Integrated Research Approach”, in *BioScience* 47(3, March):160-167.

UNEP-WCMC (United Nations Economic Programme – World Conservation Monitoring Centre). 1999. *Digital Map of Ecozones.* Cambridge, U.K.: WCMC.

USDA (U. S. Department of Agriculture). 1998. *Soil Stress Map of the World*, digitized map of the primary soil limiting factors. Electronic data received by email from the World Soil Resources (USDA).

USGS (U. S. Geological Survey). 1999. *Digital South American River Basins.* Sioux Falls, SD: EROS Data Center. Download at http://edcdaac.usgs.gov/gtopo30/hydro/sa_basins.html.

van Garderen, Kees Jan, Kevin Lee, and M. Hashem Pesaran. 2000. “Cross-sectional Aggregation of Non-linear Models”, *Journal of Econometrics* 95:285-331.

Veríssimo, Adalberto, Paulo Barreto, Ricardo Tarifa, and Christopher Uhl. 1995. “Extraction of a High-Value Natural Resource in Amazonia: The Case of Mahogany”, in *Forest Ecology and Management* 72:39-60.

Veríssimo, Adalberto, Carlos Souza Júnior, Steve Stone, and Christopher Uhl. 1998. “Zoning of Timber Extraction in the Brazilian Amazon”, in *Conservation Biology* 12(1, February):128-136.

Vosti, S.A., C.L. Carpentier, and J. Witcover. 2002. “Agricultural Intensification by Smallholders in the Western Brazilian Amazon”, IFPRI Research Report 130. Washington, D.C.: IFPRI.

Walker, Robert, Emilio Moran and Luc Anselin. 2000. “Deforestation and Cattle Ranching in the Brazilian Amazon: External Capital and Household Processes”, *World Development* 28(4):683-699.

Warnken, Philip F. 1999. *The Development and Growth of the Soybean Industry in Brazil*. Ames, IA: Iowa State University.

Wong, David. 1996. “Aggregation Effects in Geo-Referenced Data”, Ch. 5 in *Practical Handbook of Spatial Statistics*, edited by Sandra Lach Arlinghaus. Boca Raton, FL: CRC Press, pp 83-106.

Wood, Charles H. 2002. “Land Use and Deforestation in the Amazon”, Introduction to *Deforestation and Land Use in the Amazon*, edited by C. H. Wood and R. Porro. Gainesville, FL: University Press of Florida, pp 95-106

Wood, Charles H. and Roberto Porro. 2002. *Deforestation and Land Use in the Amazon*. Gainesville: University Press of Florida.