Economic development and the environment: three essays

Stefanie Sieber

Dissertation submitted for the degree of

Doctor of Philosophy in Economics at

The London School of Economics and Political Science

February 2011

UMI Number: U615342

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI
Dissertation Publishing

UMI U615342
Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

F
9459

LSE
Library
British Library of Political
and Economic Science

1275778

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others. The second chapter draws on work that was carried out jointly with equal share by Robin Burgess (LSE), Matthew Hansen (South Dakota State University), Benjamin Olken (Massachusetts Institute of Technology), and me.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

Stefanie Sieber

Abstract

The main question that motivates my PhD thesis is how economic activity in developing countries is influenced by and, in turn, affects the environment. Since these interactions can take many forms, I investigate this issue from three different angles, which necessitates both the usage of novel remote-sensing-based datasets and the development of a new theoretical framework.

Firstly, the environment can have a direct impact on economic development, the most obvious example being natural disasters like cyclones. As the incidence and intensity of these events will increase with climate change, it is crucial to estimate their short- and long-run costs and the behavioural response of producers to these large and mostly uninsured aggregate shocks. I have, thus, created a new digital database of cyclone exposure for India to estimate how farmers smooth income in the aftermath of these events.

The causality can also run the other way, as economic agents disrupt the environment. A case in point is deforestation, which is analysed in the co-authored second chapter of my thesis. In particular, we use satellite data to study how political decentralisation has affected district-level logging rates in Indonesia. Possible mechanisms include local election cycles, the move from monopoly to oligopoly or the need to raise revenue in the absence of other natural resources.

Finally, the third chapter assesses to what extent the environment can create preconditions for socioeconomic interaction. More specifically, I analyse how the introduction of heterogeneous space into the standard urban Muth-Mills model generates a residential equilibrium where the formal and informal housing markets coexist. This new setup is then used to evaluate the usual policy prescriptions for slums and demonstrates that new insights can be gained by adding the spatial component.

This thesis, therefore, explores possible links between the environment and economic development and illustrates the advantages of using methods and data sources from other disciplines.

For my grandfather, Peter Bahles senior, without whom I could have never studied in the United Kingdom

Acknowledgements

There are many whose help was invaluable for me in completing my PhD thesis. I would like to thank my primary supervisor Robin Burgess for his inspiring ideas and generous advice. My secondary supervisor Oriana Bandiera deserves my thanks for prolific discussions and the great amount of time she has devoted to my thesis, especially in its final stages. Benjamin Olken from MIT was very helpful in getting me started on my first chapter and has been a great source of learning throughout the work for our joint project, which constitutes the second chapter of this thesis.

I am also deeply indebted to the LSE for awarding me the full Tibor Scitovsky Scholarship and a one-year Economic Research Studentship that allowed me to conduct my research. Financial support from the DFID Research Program Consortium on Improving Institutions for Pro-Poor Growth in Africa and South-Asia is also gratefully acknowledged.

I would also like to thank Matthew Hansen from the Geographic Information Science Center of Excellence at SDSU for sharing his remote sensing data and expertise. The second chapter of this thesis would not have been possible without his help. Mark Broich has also provided kind and patient explanations. Furthermore, Raymond Guiteras has been extremely helpful in sharing the World Bank India Agricultural and Climate Dataset with me.

Moreover, I would like to thank my wonderful proof-readers, Christine Knoop and Jennie Condell, without whose input my work could not have become what it is. Any linguistic infelicities that may have found their way into the text despite their best efforts are, of course, my responsibility.

There are a number of others who contributed to this thesis less directly, but no less significantly. I would like to thank Dave Donaldson, Esther Duflo, Cynthia Kinnan, Guy Michaels, Henry Overman, Rohini Pande, Olmo Silva, Daniel Sturm, Anthony Venables, and seminar participants at the LSE for helpful comments and discussions. This work has also benefitted from my stay at MIT and the Jameel Abdul Latif Poverty Action Lab from September 2008 to August 2009.

Many thanks are also due to the incredibly helpful IT team at STICERD, Joe Joannes and Nic Warner, whose continued support and assistance made the empirical work for the second chapter possible. Mahvish Shaukat and Nivedhitha Subramanian have provided excellent research assistance for the same project.

In addition, I would like thank my friends for their support, assistance, and patience throughout my doctoral studies. Above all, I would like to thank my best friend Christine Knoop for her encouragement and thought-provoking impulses in the many phases of this thesis. I could not have completed it without her.

Last but not least, I would like to thank my dear family, my parents Katja and Michael Sieber, my brother Oliver and his wife Virginie, for their unwavering support that began long before and continues well beyond my PhD studies.

Contents

List of Figures

List of Tables

Preface

The issue of economic activity in developing countries being influenced by and, in turn, affecting the environment has come to the forefront of public debate. The discourse has been fuelled by recent research on anthropogenic climate change, which has shown that livelihoods all over the world will be threatened if environmental concerns continue to be ignored (IPCC, 2007a). It is therefore crucial to understand the linkages between economic activity and the environment if we are to draw wider conclusions on ways to decelerate climate change and mitigate its consequences. My PhD thesis explores three of these many conceivable linkages.

Until recently, research in this field has been restricted due to limited availability of relevant data. Now remote sensing technologies have opened up new data sources to economists and can be used to measure natural phenomena objectively in real time. This kind of data is particularly valuable for analyses of developing countries, for which the record is often incomplete. Moreover, satellite images provide a comprehensive picture of the earth's surface and therefore are able to capture the impact of both legal and illegal actions on the environment.

The first two chapters of this thesis make use of these novel data sources to analyse two possible linkages between the environment and economic activity. The environment can have a direct impact on economic development, the most straightforward example being the destruction of agricultural output caused by natural disasters. Developing countries are particularly vulnerable to these shocks, because their economies depend heavily on the primary sector (FAO, 2009). Given that the frequency and intensity of extreme weather events are predicted to increase with climate change (Christensen et al., 2007), it becomes all the more important to estimate the associated short- and long-run costs.

Chapter 1 studies the impact of one particular natural disaster, namely tropical cyclones. I construct a new digital database of cyclone exposure for India, which I use to estimate the effect of these shocks on the primary sector. In addition, the disaggregate nature of the agricultural data allows me to take the analysis further and study how producers recover from these large and mostly uninsured aggregate income shocks. More specifically, I exploit the interaction between the random timing of the cyclone hit and the district-level growing seasons to identify possible income smoothing mechanisms. I find that producers smooth income both by increasing their production in the growing season immediately following the cyclone shock and by changing their crop mix towards more resilient and nutritious crops.

However, the causality can also run the other way, as economic agents disrupt and destroy the environment. A case in point is deforestation, which is one of the main drivers of anthropogenic climate change (Denman et al., 2007). It is thus important to understand how forests can be preserved effectively, as sustainable forest management could potentially play a major role in future mitigation strategies (Stern, 2006; Nabuurs et al., 2007). This policy issue is again especially relevant for developing countries, whose weak institutions can lead to excessive illegal logging (CIFOR, 2004), which will have to be curbed if emissions are to be reduced.

Indonesia provides an insightful case study, as it is experiencing some of the most rapid deforestation rates in the world (Hansen et al., 2009). Furthermore, it has concomitantly experienced fundamental institutional changes since the overthrow of the dictator Soeharto. The co-authored second chapter of my thesis thus uses a new satellite-based forest cover dataset to investigate how decentralisation has affected deforestation. We analyse a series of possible political economy mechanisms and find that the move from a monopoly of forest extraction rights to an oligopoly has significantly increased logging. Moreover, fiscal decentralisation has incentivised local bureaucrats and politicians to raise revenue through logging. This is particularly the case prior to local elections or if districts do not have access to other lucrative natural resources, such as oil and gas.

Finally, given that climate change will destroy rural livelihoods, especially in developing countries (IPCC, 2007a), the pressure on urban areas is set to increase over the coming decades. Since city governments in these regions have limited resources to finance public housing schemes (Simha, 2006; WB, 2002), this will drive the new arrivals into the informal housing market. A variety of policy prescriptions involving slum-upgrading and titling programmes have been offered to tackle this looming housing crisis (UN-Habitat (2003), chapter 7). However, the consequences of these policies can only be fully understood if the analysis is integrated into a spatial context; an aspect which has been ignored so far.

The third and final chapter of my thesis addresses this issue by studying how the environment can create preconditions for socioeconomic interaction. More specifically, Chapter 3 provides a new theoretical framework that introduces heterogeneous space into the standard urban model of the monocentric city. This generates a residential equilibrium where the formal and informal housing markets coexist; a feature that is assumed to be exogenously given in other models (Jiminez, 1984; Brueckner and Selod, 2008). This new setup is then used to evaluate the usual policy prescriptions for slums and demonstrates that new insights can be gained by adding the spatial component. In particular, the results suggest that slum upgrading is economically inefficient and that resettlement programmes could increase social welfare.

The three chapters of this thesis thus contribute to the wider debate on how our economic decisions both are influenced by and impact upon the environment, and, by extension, anthropogenic climate change. The first chapter provides some evidence on possible coping strategies in the face of extreme weather events that disrupt the usual insurance arrangements. Given that global warming will increase the likelihood of such aggregate shocks, the focus of both academics and policy makers should shift more towards understanding income smoothing mechanisms. In contrast, the second chapter studies one of the root causes of anthropogenic climate change: deforestation.

The findings suggest that logging rates can only be lowered with a comprehensive policy approach that also addresses institutional weakness. Lastly, the third chapter deals with the consequences of actual displacement through climate change, which will transmit the impact of global warming from the rural to the urban sphere. As these population pressures will significantly increase in the coming decades, early action is warranted to create the capacities to accommodate this influx.

Given the complexity of the problem, future research should devote more effort to understanding the multifaceted relationship between the environment and economic activity. Only then will we be able to devise possible adaptation and mitigation strategies that will help lower the costs of anthropogenic climate change both for current and future generations.

1 Income Smoothing and Cyclone Damage in India

1.1 Introduction

Tropical cyclones are some of the most devastating natural disasters on the earth (Dilley et al., 2005). They are intense whirls in the atmosphere that reach wind speeds of up to 250km/h – the average speed of a TGV (Taylor, 2007). These storms usually affect large areas, as a fully-grown cyclone can have a diameter of 1,000km and move 300 to 500km in a day (IMD, 2009). In addition, they are often accompanied by storm surges and torrential rains, which inundate coastal regions and flood rivers. The resulting destruction can be apocalyptic. For instance, the 1999 Orissa cyclone – the deadliest Indian storm since 1971 – affected 17,993 villages in 14 districts (BAPS, 1999). Official figures put the death toll at 9,893 (though it is believed that as many as 15,000 people died) and estimated a cost of approximately US$4.5 billion. Moreover, around 5 million farmers lost their livelihoods, since 406,000 farm animals perished in the storm and 1.7 million hectares of crops were destroyed (USNO, 1999).

The destructive potential of tropical cyclones is bound to increase even further. In fact, climate change models predict a higher frequency and intensity of these storms over the next few decades (Henderson-Sellers et al., 1998; Kuntson and Tuleya, 2004; Emanuel, 2005; Christensen et al., 2007). Developing countries in particular will suffer disproportionately from this development. Empirical evidence has shown that the impact of natural disasters is larger for poorer, more unequal, and less democratic societies (Khan, 2005; Strömberg, 2007; Toya and Skidmore, 2007). Furthermore, their economies heavily depend on the primary sector, which is affected most by cyclones. Recent figures for India, for example, show that 56.35% of the economically active population work in agriculture whose share in GDP is 16.6% (comparable figures for the US are 1.75% and 1.1% respectively) (FAO, 2009).

Given the importance of the issue, this paper's main aim is to estimate the cost of tropical cyclones on the primary sector in developing countries. In addition, I investigate possible mechanisms through which farmers try to cope with these aggregate income shocks. I focus on one country – India – that is struck by a tropical cyclone every two to eight years on average. The great advantage of this case study is that the India Meteorological Department has recorded the positions of all cyclones since 1891. I can thus construct a novel dataset of cyclone exposure, which estimates the disaster impact for a long time horizon in a consistent and objective manner (O'Keefe and Westgate, 1976; Albala-Bertrand, 1993; Yang, 2008). This dataset is then combined with the India Agriculture and Climate Dataset of the World Bank – the most comprehensive district panel of agricultural data currently available for India. Specifically, it reports crop-level production information annually from 1956 to 1987.

Using these two datasets, I can estimate the cost of tropical cyclones on a much finer geographical scale than has been feasible previously.[1] I find that a one standard deviation increase in cyclone exposure lowers output by 7.75%. Rather surprisingly, I also find that the area planted increases contemporaneously by 3.02%; an effect which persists for three years until recovery is achieved. These seemingly contradictory results can be explained by the following observation: most districts in India have two and some even three growing seasons (GOI, 1967). That is to say, producers will be able to adjust their production later in the *same* year, if, for instance, they were struck by a cyclone in the spring. The annual data thus only captures the *net* effect of the natural disaster, where the destruction caused by the cyclone is counterbalanced by the subsequent behavioural response of the producers.

The change in behaviour in itself is warranted, because most farmers are not protected against cyclone shocks. Firstly, the Indian government only provides short-term disaster relief that does not involve any aid for recovery or development (Parasuraman and Unnikrishnan, eds, 2000). Secondly, most producers lack formal insurance against weather risk, which is rarely taken out in both developed and developing countries (Kunreuther and Pauly, 2004; Cole et al., 2009). Finally, since the cyclone affects the entire social network, access to informal finance will be disrupted (Besley, 1995a). Consequently, farmers will be unable to smooth consumption in the aftermath of the natural disaster and will have to adjust their production behaviour instead.

Due to the detailed nature of my dataset, I can identify two possible coping strategies that could explain the observed increase in the area planted. Firstly, the analysis exploits the interaction between the random timing of the cyclone shock and the district-level growing seasons. More specifically, I supplement the World Bank's dataset with district-level growing cycle information to distinguish between cyclone events that happen before and after the sowing season of each individual crop. I can then test whether production increases after the previous harvest has been destroyed. The results show a very consistent picture – a crop is destroyed if the cyclone hits prior to the harvest, but more is planted if it strikes before the sowing season. Farmers thus attempt to smooth income by substituting intertemporally across growing seasons. Note that the effectiveness of this coping strategy will be dampened, since markets are not well-integrated in India (Topalova, 2004; Duflo and Pande, 2007; Guiteras, 2007). This is confirmed in the data, as local farm harvest prices decline with the expansion in production.

The second income smoothing mechanism involves a change in the crop mix. In particular, I separate the food crops into two main groups. The first category comprises coarse cereals, such as bajra and jowar. These crops are resilient, cheap, and highly nutritious. The second group includes more risky and expensive calories like rice and wheat. If farmers need to smooth income and (most probably) nutrition

[1]The only sub-national estimate of natural disasters to date is a state-level analysis of Pugatch and Yang (2008) for Mexico.

after the cyclone shock, we would expect to see a switch towards coarse cereals. The evidence suggests that this is indeed the case – the area planted with coarse cereals expands by 4.22% resulting in a net increase in output and a decline in prices of 11.47% and 2.53% respectively. In contrast, the other food crops are largely destroyed by the cyclone (output drops by 13.93%) and are not replanted within the same year. This finding can also explain the overall increase in the area planted, since coarse cereals can be grown on marginal land whereas other food crops cannot (Sawhney and Daji, 1961). Farmers thus seem to forgo higher income from low calorie crops to meet basic nutritional needs and smooth income after the cyclone shock.

These results document the large and adverse impact of tropical cyclones on the primary sector in India. I find that producers attempt to counter the destruction – by planting more crops in the growing season directly following the cyclone shock or by changing their crop mix towards cheap calories. However, their response is not sufficient. The overall contemporaneous impact on output is still negative and sizeable. Moreover, income smoothing continues for another three years until the damage is offset completely. Disaster relief should thus not only provide short-term supportive measures, but also assistance to rebuild livelihoods. Furthermore, it is important to note that these coping strategies are specific to India. The impact of cyclonic storms will be larger for developing countries with only one growing season.

The remainder of the paper is organized as follows. In the next section, I discuss the related literature and place the contribution of this paper in a broader context. Section 1.3 explains how the cyclone exposure dataset is constructed and discusses the key features of the World Bank dataset needed for the empirical analysis. In Section 1.4, I motivate the identification strategy and describe its empirical implementation. Section 1.5.1 presents the estimates of the net effect of cyclone exposure on the primary sector. In Section 1.5.2, I study how the interaction between the district-level growing seasons and the timing of the cyclone shock affects planting decisions. Section 1.5.3 tests if the crop mix changes in response to tropical cyclones. Section 1.6 concludes.

1.2 Background

This paper is most closely related to the nascent literature on natural disasters. Research in this field has so far largely focused on estimating the impact of these aggregate shocks at the macro-level (Toya and Skidmore, 2002; Strömberg, 2007; Cuaresma et al., 2008). A prominent cross-country analysis has been carried out by Khan (2005), who shows that good institutions and democracies can play a crucial role in lowering the death toll from natural disasters. Similarly, Anbarci et al. (2005) provide evidence that richer and more equal societies have fewer fatalities.

However, even if this relationship is true across countries, recent micro-evidence paints a more nuanced picture. For instance, Pugatch and Yang (2008) show nonlinearities in the relationship between mortality from hurricanes and the level of development. They argue that reconstruction efforts in more affluent regions have induced emigration from poorer to richer states, thus mitigating the impact for both. In addition, more aid is usually provided by more accountable local governments (Besley and Burgess, 2002; Cole et al., 2008) or for natural disasters with wider media coverage (Eisensee and Strömberg, 2007). The impact also differs across income groups, since families with lower levels of education (Bluedorn and Cascio, 2005) and income (Foster, 1995) suffer most.

The present study contributes to this literature by estimating the effect of tropical cyclones on the primary sector in India. By exploiting the finesse of my cyclone dataset, I am able to estimate this effect for each district – a level of disaggregation which has not been achieved before. The reasons for focusing on the primary sector are as follows. Firstly, agricultural production is most exposed to cyclonic storms. The strong winds, and accompanying rains and floods will primarily destroy crops and kill livestock (IMD, 2009). Secondly, more than half of the economically active population in India (and in most other developing countries) works in agriculture today (FAO, 2009). Therefore, the full impact of this natural disaster can only be understood if the primary sector is taken into account.

Moreover, to the best of my knowledge this paper attempts to disentangle possible coping strategies for the first time. So far, little empirical evidence exists on how producers recover from aggregate income shocks like natural disasters. For instance, the related literature on large temporary shocks merely studies their long-run impact. A case in point is the analysis of Miguel and Roland (2006). They show that the extensive bombing of Vietnam has not permanently changed poverty rates, consumption patterns, infrastructure, literacy, and population density. However, they cannot identify the mechanisms through which recovery has been accomplished. Similarly, Davis and Weinstein (2002) provide evidence that the dropping of the atomic bombs on Hiroshima and Nagasaki has not had a long-run impact on the distribution of economic activity in Japan. But again evidence on possible coping strategies is scarce.[2]

Given that natural disasters repeatedly strike the same region (Dilley et al., 2005), we would expect that forward-looking economic agents will insure themselves against these shocks. Nonetheless, empirical evidence has shown that households underinsure against weather risk due to incomplete information about the event probability (Kunreuther and Pauly, 2004), credit constraints, and lack of trust (Cole et al., 2009). In the case of India, government assistance is not sufficient to offset the negative consequences of natural disasters either. That is, relief operations largely focus on short-term supportive measures and not on recovery and development (Parasuraman and Unnikrishnan, eds, 2000; HPC, 2001).[3]

[2]In contrast to these temporary one-off shocks, permanent changes in the productive capacity of a region should lead to a behavioural response. For example, Hornbeck (2008) finds that migration patterns and population growth changed substantially in response to the extensive topsoil erosion in the American Dustbowl.

[3]This attitude has shifted only recently after the super-cyclone in Orissa in 1999 and the major

Alternatively, producers could insure themselves more indirectly through borrowing and saving. The academic literature has provided considerable evidence on consumption smoothing through both formal and informal credit arrangements in developing countries (Deaton, 1992; Udry, 1994; Townsend, 1994; Grimard, 1997; Fafchamps and Lund, 1997; Dubois, 2000). Yet, the findings suggest that this is far from complete. Moreover, informal insurance networks are particularly vulnerable to geographically co-moving shocks like tropical cyclones. In particular, they usually rely on local information networks and community sanctions to overcome information asymmetries, enforcement problems and transaction costs (Besley, 1995a). Since the cyclone will affect the entire social network, producers will be severely limited in their ability to smooth consumption through borrowing.

In such an environment of repeated and largely uninsured aggregate shocks, producers will have to fall back on income smoothing mechanisms (Morduch, 1995). This literature is much more limited, as most of the research to date has focused on household-level borrowing and saving behaviour. The existing evidence suggests that farmers attempt to smooth income by selling their assets or livestock or by increasing their labour supply or off-farm employment (Rosenzweig and Wolpin, 1993; Dercon, 2000; Jayachandran, 2006). However, this will be more difficult after a cyclone shock, which will have damaged assets, killed domestic animals, and severely disrupted the local economy. Moreover, the effectiveness of these coping strategies will be reduced if markets are not well integrated. This is the case in India (Topalova, 2004; Duflo and Pande, 2007; Guiteras, 2007), where we would expect general equilibrium effects to drive down local prices and wages.

This paper studies two alternative income smoothing mechanisms: namely, changes in the crop mix and an increase in agricultural production in the subsequent growing season. Naturally, these responses will also have local price effects, which will mitigate their impact. However, they might be easier to implement after a natural disaster. To identify these coping strategies, I take advantage of the detailed nature of my dataset. More specifically, I can determine whether a cyclone shock has occurred before or after the district-level sowing season of a given crop. The behavioural response will be identified by cyclone hits prior to the sowing season. In contrast, shocks after planting but before the harvest will capture the cyclone's destruction. In addition, I am able to investigate whether producers adjust their crop mix by switching toward cheap calories.

1.3 Data

The main difficulty in estimating the cost of tropical cyclones is to obtain a consistent and objective measure of the damage caused. Section 1.3.1 outlines how I solve this

earthquake in Gujarat in 2001. Since then an initiative has been launched by the High Powered Committee on Disaster Management to develop a comprehensive approach to all natural disasters that involves not only disaster relief, but also mitigation and rehabilitation (HPC, 2001).

problem by using two complementary datasets to predict cyclone exposure. This procedure is explained and evaluated in detail in Section 1.3.2. To identify the impact of tropical cyclones on the primary sector, I also need to consider the timing of the cyclone shock and, most importantly, its interaction with the growing cycle of each crop. Therefore, I augment my main agricultural database – which will be discussed at length in Section 1.3.3 – with district-level sowing and harvesting times. Both datasets together then help formulate the identification strategy outlined in Section 1.4. They also prepare the ground for the empirical analysis, the results of which are presented in Section 1.5.

1.3.1 Cyclone data sources

To capture the impact of tropical cyclones, I use meteorological measurements. The key advantage of this type of data is its objectivity. That is, alternative data sources, such as damage assessments done by local or national governments, are known to exaggerate the disaster impact (O'Keefe and Westgate, 1976; Albala-Bertrand, 1993; Yang, 2008). Furthermore, developing countries might have underreported natural disasters in the past. Consequently, any increases in the frequency of events may be solely due to better reporting practises (Strömberg, 2007). On-the-ground assessment of the damage is equally problematic. For this approach, both pre- and post-disaster measurements are necessary, which are often incomplete and inaccurate. Meteorological data therefore presents a unique opportunity to measure the disaster impact objectively.

The raw cyclone data records the position of the centre of the storm every few hours. These coordinates make up the so-called cyclone track, which traces out the storm's movement across time and space. To estimate the actual extent of the cyclone from this data, a theoretical framework is required.[4] These models generally describe the wind and pressure fields of the storm. In addition, they calculate the changes in its windfall dynamics, as the cyclone moves across the ocean and hits land. The final outputs are the so-called wind speed buffers that are constructed around each cyclone track. These calculations require the cyclone track itself, measurements of the central pressure and wind speed, and a host of auxiliary parameters. Since wind speed measurements are based on satellite images, these models cannot be applied prior to the 1970s when this data first became available.[5] This is problematic, since the agricultural data used in this analysis only covers 1956 to 1987.[6]

[4]See Lovell (1990) for a comprehensive review. Frequently used models build on the work of Holland (1980), Holland (1997), De Maria et al. (1992), or Merrill (2000).

[5]E-mail correspondence with Pascal Peduzzi from the UNEP/DEWA/GRID-Europe on the 24.03.2009.

[6]Duflo and Pande (2007) have collected district-level data for the six major crops (bajra, jowar, maize, rice, sugar and wheat) up to 1999. However, their dataset is incomplete and lacks information on several major states, including Bihar, Orissa, and West Bengal. It also does not provide information on the minor crops that are crucial for my analysis. I therefore only work with data from the World Bank, which is the most comprehensive database on Indian agriculture currently available.

To solve this problem, I use two cyclone datasets that complement each other in terms of precision and data coverage. The first dataset is the so-called eAtlas of the India Meteorological Department (IMD), which is a digital database of all storms and depressions in the Bay of Bengal and the Arabian Sea from 1891 to 2008.[7] Since this dataset covers a long time horizon, only cyclone track information is available. However, the data is comparable across years, as the IMD has applied a consistent classification throughout (IMD, 2008). The second more sophisticated dataset has been designed by UNEP/DEWA/GRID-Europe PREVIEW (UNEP).[8] They employ a state-of-the-art GIS model based on Holland (1997) to calculate wind speed buffers for all cyclone events that occurred worldwide between 1975 and 2007.

Clearly, the first dataset has the great advantage of covering a long time period. In contrast, the second provides a much more precise estimate of cyclone exposure, albeit for a shorter time horizon. To take advantage of both datasets, I use the period of overlap to predict the percentage of the district's area affected by a cyclone for 1891 to 2008.[9] Note that this prediction – which will be explained in more detail in Section 1.3.2 below – only proxies for the extent of the cyclone, not the actual level of destruction. Yang (2008) has therefore argued that the cyclone variable should also be weighted by population. Unfortunately, this is not possible in this context, as the Indian Census only takes place every ten years. Interpolation of the intervening years is not desirable, because it fails to capture any declines in the population due to the cyclone itself.

1.3.2 Measuring cyclone exposure at the district level

To predict cyclone exposure, I first construct 16 district-level variables from the cyclone tracks of the eAtlas. These variables attempt to mimic the main features of the UNEP's wind speed buffers and will be used in the prediction[10]. To this end, both raw data sets are displayed in Figure 1.1. The map, for example, clearly shows that stronger cyclones are associated with both larger wind speed buffers and longer cyclone tracks. I thus proxy for the size of the wind speed buffers by summing the cyclone track length within a district. To capture possible non-linearities in this relationship, I square the track length variable and use it as an additional regressor in the prediction. The other 14 variables are constructed similarly and explained in full in Appendix A.3.

These 16 controls are subsequently regressed on the dependent variable constructed

[7]The data can be ordered at `http://www.imdchennai.gov.in/cyclone_eatlas.htm`. Appendix A.3 contains additional information on this database.

[8]The dataset can be downloaded from `http://preview.grid.unep.ch/index.php?preview=data&events=cyclones&lang=eng`. More information on the data is provided in Appendix A.2.

[9]I use the district boundaries of 1966 throughout the analysis. This makes the cyclone data compatible with the World Bank's India Agriculture and Climate dataset, which is my main source of data for the primary sector (see Section 1.3.3 for more detail).

[10]Note that the prediction does not change significantly if instead I use the first three principal components of these variables, which explain 82.95% of the total variation.

from the UNEP data. The latter measures the percentage of the district's area covered by a wind speed buffer for the period of overlap, 1977-2003[11]. The dependent variable is then predicted out of sample for 1891 to 2008. This yields the so-called *cyclone exposure* variable, which will be the main control of interest in the empirical analysis. Note that the standard errors will have to be bootstrapped in the regressions, since the *cyclone exposure* variable is a predicted regressor. Finally, I have recorded the date information for each of the 16 eAtlas controls. Therefore, I can determine when a cyclone hits a given district, which will be important for the identification strategy later on.

Nonetheless, there are two problems with using the eAtlas data for this prediction. Firstly, it codes storms with wind strength of 48 knots (equivalent to 88.9km/h) or more as cyclones, whereas the UNEP data uses a much higher threshold of 96 knots (177.8km/h). Since no wind speed measurements are available for the eAtlas, I cannot distinguish weaker from stronger storms. Therefore, all of the cyclone events will have to be included in the prediction. As a consequence, there is a risk of overpredicting exposure, especially at the low end of the distribution. I therefore construct a second *severe cyclone exposure* variable, which restricts the prediction to being greater or equal to 25% of the district's territory.[12]

To assess the quality of these predictions, I first compare the dependent variable with the *cyclone exposure* variable for the period of overlap. The correlation between these two variables is displayed in Figure 1.2a. It clearly shows that the eAtlas data often predicts damage, when none was recorded by the UNEP wind speed buffers. This leads to a Pearson's correlation coefficient of 0.29. The coefficient increases to 0.44, if the correlation is only calculated for strictly positive values of the dependent variable (Figure 1.2b), and to 0.46 for the *severe cyclone exposure* variable (Figure 1.2c). Similarly, Spearman's rank correlation coefficient increases from 0.13 to 0.45 to 0.47 as the sample is restricted further.

The improved fit can also be seen when comparing the distributions of the predicted variables with the dependent variable for the period of overlap. Figure 1.3a shows that the distribution of the UNEP data is twin-peaked with a large mass at very low and high levels of the percent area affected. Thus, it has a large standard error of 3.44 around its mean of 33.46%. The distribution of the predicted *cyclone exposure* variable is instead centred on its mean of 22.74% with a standard error of only 0.47 (Figure 1.3b). Yet, once the distribution is truncated for the *severe cyclone exposure* variable (Figure 1.3c), it more closely resembles the UNEP data – it is then skewed to the right and its mean and associated standard error almost double to 39.41% and 0.83 respectively.

These results seem to suggest that the *severe cyclone exposure* variable better fits

[11] For the period of overlap tropical cyclones only reach the Indian shore between 1977 and 2003.

[12] This threshold effectively restricts the distribution of the predicted variable to the top 30% (conditional on being positive). Any increases in this threshold lower the sample size significantly.

the UNEP data and should thus be used in the regression analysis. However, it is important to note that the truncation considerably limits the number of district-level cyclone events. For the final sample period ranging from 1956 to 1987, the number of district-level cyclone events drops from 1138 to 273 through the restriction. This will significantly lower the precision of the estimates in the empirical analysis. These two factors will have to be weighed against each other before making a final decision on which variable to use.

The second problem with the eAtlas data concerns measurement error. Firstly, I need to use a much simpler model than the UNEP team to estimate the percentage of the district's area affected. Moreover, the eAtlas cyclone tracks only roughly follow the movement of the UNEP wind speed buffers, which can be clearly seen in Figure 1.1. To address this issue, I construct a third variable, which solely relies on the eAtlas data. More specifically, I simply calculate the district's minimum distance to the nearest cyclone track. However, this *minimum distance* variable only provides a lower bound estimate of cyclone exposure, since its correlation with the UNEP wind speed buffers is only 0.12 in absolute value as opposed to 0.44 for the *cyclone exposure* variable. It will thus be used as a robustness check in the empirical analysis.

Table 1.1 displays the summary statistics for all three eAtlas-based cyclone variables for the sample period used in this analysis, 1956-1987. The data is reported for all observations in Panel A and for the respective cyclone samples in Panel B. The overall probability of being exposed to a cyclone is very low; on average only 2.37% of the district's area is exposed to the cyclones as measured by the *cyclone exposure* variable – a figure that drops to 1.05% for the *severe cyclone exposure* variable. Similarly, the *minimum distance* to the nearest cyclone is large on average at 775.40km. However, if a district is struck by a cyclone, exposure increases dramatically to 21.41% and 39.54% for the *cyclone exposure* and *severe cyclone exposure* variable respectively. The average *minimum distance* to the nearest cyclone track also drops significantly to 237.46km.

1.3.3 Data for the primary sector

The outcome variables for this analysis were retrieved from the World Bank's India Agricultural and Climate Dataset (WB).[13] This is a comprehensive district-level database for the primary sector that provides production information on an annual basis for 1956 to 1987 – note that a year is defined as an agricultural year, which roughly runs from April to the following March. The sample includes 271 districts within 13 states (which use the 1966 boundaries) covering almost 80% of the Indian territory.[14] I only lack information on two agriculturally important states, namely Kerala and Assam. Nonetheless, this is not a major concern for the current analysis, since both

[13]The data can be downloaded from `http://ipl.econ.duke.edu/dthomas/dev_data/index.html`. For more information on the dataset see Appendix A.4.

[14]The full list of states included in the analysis is provided in Appendix A.1.

states are not hit by a tropical cyclone during the sample period. Consequently, they would have only contributed to the estimation of the fixed effects.

The great advantage of this dataset is that it contains crop-level information on farm harvest prices, agricultural output, and area planted for each district. This data is available for 6 major and 14 minor crops, which cover a wide variety ranging from cereals to fibre crops.[15] It also reports price and quantity information for variable inputs, such as fertilizer, agricultural labourers, bullocks and tractors. Unfortunately, this data is only available at the district level. This makes it impossible to calculate crop-level profits or to analyse potential changes in the production technology in response to the shock. Moreover, the labour and capital data is collected only in decennial and quinquennial censuses respectively. Therefore, it fails to capture the disaster impact, making it unsuitable for the present study.

Instead, the empirical analysis focuses on the crop-level variables measuring the area planted, output and farm harvest prices. Note that I adjust prices with the agricultural GDP deflator for 1980. I also construct three additional outcome variables. First of all, I capture the value of production by calculating earnings as the product of prices and quantities. Secondly, to measure productivity, I compute yields that are equal to the ratio of earnings to the area planted. Finally, I proxy for the nutritious value of the crops by converting the output data into its calorie content. These calculations use weights obtained from MeDINDIA – a premier health portal in India.[16]

Nonetheless, more information is needed to identify the impact of tropical cyclones on the primary sector. The reason for this is that cyclone events occur throughout the agricultural year.[17] These shocks thus have a differential effect across crops depending on their position in the growing cycle. The situation is further complicated by the fact that most Indian districts have two and some even three growing seasons (GOI, 1967). Consequently, some crops within the *same* agricultural year are affected by the cyclone, whereas others are planted only after the shock. To disentangle these effects, the WB dataset is augmented with a district-level crop calendar (GOI, 1967).[18] This growing season information allows me to distinguish which crops in my sample were sown before or after a given cyclone shock.[19]

Section 1.4 now explains how this distinction helps identify the impact of tropical

[15]The six major crops are bajra, jowar, maize, rice, sugar, and wheat. The 14 minor crops are barley, cotton, gram, groundnut, jute, other pulses, potatoes, ragi, rapeseed and mustard, sesamum, soy, sunflowers, tobacco, and tur.

[16]The calorie information was downloaded on 22nd of July 2010 at `http://www.medindia.net/patients/foodcalories/index.asp`

[17]For the *cyclone exposure* variable, storms occur from February-December, whereas for the *severe cyclone exposure* variable, shocks occur from April-December.

[18]Appendix A.5 provides more information on the construction of this dataset.

[19]Note that the final dataset only includes crops for which sowing and harvesting times are reported in the crop calendar, or for which I could calculate the sowing and harvesting times from the state-level data. For details of these calculations refer to Appendix A.5. I have also dropped non-annual crops from the analysis, as cyclones in several years could impact on their production and it is difficult to disentangle these effects.

cyclones on agricultural production and crop mix choices.

1.4 Identification strategy

1.4.1 Motivation

The identification strategy is built on the following two observations about tropical cyclones. Firstly, the movement and intensity of each storm is solely determined by climatic and oceanic conditions (IMD, 2009). The cyclone shock is therefore exogenous to the primary sector allowing me to estimate its causal impact on agricultural output and production choices. Note that this effect is heterogeneous across districts due to the non-random movement of tropical cyclones. For example, in this dataset, 8.4% of all districts suffer one third of all cyclone shocks prior to 1956. We would thus expect producers in these high-risk regions to anticipate the event and adjust their production behaviour accordingly. However, since they constitute such a small fraction of the sample, the estimates are too imprecise to identify a differential effect. I therefore restrict myself to estimating the average cost and behavioural response.

Secondly, tropical cyclones form throughout the year and can affect the same district in different months.[20] As has been explained in Section 1.3.3, the impact of the shock thus differs across crops within the *same* agricultural year due to the presence of multiple growing seasons. Specifically, some crops are hit by the cyclone prior to their harvest. Their output is hence destroyed. However, there is a second group of crops that has not been sown yet. Given that cyclone shocks are largely uninsured, producers can thus respond in a variety of ways. For instance, they might intensify their input use in the subsequent growing season to raise yields. Alternatively, farmers could expand production by using land that was supposed to be fallow[21]. They might also plant more resilient crops such as coarse cereals that can flourish on marginal land (Sawhney and Daji, 1961). These crops have the added advantage of a high calorie content[22], which will help farmers smooth both income and nutrition in the aftermath of the natural disaster.

Given that producers can respond to the shock within the *same* year, a simple regression of cyclone exposure on agricultural output will not estimate the total level of destruction. Instead it will provide a measure of the *net* effect of a tropical cyclone. That is, it will capture both the damage caused *and* the subsequent behavioural response of the producers. These two effects counterbalance each other and it is not

[20]Cyclone shocks occur from February to December for the *cyclone exposure* and *minimum distance* variables and from April to December for the *severe cyclone exposure* variable. However, the prime cyclone season runs from May until November, with more than three quarters of all shocks occurring during that time period for all three cyclone measures.

[21]Summer and winter crops are often planted in rotation, which can include fallow periods. For example, the maximum production of barley can only be achieved if the land is not cropped in the preceding summer (Sawhney and Daji (1961), p. 116). The suitability of a given combination of crops depends on local climate, soil characteristics, and the availability of irrigation facilities.

[22]The average calorie content of coarse cereals is 343.6 calories per 100g, whereas the corresponding figure for all other food crops is 294.8 calories per 100g.

clear *a priori* which effect should outweigh the other. It is therefore an empirical question to determine if the change in behaviour is sufficient to offset the impact of the natural disaster. I can estimate this overall effect by regressing the cyclone exposure variable on the production data for the entire universe of crops.

I can even take the analysis one step further and break up the overall effect into its individual components. To do so, I exploit the interaction between the timing of the cyclone shock and the multiple growing seasons in my sample. For example, I distinguish between cyclones that occurred before and after the sowing time of the so-called winter crops – i.e., crops planted in the second half of the agricultural year. If these two cyclone variables are then regressed on the output of winter crops *only*, I can separately identify the cyclone impact and coping strategy. That is, if the district was hit by a cyclone at the beginning of the year, farmers should have smoothed income by expanding production and thus output in the second growing season. However, shocks that occurred towards the end of the year should have destroyed the standing winter crops with no room for recovery. The converse argument can be applied to the summer crops, where the destruction of the harvest in the previous year should have led to an increase in production in the next growing season.

Furthermore, I investigate if farmers change their crop mix towards coarse cereals in response to the shock. As has been argued above, producers might choose to plant more coarse cereals after a cyclone shock to ensure a steady income and cheap calories. However, by doing so farmers forgo higher income. That is, coarse cereals are generally of low economic value – on average, they fetch prices half the size of other food crops in my sample. Therefore, we would expect this change in the crop mix to be only temporarily. I can test for this coping strategy by splitting the data into coarse cereals[23] and all other food crops[24]. I then estimate the *net* effect of cyclone exposure for both groups separately and test for the persistence of the shock.

Section 1.4.2 now describes how this identification strategy is implemented.

1.4.2 Empirical Implementation

Firstly, I estimate the net effect of tropical cyclone on agricultural production with an OLS fixed effects regression, as follows:

$$y_{cdt} = \alpha + \beta cyclone_{dt} + \lambda_d + \mu_c + \gamma_t + \epsilon_{cdt}, \tag{1}$$

where y_{cdt} is either the area planted, output, farm harvest price, earnings, yields, or calorie content of crop c in district d in year t, and $cyclone_{dt}$ is one of the three district-level cyclone exposure variables discussed in Section 1.3.2. The coefficient β

[23]The following crops are included in the coarse cereal category: bajra, barley, gram, jowar, and ragi.

[24]The following crops are included in the other crop category: groundnut, jute, maize, potato, rapeseed and mustard, rice, sesamum, sugar, tur, and wheat.

thus captures the net causal impact of cyclone damage on agricultural production and crop mix choices. The regression also controls for time-invariant crop and district fixed effects, λ_d and μ_c respectively.[25] In addition, γ_t controls for common macro shocks[26], because agricultural policy is set by the central government. Given that the *cyclone exposure* variables are predicted regressors, the standard errors are clustered at the district level and bootstrapped. Lastly, the regression is weighted by the share in initial production of crop c in district d, since there is substantial variation in output across crops and districts – the 5th percentile of the output variable is 44 tons, whereas the 95th percentile is 150,800 tons.[27]

To investigate the dynamics of cyclone's net impact, I simply include three lags of the cyclone exposure variable into equation (1):

$$\begin{aligned} y_{cdt} =& \alpha + \beta_{10} cyclone_{dt} + \beta_{11} cyclone_{dt-1} + \beta_{12} cyclone_{dt-2} + \beta_{13} cyclone_{dt-3} \\ &+ \lambda_d + \mu_c + \gamma_t + \epsilon_{cdt} \end{aligned} \tag{2}$$

The results are robust to the inclusion of more lags. Since the effects generally lose significance with the second or third lag, I focus the analysis on this specification. On the one hand, we would predict that the coefficients on the lags are statistically significantly different from zero, as producers take time to recover from the shock. Yet, these coefficients should lose significance over time, as the empirical literature on large temporary shocks does not find lasting changes in production behaviour (Davis and Weinstein, 2002; Miguel and Roland, 2006).[28]

Furthermore, I can carry out a falsification exercise to verify that I am indeed capturing the impact of an exogenously determined natural disaster. Specifically, I include three leads of the cyclone exposure variable into regression (2). Their coefficients should be jointly insignificantly different from zero, if I am indeed capturing the impact of an exogenously determined natural disaster. That is, there should be no pre-trends in the data that are captured by the cyclone variable after controlling for common macro shocks.

Secondly, I disentangle the direct impact of the cyclone on agricultural production from the subsequent behavioural response of the producers to shock. To do this,

[25] The inclusion of crop-by-district fixed effects does not change the results substantially. This is not surprising, as each crop category comprises several varieties, which the dataset does not distinguish separately (Sawhney and Daji, 1961). Thus, there should not be much variation in the *observed* suitability of one crop across districts, as farmers have already chosen the optimal variety. Moreover, districts are small geographical units, so that most of the variation is absorbed by λ_d.

[26] The results are robust to the inclusion of crop-by-year fixed effects. This finding makes intuitive sense, as it is a well-known fact that markets are not well-integrated in India (Topalova, 2004; Duflo and Pande, 2007; Guiteras, 2007). In fact, my results below show that farm harvest prices do respond to local shocks. Hence, it is unlikely that crop-level shocks equally affect all districts within India.

[27] This weight is calculated as the share in total production of crop c in district d for 1956-1958. If the crop was not produced during this time period the output level in the first year that it was planted is used for the calculation of the share.

[28] As has already been mentioned in footnote 2, permanent production shocks can have general equilibrium effects, as behavioural responses are adjusted (Hornbeck, 2008).

I estimate the following weighted OLS fixed effects regression – with clustered and bootstrapped standard errors – for the sample of winter and summer crops separately:

$$\begin{aligned} y_{cdt} =& \alpha + \beta_1 \textit{pre-sowing cyclone}_{dt} + \beta_2 \textit{pre-harvest cyclone}_{dt} + \\ & \lambda_d + \mu_c + \gamma_t + \epsilon_{cdt}, \end{aligned} \tag{3}$$

where *pre-sowing cyclone*$_{dt}$ and *pre-harvest cyclone*$_{dt}$ measure the cyclone exposure for all event prior to the sowing and harvesting season of the winter or summer crops respectively. The coefficient β_1 thus captures the behavioural response of the producers. In contrast, the coefficient β_2 measures the direct impact of the cyclone on the outcome variable. To estimate the persistence of these effects, I include three lags of *pre-sowing cyclone*$_{dt}$ and *pre-harvest cyclone*$_{dt}$ into equation (3).

Finally, to test for changes in the crop mix, I simply run equations (1) and (2) on the sample of coarse cereals and all other food crops one at a time. The sample of coarse cereals comprises bajra, barley, gram, jowar, and ragi. The other food crops include groundnut, jute, maize, potato, rapeseed and mustard, rice, sesamum, sugar, tur, and wheat.

Section 1.5 now presents the results of the empirical analysis.

1.5 Results

1.5.1 The net effect of tropical cyclones on agricultural production

I begin by estimating equations (1) and (2) using the *cyclone exposure* variable and the entire sample of crops. These results are presented in Table 1.2. Columns 1-6 provide the estimates for the area planted, output, farm harvest prices, earnings, yields, and calorie content respectively. Panel A displays the estimates for the contemporaneous effect (obtained from estimating equation (1)) and Panel B for the medium-run impact (i.e., the estimates of equation (2)). Note that the regressions for the calorie content exclude zero-calorie crops like cotton and tobacco, which is why column 6 has fewer observations than columns 1-5.[29]

As expected, I find an overall reduction in output if a district is struck by a cyclone. In fact, column 2 of Panel A shows that a one standard deviation increase in the *cyclone*$_{dt}$ variable lowers the quantity produced by 7.75%. Rather surprisingly, I also find that the area planted increases by 3.02% in response to the cyclone event (Panel A, column 1). This finding shows that farmers are indeed able to counter the aggregate income shock within the *same* agricultural year by expanding their production – a response which will be explored in more detail in Sections 1.5.2 and 1.5.3. However, it is important to note that this change in behaviour is not sufficient to offset the damage caused by the tropical cyclone. In fact, the reduction in output estimated above captures the *net* impact, where the traces of destruction have been

[29]The results in columns 1-5 are robust to the exclusion of the zero-calorie crops.

lessened but not removed by the behavioural response. Moreover, given that these two effects are diametrically opposed, the net impact on prices, earnings, yields, and calorie content is indeterminate and insignificant.

Panel B estimates the dynamic impact of cyclone exposure on agricultural production by including three lags of the $cyclone_{dt}$ variable into the regression. The medium-run behavioural response – estimated by calculating the sum of the immediate effect and all three lags – is even larger than in the main specification. At the end of three years the area planted has increased by 9.49% (Panel B, column 1, sum of lags). This continued expansion in production translates into a similar rise in output, earnings, and calories in the three years after the shock (the sum of the three lags *only* increases by 13.19%, 22.97%, and 13.14% respectively). However, due to the negative contemporaneous effect of the cyclone, the medium-run impact is positive but insignificant (Panel B, columns 2, 4, and 6, sum of lags). That is to say, recovery from a cyclone shock is slow – the behavioural response only manages to offset the damage after three years. This is not surprising given that an average cyclone shock affects 21.41% of the district's area. Finally, it is interesting to notice that the increase in output also leads to a medium-run reduction in local prices of 2.33% (Panel B, column 3, sum of lags). This last finding supports the long-standing assumption in the academic literature that markets are not well-integrated in India (Topalova, 2004; Duflo and Pande, 2007; Guiteras, 2007).

These results are robust to using alternative measures of cyclone exposure. To illustrate this, Table A.1 in the appendix shows the estimates of the contemporaneous effect for the *severe cyclone exposure* variable in Panel A and the *minimum distance* variables in Panel B. The estimates for the *severe cyclone exposure* variable are of similar magnitude and sign. However, they are less precisely estimated, because the number of cyclone shocks is restricted considerably through the truncation of the predicted values. I therefore prefer the *cyclone exposure* variable. The findings for the *minimum distance* variable are also of similar magnitude but opposite sign, because this variable is smaller the closer the district is to the cyclone track. Nonetheless, given that it is less correlated with the UNEP data than the *cyclone exposure* variable[30], I use the latter for the remainder of the empirical study.

Before proceeding with the analysis, it is important to establish that I am indeed capturing the impact of tropical cyclones. I therefore carry out a falsification exercise by including three leads of the $cyclone_{dt}$ variable into the regression. Table 1.3 presents these results. Reassuringly, the coefficients of the leads are jointly insignificant for all six outcome variables. Furthermore, the main results from Table 1.2 remain largely unchanged: tropical cyclones destroy output, but farmers respond by increasing production. However, this behavioural response is not sufficient to counterbalance the damage caused in the year of the shock. Districts only recover after three years. This

[30]The correlation with the UNEP wind speed buffers is only 0.12 in absolute value for the *minimum distance* variable, but 0.44 for the *cyclone exposure* variable.

evidence enhances confidence that I am indeed capturing the causal impact of cyclone damage on agricultural production.

Further analysis is now needed to disentangle the ways in which farmers cope with these large and mostly uninsured aggregate income shocks. By taking advantage of the detailed nature of my dataset, I can identify two possible coping strategies: income smoothing across growing seasons and changes in the crop mix towards cheap calories discussed in Sections 1.5.2 and 1.5.3 respectively.

1.5.2 Coping strategy 1: income smoothing across growing seasons

As has been explained in Section 1.4, I can identify the destruction caused by the cyclone and the subsequent behavioural response separately by following a simple logic. Specifically, I use the date of the natural disaster and the district-specific growing seasons to distinguish between cyclone events that occurred before or after the sowing time of a given crop. Shocks prior to planting will identify the behavioural response. In contrast, the cyclone's destructive potential will be captured by all events that happened after sowing but before the harvest. Given that there are about 20 crops in the sample, I group them according to the two main growing seasons in India; namely, the summer crops that are planted at the beginning of the year and the winter crops, which are sown in second half. This identification strategy has already been formulated in equation (3) and is now estimated for each group separately.

The results show a very consistent picture – crops are destroyed if the cyclone hits prior to the harvest, but more are grown if it strikes before the sowing season. This is clearly illustrated in Table 1.4, which presents the findings for the winter crops only.[31] The results in Panel A show that farmers expand their production of winter crops if a cyclone hits prior to their sowing season; a one standard deviation increase in the *pre-sowing* $cyclone_{dt}$ variable increases the area planted, output, earnings, and calorie content by 9.97%, 18.77%, 11.16%, and 13.83% respectively (columns 1, 2, 4, and 6, cyclone before sowing). The effect on yields is insignificant, because both the area planted and output increase. These effects persist in the medium-run as shown in Panel B (columns 1, 2, 4, and 6, cyclone before sowing, sum of lags).

On the contrary, cyclone events before the harvest only affect output. We would expect this coefficient to be larger than the net effect estimated in Table 1.2, which also captures the mitigating behavioural response. This is the case, as a one standard deviation increase in the *pre-harvest* $cyclone_{dt}$ variable lowers the quantity produced by 11.55% (Panel A, column 2, cyclone before harvest) – the net effect is only −7.75%. As a second consistency check, it is important to notice that there is no effect on the area planted (Panel A, column 1, cyclone before harvest). This makes intuitive sense, since the *pre-harvest* $cyclone_{dt}$ variable only captures the effect of tropical cyclones

[31] A crop is defined as a winter crop if it is sown in a given district after the first week of August. These crops are then harvested in the winter time or early spring of the following year. The results are robust to changing the cut-off date.

after planting decisions have been made. Finally, production does not change in the medium run – the estimates for all six outcome variables are insignificant in Panel B. This provides suggestive evidence that farmers only adjust their behaviour in the growing season, which is at a lower risk from cyclone shocks.

Analogous results are obtained for the summer crops shown in Table 1.5. In fact, the highly significant coefficients in Panel A exhibit a similar pattern: a one standard deviation increase in the *pre-sowing* $cyclone_{dt}$ variable increases the area planted, output produced, earnings, and calorie content by 5.34%, 22.63%, 23.10%, and 21.46% respectively (columns 1, 2, 4, and 6, cyclone before sowing). This time the cyclone shock also has a negative and highly significant impact on local prices, which decline by 1.33% (columns 3, cyclone before sowing). These effects persist in the medium run and are even stronger and more significant than for the winter crops. Furthermore, it is interesting to note that yields now also increase in the medium run (column 5, cyclone before sowing, sum of lags). This suggests that the production of summer crops expands both along the extensive *and* intensive margin. Lastly, similar to Table 1.4 a tropical cyclone hit before the harvest only lowers production by 13.66% (Panel A, column 2, cyclone before harvest). There is, again, no persistent impact.

The evidence clearly shows that cyclones destroy output and that this effect is larger than the overall results would suggest. In addition, producers attempt to smooth income after a cyclone shock by expanding production in the subsequent growing season of the *same* agricultural year. Specifically, they plant high calorie crops on new land – an effect which persists in the medium run. The following section now explores this change in the crop mix in more detail.

1.5.3 Coping strategy 2: changes in the crop mix towards cheap calories

This section explores the second income smoothing mechanism that I can identify using my detailed dataset. In particular, it investigates if farmers adjust their crop mix towards coarse cereals after a cyclone shock. As has been argued above, these crops need few inputs and can grow on a variety of soils, including low-quality land. This will make it easier for producers to smooth income if the tropical cyclone has destroyed their capital and variable inputs. Moreover, coarse cereals have a high calorie content on average. Farmers will thus be able to meet basic nutritional needs after the natural disaster struck by planting these crops. Nonetheless, since they have a low economic value, we would expect any changes in the crop mix to be only temporary. To study this coping strategy, I estimate equations (1) and (2) for the sample of coarse cereals and other food crops successively.

Table 1.6 presents the results for the coarse cereals. Similar to Table 1.2, the estimates capture the *net* effect of cyclone exposure. Consistent with our expectations, the findings clearly show that farmers choose to produce more coarse cereals after a cyclone shock. As a matter of fact, they increase production so much that the damage is offset completely, i.e. the area planted, output, and calorie content experience a net increase of 4.22%, 11.47%, and 11.70% respectively (Panel A, columns 1, 2, and 6). This rise in output leads to a contemporaneous drop in local farm harvest prices of 2.53% (Panel A, column 3). Given these opposing forces, the impact on earnings and yields is again indeterminate and insignificant (Panel A, columns 4 and 5).

Furthermore, the estimates of the dynamic effects in Panel B show that farmers continue to put more effort into the production of coarse cereals in the years after the shock. However, this increase is solely along the intensive margin, as output and calorie content rise (and prices fall) without a concomitant increase in the area planted (Panel B, columns 2, 3, and 6, sum of lags). This effect could be driven by a more intensive application of fertilizer, which can substantially increase yields of coarse cereals (Sawhney and Daji, 1961). Unfortunately, I am unable to investigate this mechanism further, as the input data is not available at the crop level.

The evidence for all other food crops is presented in Table 1.7. It complements the story of income and nutrition smoothing. In particular, Panel A shows that these crops are largely destroyed by the cyclone shock and *not* replanted within the same year. The destruction is sizeable – a one standard deviation increase in the $cyclone_{dt}$ variable reduces output by 13.93% (Panel A, column 2). There is no significant contemporaneous effect on any of the other outcome variables. It should also be noted that the production of other food crops is not adjusted in the medium run. The results in Panel B of Table 1.7 show no statistically significant sum of lags for any of the six outcome variables. This is not surprising, since these crops have yields that are on average three times higher than those of the coarse cereals and fetch prices that are twice as large. Profit-maximising farmers will therefore prefer to grow other food crops.

Summarizing the results of the last three sections yields a consistent story of the impact of tropical cyclones on the primary sector in India. First of all, the destruction caused by a tropical cyclone is sizeable. Given that farmers are largely uninsured against these aggregate income shocks, they have to change their production behaviour. Due to there being multiple growing seasons in India, they will be able to expand production in the growing season directly following the cyclone shock. They also change their crop mix towards cheap calories, such as coarse cereals, to smooth both income and nutrition. This will lessen the impact of the natural disaster, but will not be sufficient to offset the damage in the year of the cyclone hit. However, producers continue to smooth income after the shock until recovery has been achieved about three years after the hit.

1.6 Conclusion

This paper has shown that the impact of tropical cyclones on the primary sector in India is devastating. Using a new digital database of cyclone exposure, I show that farmers need on average three years to recover from the aggregate income shock. Given that natural disasters are largely uninsured, producers have to change their production behaviour in an attempt to smooth income and nutrition. By exploiting the detailed nature of my dataset I can identify two coping strategies. On the one hand, producers smooth income intertemporally by expanding their production in the growing season directly following the cyclone shock. On the other, they change their crop mix towards coarse cereals, which are highly resilient crops that provide cheap calories.

These results clearly show that the focus of disaster relief on short-term supportive measures is insufficient. The first relief operations should be seamlessly followed by medium-run aid to help rebuild sustainable livelihoods and smooth income. The attitudes towards introducing a development component into disaster relief have shifted only recently in India. Since the aforementioned super-cyclone struck Orissa in 1999 and a major earthquake hit Gujarat in 2001, an initiative has been launched by the High Powered Committee on Disaster Management to develop a more comprehensive approach towards natural disasters (HPC, 2001). The findings of this paper stress the need to push ahead with this initiative, especially since climate change will enhance the frequency and intensity of cyclonic storms.

However, more research is needed to identify effective mitigation and rehabilitation methods. Furthermore, due to data limitations the analysis cannot study the behavioural response of districts that are historically more exposed to cyclone damage. The coping strategies and recovery process might be quite different in more vulnerable communities. More disaggregated data is needed to investigate this further. Finally, it is important to note that the possibilities to smooth income and nutrition are exceptional in India. The impact of tropical cyclones will be more pronounced in developing countries where there is only one growing season and land is already fully exploited. It thus becomes even more important to help producers recover from the shock and protect themselves more effectively against future shocks.

A Data appendix

This appendix provides information complementary to Section 1.3 on the variables used in this paper.

A.1 The district sample

The district sample used in this analysis is the same as in the World Bank's India Agriculture and Climate Dataset. It covers 271 districts within thirteen states of India (Andhra Pradesh, Bihar, Gujarat, Haryana, Karnataka, Maharashtra, Madhya Pradesh, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal). Kerala and Assam are the only major agricultural states that are not included in the dataset. However, both states were not hit by a cyclone during 1956-1987, which is the time horizon of the agricultural data. It also provides no information on Himachal Pradesh, Jammu and Kashmir, the Union Territories, and the small states in the Northeast of India. This does not limit the analysis, since these states are not very important agriculturally and only Pondicherry was hit by a cyclone during the sample period. The dataset corrects for any changes in the district boundaries by using the 1966 borders throughout the sample period.

A.2 The tropical cyclone wind speed buffers dataset, 1975-2007

The tropical cyclone wind speed buffers have been designed by UNEP/DEWA/GRID-Europe PREVIEW (Project for Risk Evaluation Vulnerability Information and Early Warning) for the Global Assessment Report on Risk Reduction (ISDR, 2009). The data is made available on the Global Risk Data Platform and can be downloaded at `http://preview.grid.unep.ch/index.php?preview=data&events=cyclones&lang=eng`. This database compiles cyclone tracks of more than 96 knots for 1975-2007 from the Regional Specialized Meteorological Centers of the World Meteorological Organization and the Tropical Cyclone Warning Centers. The wind speed buffers are then estimated with a state-of-the-art GIS model based on (Holland, 1997).

I import this data into the ArcGIS software and extract the buffers for the Indian subcontinent. These polygons are then overlaid with a district map of 1966, which corresponds to the boundaries used by the World Bank's India Agriculture and Climate Dataset. On this basis, I am able to calculate the area covered by the wind speed buffers for each district. Furthermore, I use the 1966 boundaries to calculate the district's total area. Both of these variables are exported into STATA, where the cyclone data is collapsed by month and year. Subsequently, I calculate the percentage of the district's area covered by the wind speed buffers. It is possible that the area affected is larger than the total size, since more than one cyclone can hit the same district in a given month. However, since this is only the case for 3.47% of the observations, I restrict this variable to being less than or equal to 100%.

A.3 The eAtlas cyclone tracks dataset, 1891-2008

The cyclone track information for 1891-2008 has been obtained from a digital database of the India Meteorological Department (IMD). The data can be ordered at `http://www.imdchennai.gov.in/cyclone_eatlas.htm`. This is an electronic version of the hard copy editions of its widely referred atlas *Tracks of Storms and Depressions in the Bay of Bengal and the Arabian Sea*, which were published by IMD in the years 1964, 1979 and 1996.

Since this dataset covers a long time horizon, only cyclone track information is available. The IMD has developed a consistent classification for all years to distinguish cyclones from depressions and minor storms. More specifically, a class 1 storm corresponds to wind speed of less than 33 knots, which is characteristic of depressions. Cyclonic storms with a strength of 33-47 knots are coded as class 2 disturbances and all events with wind speeds of more than 47 knots are allocated to the class 3 group, which corresponds to severe cyclones. This analysis only focuses on the class 3 events, since they most closely correspond to the commonly used definition of a cyclone.

The raw data is supplied in terms of coordinates that mark the start and end point of a line segment, which is allocated either to class 1, 2 or 3. In addition, information is provided on the start and end date of each line segment. I import this data into the ArcGIS software and create unified cyclone tracks for the class 3 disturbances. I then overlay this data with the 1966 district map used for the wind speed buffers and calculate the total track length for the district itself and all the districts within a 400km radius. This data is exported into STATA and collapsed by month and year.

To select suitable right-hand-side variables for the prediction, it is helpful to examine the UNEP data in Figure 1.1. Several features are noteworthy:

- *The wind speed buffers are much larger for stronger cyclones, which suggests a possible non-linear relationship between the area affected and the strength of the storm.* I calculate the total track length of a cyclone within a district. This variable measures the exposure to the storm. Furthermore, I square this measure to model possible non-linear effects.
- *A single cyclone usually affects several districts at once. The maximum distance from the 'eye of the storm' to the outmost boundary of the wind speed buffers is about 400km. This impact should be smaller for districts further away, as Figure 1.1 clearly shows that wind speeds decline monotonically with distance from the 'eye'.* I construct the total track length variable for all districts within a 400km radius. I then sum this data across two mutually exclusive groups: the first set of districts comprises the nearest neighbours that share parts of the district's boundary. The second group includes all those districts that are within a 400km radius, but do not directly adjoin the district. We would expect that cyclone shocks that hit the nearest neighbours should have a much larger impact on the

district. I square the total track length variables for both neighbour groups to model possible non-linear effects. This yields four more variables.

- *The wind speed buffers do not penetrate more than 160km into the country.* I restrict the eAtlas data accordingly.

I construct 10 additional variables, which capture the fact that cyclones in India are on average much stronger in autumn (IMD, 1979). That is, I create a dummy equal to 1 if the district is hit by a cyclone in a given year and month. This variable is then interacted with a dummy that switches on in or after September. I also calculate the total number of cyclone hits for the two neighbourhood groups respectively to measure exposure. I then create similar interactions with the autumn dummy. Furthermore, I capture possible non-linear effects by squaring the total number of cyclone hits for the two neighbourhood groups and interacting them with the autumn dummy.

These 16 variables are then regressed on the percentage of the district covered by the UNEP wind speed buffers for 1977-2003.[32] The estimated coefficients are used to predict the percentage of the district's area exposed to a cyclone shock out of sample. Note that I have to set 8.25% of the predicted values equal to zero, since they are estimated to be negative. This *cyclone exposure* variable is then also restricted to damage more than 25% of the district's territory to determine the *severe cyclone exposure* variable. Both variables are then summed across months to create annual measures to match the agricultural data.

Finally, I construct the third cyclone variable, which is solely based on the eAtlas cyclone track data. In ArcGIS, I calculate the Euclidian distance to the nearest cyclone track for 125 x 125 metre cells. I then use the zonal statistics tool to calculate the district's minimum distance to the nearest cyclone track. The data thus obtained is imported into STATA and collapsed by year.

A.4 The data for the primary sector, 1956-1987

The data for the primary sector comes from the World Bank's India Agricultural and Climate Dataset, which has been constructed by Sanghi, Kumar and McKinsey and used in Dinar et al. (1998). It can be downloaded at `http://ipl.econ.duke.edu/dthomas/dev_data/index.html`.[33] This dataset is the most comprehensive dataset on the agricultural sector in India. It covers the time period of 1956-1987, where one year refers to an agricultural year.

This database contains district-level information on farm harvest prices, agricultural output and area planted broken down for six major (bajra, jowar, maize, rice, sugar, and wheat) and 14 minor crops (barley, cotton, gram, groundnut, jute, other pulses, potatoes, ragi, rapeseed and mustard, sesamum, soy, sunflowers, tobacco, and

[32] Despite the fact that the period of overlap ranges from 1975-2007, cyclones only reach the Indian shore between 1977 and 2003.

[33] More information on the dataset is also provided at `http://ipl.econ.duke.edu/dthomas/dev_data/datafiles/india_agric_climate.htm`.

tur). The dataset also contains district-level price and quantity information for inputs, such as fertilizer, agricultural labour, cultivators, tractors and bullocks. However, the data on the labour and capital inputs is interpolated, which makes it unsuitable for this analysis. Therefore, the analysis only uses the crop-level data on area planted, output, prices (deflated to 1980 INR), earnings (defined as the product of deflated prices and output) and yields (defined as the ratio of earnings to the area planted).

The output data is also converted into its calorie content by using information obtained from MeDINDIA, which is a premier health portal in India. The calorie content per 100g was downloaded for each individual crop on the 22nd of July 2010 at `http://www.medindia.net/patients/foodcalories/index.asp`. For several types of crops more than one calorie estimate was available. For instance, the website distinguishes between several types of gram, whereas the WB dataset only provides information on the overall category. I have thus taken the average of these values. In addition, I have used the value of unprocessed crop for maize, rice, and wheat.

A.5 The crop calendar

The raw data was entered from the Indian Crop Calendar (GOI, 1967). This publication lists the dates of sowing, harvesting and marketing times for each crop and district. The listed dates are averages, so that actual crop- and district-specific dates for these agricultural activities differ from plot to plot and from year to year. The crops listed in this publication include the 25 most important crops in India, but only the most relevant crops were listed for each district. The Crop Calendar does not discuss how the crop selection decision was determined at the district level. Not all 339 districts in India (in 1961) were accounted for in the Crop Calendar, but most were. Missing data was replaced by the state-level minimum (maximum) for the start (end) of the respective crop and season.

The sowing and harvesting dates in the Crop Calendar were listed in a variety of forms. Usually a range of dates was provided in which sowing or harvesting could occur. This range of dates could either be precisely listed in terms of dates or simply list a range of months. I convert the start and end points of the date range into individual calendar dates by taking the largest possible window. For example, sowing in 'April-May' would be coded as a sowing start date of '1st of April' and a sowing end date as '31st of May'. The precise calendar dates are then converted into the serial number for that day of the year, counting Jan 1st as 1 and December 31st as 365. Note that the agricultural season in India runs from roughly May to the following April. Hence, many of the sowing/growing/harvesting windows span two different calendar years. The numbering is continued accordingly.

Figure 1.1: UNEP wind speed buffers and eAtlas cyclone tracks, 1977-2003

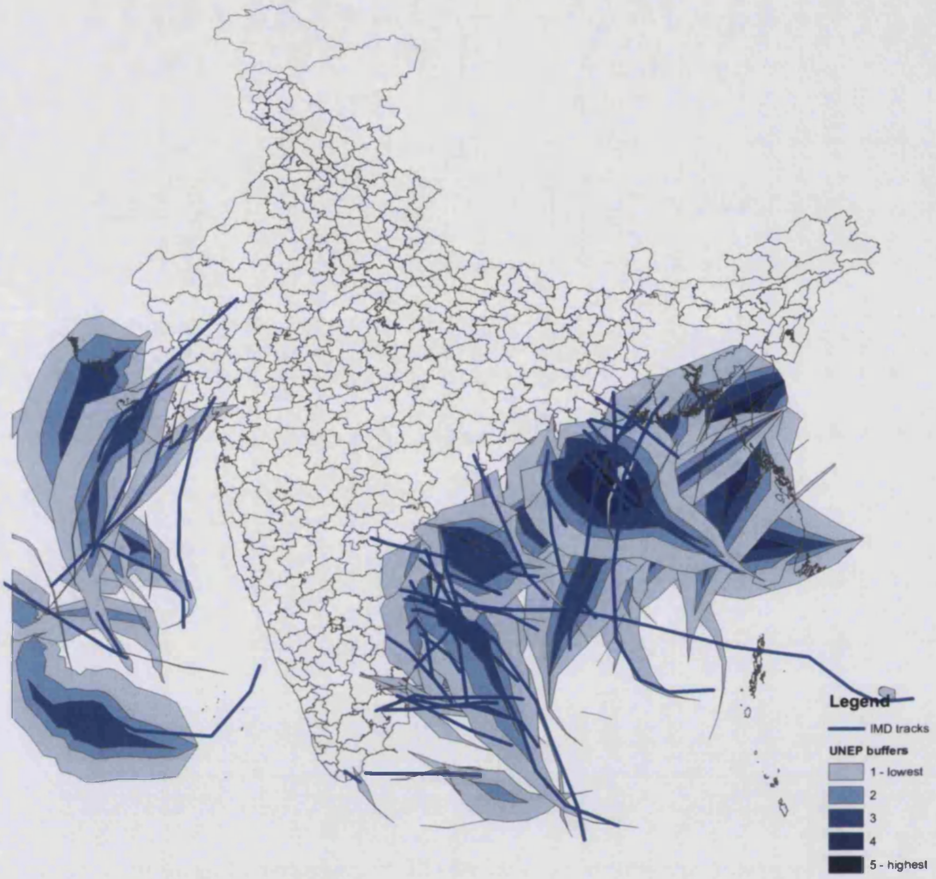


Figure 1.2: Correlation between the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(a) % area exposed, N=906

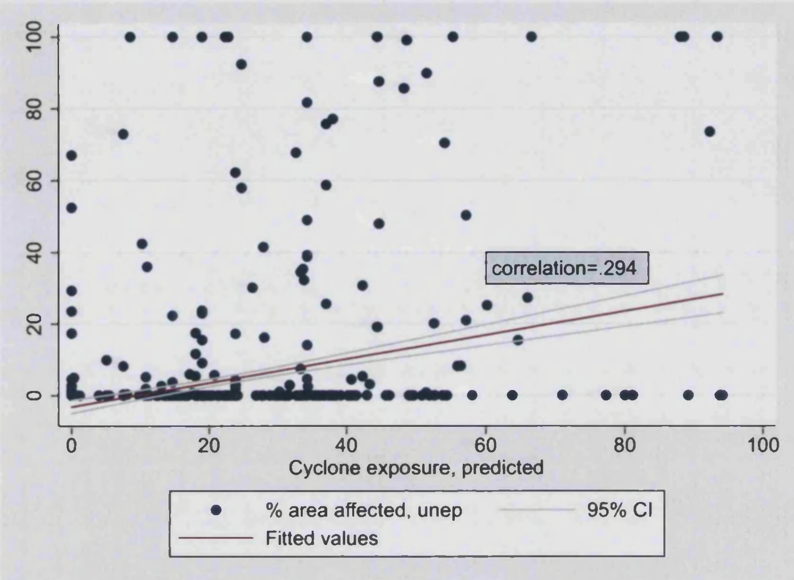


Figure 1.2: Correlation between the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(b) % area exposed > 0 for the UNEP data, N=108

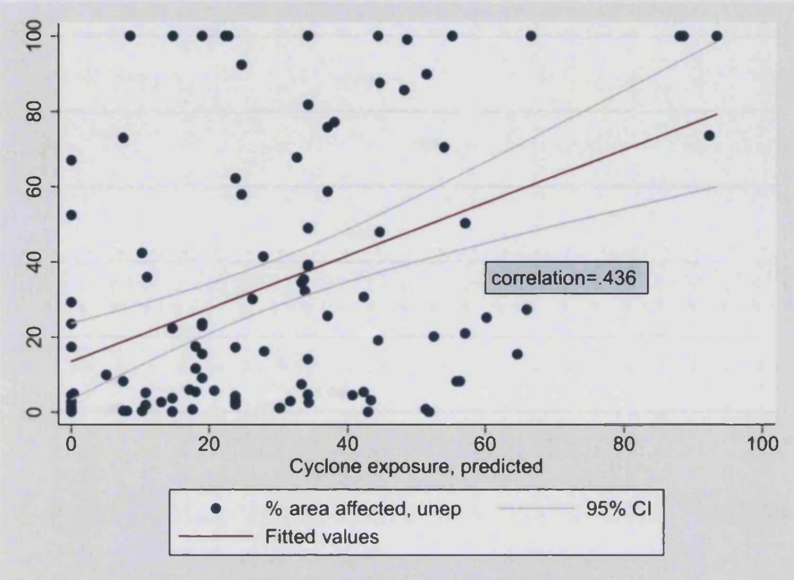


Figure 1.2: Correlation between the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(c) % area exposed ≥ 25 for the UNEP data and eAtlas prediction, N=33

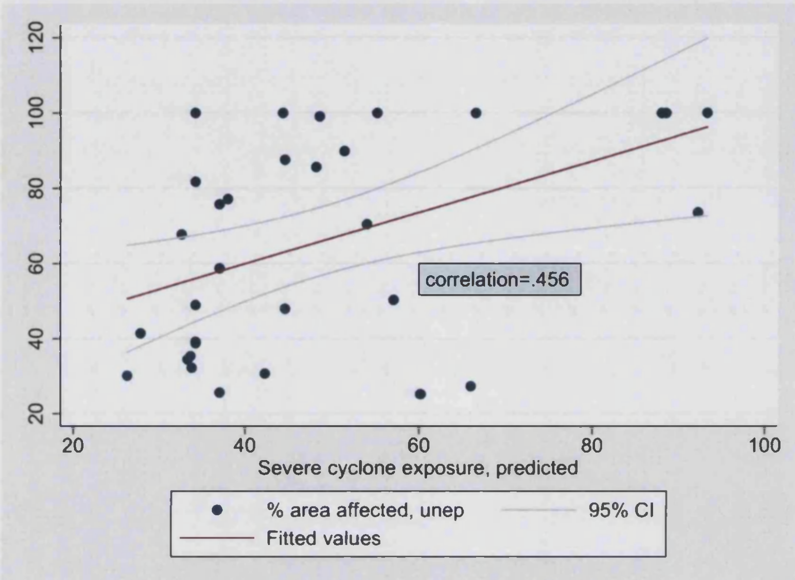


Figure 1.3: Distribution of the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(a) % area exposed > 0 for the UNEP data, N=108

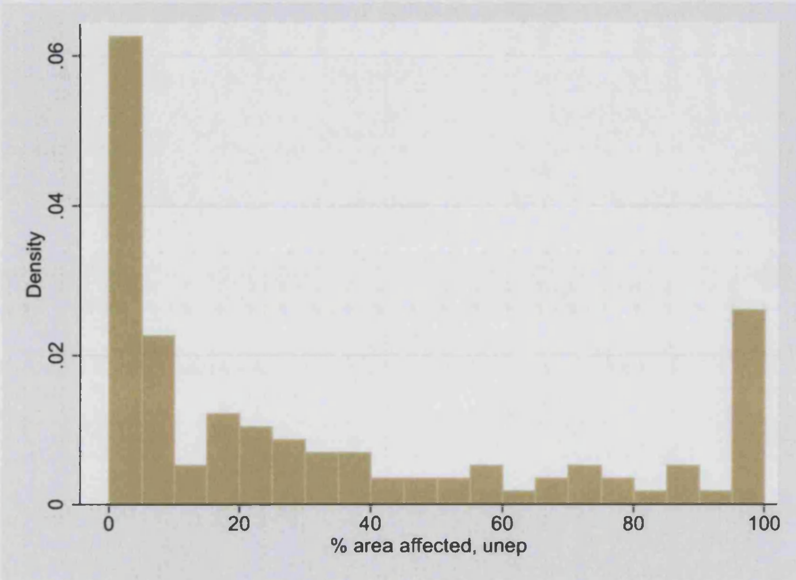


Figure 1.3: Distribution of the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(b) % area exposed > 0 for the eAtlas prediction, N=855

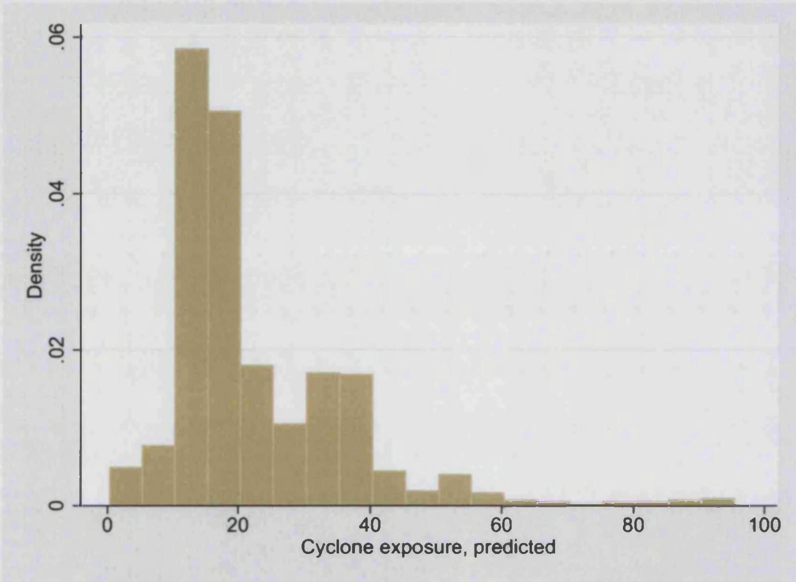


Figure 1.3: Distribution of the % area exposed as measured by the UNEP wind speed buffers and as predicted by the eAtlas cyclone tracks, 1977-2003

(c) % area exposed $\geq$ 25 for the eAtlas prediction, N=258

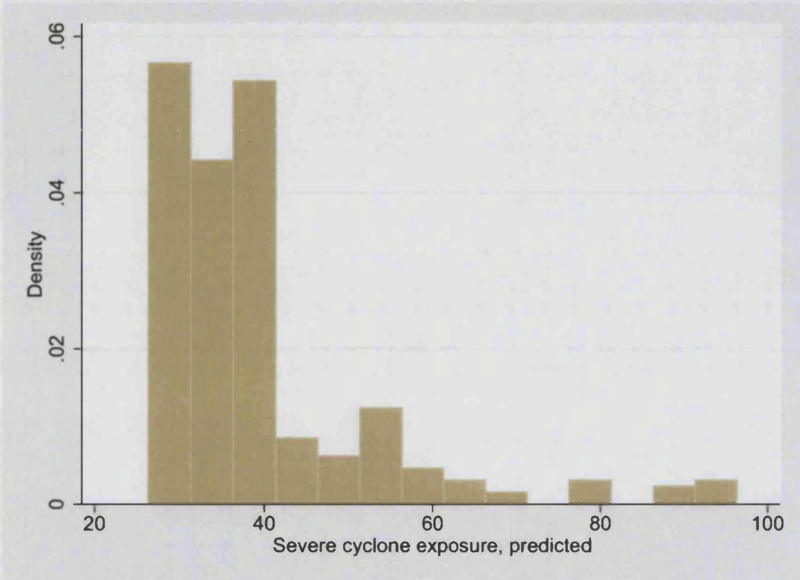


Table 1.1: Summary statistics of the cyclone exposure variables, 1956-87

Panel A: Total sample

	(1)	(2)	(3)
Data Source	Cyclone exposure	Severe cyclone exposure	Minimum distance
Mean	2.368	1.049	775.402
Std deviation	7.964	6.751	525.177
Minimum	0	0	0
Maximum	93.245	93.245	2242.576
N	10287	10287	10287

Panel A: Cyclone events only

	(1)	(2)	(3)
Data Source	Cyclone exposure	Severe cyclone exposure	Minimum distance
Mean	21.408	39.539	237.461
Std deviation	12.878	13.999	163.460
Minimum	2.244	25.107	0
Maximum	93.245	93.245	625.848
N	1138	273	1138

Notes: The *cyclone exposure* and *severe cyclone exposure* variables are measured in percent. The *minimum distance* variable is measured in kilometres. Data sources and construction are described in full in Appendix A.

Table 1.2: The net effect of a cyclone shock

Panel A: The contemporaneous effect

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	70.634**	-168.947*	-0.077	-136.167	-1.031	-37.950
	(32.187)	(101.963)	(0.052)	(203.595)	(1.886)	(43.776)
N	104,288	104,288	104,288	104,288	104,288	89,056
R-squared	0.731	0.629	0.640	0.689	0.247	0.624
Mean LHS var	30,168	28,066	269	51,059	2,649	11,368
	(69,282)	(76,537)	(210)	(137,751)	(13,851)	(29,079)
Mean Cyclone	21.408	21.408	21.408	21.408	21.408	21.408
	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)

Panel B: The persistence of the cyclone shock

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	65.117**	-180.195*	-0.061	-187.982	-1.142	-42.571
	(30.620)	(101.450)	(0.053)	(197.584)	(1.777)	(43.219)
Lag 1	75.174**	221.999***	-0.130***	433.905***	0.134	94.320***
	(35.522)	(54.991)	(0.039)	(150.699)	(1.123)	(21.341)
Lag 2	81.884***	148.040**	-0.209***	327.191	5.122**	60.948**
	(28.304)	(62.600)	(0.048)	(288.282)	(2.358)	(26.124)
Lag 3	0.082	-82.497	-0.088*	149.689	1.729	-39.257*
	(35.130)	(66.612)	(0.046)	(243.392)	(1.579)	(22.777)
Sum of lags	222.257**	107.347	-0.487***	722.802	6.842	73.440
	(98.484)	(175.674)	(0.114)	(663.568)	(4.739)	(74.638)
Joint p	0.014	0.000	0.000	0.011	0.084	0.000
N	104,288	104,288	104,288	104,288	104,288	89,056
R-squared	0.731	0.629	0.641	0.689	0.247	0.624

Notes: The *cyclone exposure* is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table 1.3: A falsification exercise

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	63.950**	-200.615**	-0.089*	-284.618	-1.234	-49.359
	(28.649)	(98.136)	(0.052)	(210.865)	(1.690)	(42.520)
Lag 1	73.899**	229.777***	-0.105***	469.779***	0.072	96.443***
	(35.906)	(56.580)	(0.040)	(157.727)	(1.169)	(21.808)
Lag 2	80.461***	127.659*	-0.235***	230.660	5.015**	54.110**
	(26.448)	(65.037)	(0.049)	(290.973)	(2.398)	(26.876)
Lag 3	-1.833	-75.226	-0.064	179.555	2.638*	-37.308*.
	(34.018)	(63.208)	(0.046)	(238.230)	(1.531)	(21.472)
Lead 1	15.337	58.105	0.177	209.678	0.608	14.119
	(34.155)	(92.156)	(0.147)	(193.757)	(1.821)	(34.509)
Lead 2	24.998	27.798	-0.151	144.153	1.444	15.146
	(34.258)	(48.192)	(0.136)	(155.401)	(1.599)	(18.666)
Lead 3	7.676	142.495*	-0.019	226.059	0.995	48.985
	(40.036)	(74.570)	(0.051)	(177.261)	(1.250)	(27.891)
Sum of lags	216.476**	81.596	-0.493***	595.377	6.492*	63.885
	(93.120)	(172.362)	(0.112)	(668.487)	(3.677)	(73.660)
Joint p lags	0.002	0.000	0.000	0.014	0.080	0.000
Joint p leads	0.793	0.207	0.307	0.405	0.717	0.228
Joint p	0.001	0.000	0.000	0.006	0.046	0.000
N	104,288	104,288	104,288	104,288	104,288	89,056
R-squared	0.731	0.629	0.641	0.689	0.247	0.624

Notes: The *cyclone exposure* is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table 1.4: The impact of cyclones on winter crops

Panel A: The contemporaneous effect

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone before sowing	282.457**	645.684**	-0.011	625.270**	-15.919	163.349***
	(132.097)	(295.005)	(0.148)	(256.348)	(11.601)	(45.805)
Cyclone before harvest	-71.464	-206.946*	-0.136	-292.612	13.030	-55.945
	(53.526)	(105.781)	(0.097)	(300.605)	(13.229)	(34.900)
N	34,144	34,144	34,144	34,144	34,144	31,520
R-squared	0.791	0.531	0.729	0.580	0.580	0.538
Mean LHS var	28,197	34,217	265	55,718	2,766	11,747
	(57,933)	(87,593)	(237)	(126,290)	(6,789)	(29,923)
BEFORE SOWING						
Mean Cyclone	20.060	20.060	20.060	20.060	20.060	20.060
	(9.949)	(9.949)	(9.949)	(9.949)	(9.949)	(9.949)
BEFORE HARVEST						
Mean Cyclone	26.882	26.882	26.882	26.882	26.882	26.882
	(19.096)	(19.096)	(19.096)	(19.096)	(19.096)	(19.096)

Panel B: The persistence of the cyclone shock

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone before sowing	314.276**	791.583**	0.008	791.405***	-15.803	200.130***
	(139.585)	(346.068)	(0.157)	(295.048)	(11.881)	(54.745)
Cyclone before harvest	-40.503	-155.642*	-0.117	-216.406	10.952	-36.483
	(51.810)	(93.808)	(0.103)	(273.642)	(12.809)	(28.877)
BEFORE SOWING						
Sum of lags	910.358**	3549.928*	-0.202	3134.418**	-58.313	778.903***
	(372.960)	(1853.006)	(0.382)	(1353.758)	(49.933)	(207.951)
Joint p sowing	0.097	0.020	0.249	0.037	0.364	0.001
BEFORE HARVEST						
Sum of lags	-206.901	-383.647	-0.670	-1439.218	39.638	-288.464
	(138.287)	(618.848)	(0.409)	(980.837)	(26.382)	(188.914)
Joint p harvest	0.413	0.153	0.037	0.066	0.116	0.022
N	34,144	34,144	34,144	34,144	34,144	31,520
R-squared	0.791	0.531	0.729	0.580	0.581	0.539

Notes: The *cyclone exposure* variable is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. A crop is defined as a winter crop if its sowing time starts after the first week of August. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table 1.5: The impact of cyclones on summer crops

Panel A: The contemporaneous effect

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone before sowing	88.947**	303.744***	-0.193***	603.190***	0.524	128.179***
	(43.599)	(66.973)	(0.049)	(186.703)	(1.287)	(26.618)
Cyclone before harvest	52.312	-198.447*	-0.034	-162.452	-2.234	-49.461
	(33.016)	(107.286)	(0.057)	(219.496)	(1.487)	(45.708)
N	70,144	70,144	70,144	70,144	70,144	57,536
R-squared	0.817	0.754	0.629	0.761	0.224	0.729
Mean LHS var	31,127	25,072	271	48,791	2591	11,159
	(74,163)	(70,337)	(196)	(142,944)	(16,211)	(28,604)
BEFORE SOWING						
Mean Cyclone	26.881	26.881	26.881	26.881	26.881	26.881
	(18.682)	(18.682)	(18.682)	(18.682)	(18.682)	(18.682)
BEFORE HARVEST						
Mean Cyclone	24.819	24.819	24.819	24.819	24.819	24.819
	(17.264)	(17.264)	(17.264)	(17.264)	(17.264)	(17.264)

Panel B: The persistence of the cyclone shock

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone before sowing	71.958*	260.962***	-0.161***	489.226***	-0.115	112.734***
	(40.209)	(60.424)	(0.047)	(179.494)	(1.299)	(24.625)
Cyclone before harvest	37.907	-190.127*	-0.038	-166.545	-2.022	-46.009
	(33.887)	(112.677)	(0.058)	(228.955)	(1.583)	(47.834)
BEFORE SOWING						
Sum of lags	235.688**	461.670***	-0.378***	1687.151**	13.970***	177.721***
	(116.969)	(149.457)	(0.129)	(661.826)	(4.869)	(55.661)
Joint p sowing	0.082	0.000	0.000	0.001	0.045	0.000
BEFORE HARVEST						
Sum of lags	80.109	642.078	0.411	1607.884	54.181	309.636
	(186.114)	(541.156)	(0.291)	(1160.888)	(17.020)	(222.057)
Joint p harvest	0.046	0.093	0.514	0.468	0.203	0.337
N	70,144	70,144	70,144	70,144	70,144	57,536
R-squared	0.817	0.754	0.630	0.761	0.224	0.729

Notes: The *cyclone exposure* variable is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. A crop is defined as a summer crop if its sowing time starts before the second week of August. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table 1.6: The impact of cyclones on coarse cereals

Panel A: The contemporaneous effect

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	121.061*	182.549***	-0.260***	67.097	-0.562	64.906***
	(68.783)	(63.572)	(0.072)	(147.146)	(0.916)	(22.415)
N	26,528	26,528	26,528	26,528	26,528	26,528
R-squared	0.826	0.579	0.487	0.577	0.278	0.578
Mean LHS var	45,496	25,242	163	40,030	970	8,805
	(88,937)	(46,661)	(58)	(75,619)	(995)	(16,321)
Mean Cyclone	23.485	23.485	23.485	23.485	23.485	23.485
	(15.865)	(15.865)	(15.865)	(15.865)	(15.865)	(15.865)

Panel B: The persistence of the cyclone shock

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	131.494*	200.374***	-0.281***	73.622	-0.546	71.165***
	(70.622)	(64.159)	(0.078)	(149.156)	(0.999)	(22.653)
Sum of lags	-2.455	624.290**	-1.264***	7.712	-2.823	221.748**
	(251.522)	(301.322)	(0.158)	(503.789)	(2.364)	(106.494)
Joint p	0.042	0.001	0.000	0.134	0.000	0.001
N	26,528	26,528	26,528	26,528	26,528	26,528
R-squared	0.826	0.579	0.491	0.578	0.278	0.578

Notes: The *cyclone exposure* variable is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. The following crops are included in the coarse cereal category: bajra, barley, gram, jowar and ragi. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table 1.7: The impact of cyclones on all other food crops

Panel A: The contemporaneous effect

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	48.903	-278.174**	-0.063	-259.045	-0.852	-67.372
	(38.637)	(127.485)	(0.063)	(256.783)	(2.200)	(52.990)
N	62,528	62,528	62,528	62,528	62,528	62,528
R-squared	0.860	0.694	0.608	0.741	0.670	0.684
Mean LHS var	26,844	35,038	242	63,866	3,085	12,455
	(62,210)	(92,771)	(119)	(167,307)	(5,392)	(32,975)
Mean Cyclone	25.160	25.160	25.160	25.160	25.160	25.160
	(17.551)	(17.551)	(17.551)	(17.551)	(17.551)	(17.551)

Panel B: The persistence of the cyclone shock

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Cyclone exposure	41.904	-281.309**	-0.058	-322.650	-1.357	-68.597
	(36.528)	(126.407)	(0.064)	(240.568)	(2.108)	(52.010)
Sum of lags	168.266	-166.198	-0.248	572.705	6.530	-17.356
	(115.080)	(223.637)	(0.183)	(835.625)	(5.100)	(92.264)
Joint p	0.003	0.000	0.000	0.016	0.087	0.000
N	62,528	62,528	62,528	62,528	62,528	62,528
R-squared	0.860	0.694	0.608	0.741	0.670	0.684

Notes: The *cyclone exposure* variable is measured in percent. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. The following crops are included in the other crop category: groundnut, maize, potato, rapeseed and mustard, rice, sesamum, sugar, tur, and wheat. Data sources and construction are described in full in Appendix A. The standard errors are clustered at the district level and bootstrapped. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. *** 0.01, ** 0.05, * 0.10

Table A.1: The net effect of a cyclone shock: alternative cyclone variables

Panel A: Severe cyclone exposure variable

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Sev cyclone exp	72.257**	-157.345	-0.046	-129.021	0.028	-32.198
	(29.576)	(101.554)	(0.047)	(203.123)	(1.682)	(43.628)
N	104,288	104,288	104,288	104,288	104,288	89,056
R-squared	0.731	0.629	0.640	0.689	0.247	0.624
Mean LHS var	30,168	28,066	269	51,059	2,649	11,368
	(69,282)	(76,537)	(210)	(137,751)	(13,851)	(29,079)
Mean Cyclone	21.408	21.408	21.408	21.408	21.408	21.408
	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)

Panel B: Minimum distance variable

	(1)	(2)	(3)	(4)	(5)	(6)
LHS var	Area	Output	Price	Earnings	Yield	Calories
Min distance	-0.191	2.061**	0.002***	4.244**	0.048**	0.616*
	(0.317)	(0.896)	(0.000)	(1.972)	(0.019)	(0.359)
N	104,288	104,288	104,288	104,288	104,288	89,056
R-squared	0.731	0.629	0.640	0.689	0.247	0.624
Mean LHS var	30,168	28,066	269	51,059	2,649	11,368
	(69,282)	(76,537)	(210)	(137,751)	(13,851)	(29,079)
Mean Cyclone	21.408	21.408	21.408	21.408	21.408	21.408
	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)	(12.878)

Notes: The *severe cyclone exposure* variable is measured in percent. The *minimum distance* variable is measured in kilometres. The area data is measured in hectares, the output data in tons, and the farm harvest price and earnings data in 1980 Indian Rupees. The yield variable is defined as the ratio of earnings to area planted. The calorie content is obtained by converting the output volume into its calorie equivalent. Data sources and construction are described in full in Appendix A. The standard errors and the standard deviations of the dependent variable and the *cyclone exposure* variable are shown in parentheses. The standard errors are clustered at the district level and bootstrapped. *** 0.01, ** 0.05, * 0.10

2 The Political Economy of Tropical Deforestation[34]

2.1 Introduction

Satellite imagery reveals vast expanses of forest extending across the Amazon Basin, the Congo Basin, and South East Asia. These tropical forests are seen as critical to slowing global climate change (Stern, 2006; Nabuurs et al., 2007). They are also particularly rich in biodiversity. Last but not least, they are surrounded by large and relatively poor populations who could derive significant economic benefits over the medium to long run if the forests are sustainably managed (Curran et al., 2004).

And yet it is precisely these tropical forests that are being cut down. Repeated satellite imagery reveals that tropical forests, unlike the great forests in the Northern hemisphere, have been experiencing rapid rates of deforestation (Hansen and DeFries, 2004). It is estimated that relative to a baseline of 1900 the majority of tropical forest has already been felled, with the rate of deforestation accelerating in the last two decades (Holmes, 2002; FWI/GFW, 2002; Hansen et al., 2008). Today tropical deforestation accounts for almost 20% of global emissions of greenhouse gases (Hooijer et al., 2006; IPCC, 2007b; Kindermann et al., 2008). This is more than is contributed globally by the transportation sector as a whole, and is roughly equivalent to the total greenhouse gas contribution of the United States. In fact, tropical deforestation places Indonesia just behind the US and China as the third largest producer of greenhouse gases worldwide.

Understanding what lies behind tropical deforestation has therefore become a pressing issue. And here political economy factors are coming to the fore. While there is an extensive literature on the optimal management of forest resources (e.g., Dasgupta and Heal (1974), Samuelson (1976), Dasgupta (1982), Brown (2000)), and while most countries' official policy seeks to implement these types of sustainable logging systems, actual practice diverges significantly from best practice. Local bureaucrats and politicians have much to gain by allowing logging to take place outside official concessions (Barr et al., 2006) or by sanctioning the transport and processing of illegally harvested logs (Casson, 2001a). On net, in many cases over 50% of the wood yield involves some illegal action – the figure for Indonesia, for example, is estimated at 60-80% (CIFOR, 2004). In this context, viewing deforestation as the result of optimal forest extraction policies implemented by a central planner misses the reality of what happens on the ground. Instead, what matters are the incentives that local politicians and bureaucrats face to either protect tropical forests or to allow their destruction.

[34]This chapter draws on work that was carried out jointly with equal share by Robin Burgess (LSE), Matthew C. Hansen (SDSU), Benjamin Olken (MIT) and me.

This paper investigates how local political economy incentives affect deforestation. We focus on one country – Indonesia – which is home to one of the largest and most valuable tropical forest reserves in the world (FWI/GFW, 2002). Although all Indonesian forests are legally controlled by the national forest ministry, local district governments have a substantial *de facto* role, particularly as the gatekeepers for illegal logging. By using imagery from the MODIS satellite, which was put into orbit in December 1999, we are able to monitor, at a 250m by 250m resolution, what has happened to forest cover on an annual basis across the whole of Indonesia for the period 2000 to 2008 (Hansen et al., 2009). The fineness at which we can monitor forests also allows us to compare and contrast deforestation across localities and in four land use zones – the Production and Conversion Forest where logging is legal (within specific concessions) and the Conservation and Protection Forest (where logging is strictly illegal).

Using this data, we investigate how the incentives faced by local bureaucrats and politicians affect the rate of deforestation. First, building on insights from the industrial organization of corruption (Shleifer and Vishny, 1993; Olken and Barron, 2009), we show that Indonesian district governments appear to engage in Cournot competition with one another in the market for wood. To demonstrate this, we take advantage of the fact that between 1998 and 2008, the number of districts in Indonesia increased by 65%, from 292 to 483, with districts splits occurring at different times in different parts of the country. As wood markets in Indonesia are defined by geographical features (mountains and rivers) which map well onto provincial boundaries, we can test the Cournot model by examining whether increased competition between districts (as proxied by increasing numbers of districts within a province) affects the propensity to extract forest resources. The fineness of our data also enables us to distinguish between the effects of one's own district splitting (which also introduces a degree of disorganization as state capacity is re-established) and other districts splitting (which raises competition within the provincial wood market and affects incentives to log now versus later).

Using this approach, we find that districts do indeed engage in Cournot competition with one another, so that splitting up control of the forests into smaller and smaller jurisdictions leads to increased logging overall. Using the MODIS satellite data, we estimate that subdividing a province by adding one more district increases the overall deforestation rate in that province by 7.89%, with the increase coming at equal rates in forest zones where logging may be legal or illegal (the Production and Conversion Forest) and zones where all logging is illegal (the Conservation and Protection Forest). We find similar magnitudes when we examine official production records from forest concessions, validating the satellite data and suggesting that some of the increase may be occurring through district governments facilitating additional extraction in the Production Forest. Consistent with the Cournot model, we show that the increase in political jurisdictions drives down prices in the local wood market: adding one more district to a province reduces local prices by 3.29%, implying a local demand elasticity for logs between 2.3 and 4, depending on specification.

Second, we test whether local election pressures influence the rate of deforestation. Starting in 2005, local district heads began to be chosen through direct popular elections rather than being indirectly selected by the local legislature. When direct elections first arrived in a district was determined by when the district head's term came to an end, and the timing of these terms, in turn, was determined by the timing of district head appointments under Soeharto (Skoufias et al., 2010). This introduces asynchronicity in district elections which is plausibly orthogonal to patterns of forest loss and which we exploit to examine whether logging, and in particular illegal logging, increases in the run-up to these elections. Using this approach, we document a "political logging cycle" where local governments become more permissive *vis a vis* logging in the years leading up to elections. We find that deforestation in zones where all logging is illegal increases by as much as 42.7% in the year prior to an election. Although logging then falls after the election, on net direct elections increase deforestation in zones where all logging is illegal by 51.5% over the entire election cycle.

Third, just as the rents from facilitating logging may become more or less valuable depending on where governments are in the political cycle, their value (and hence the incentive to allow logging) will depend on what alternative sources of rents governments have access to. Oil and gas reserves are highly unevenly distributed across Indonesia and the revenue sharing rules put in place by post-Soeharto governments, which give greater weight to the districts and provinces where these resources emanated from, mean that the distribution of revenue from these sources is also highly unequal. We exploit the variable availability of oil and gas revenues over time and space to examine whether they blunt or sharpen incentives to extract forest resources both immediately after these hydrocarbon resources become available and over the medium term. Consistent with other examples in the economics of corruption (Olken, 2007; Niehaus and Sukhtankar, 2009), we find that these two alternate sources of rents are substitutes in the short-run. In the medium term, however, we find that they become complements.

These results document that the incentives faced by local politicians and bureaucrats – the potential rents they can obtain from restricting logging vs. allowing more, the timing of rent extraction with regard to political needs, and the availability of alternative sources of rents – strongly affect patterns of deforestation in Indonesia. If optimal logging rules were being followed, none of these factors should matter. The fact that they do highlights the lack of full control central governments have over natural resources in developing countries, and suggests that incorporating the incentive compatibility constraints for local agents of the state is crucial to designing effective forestry policies.

The remainder of this paper is organized as follows. In the next section we discuss the background on how political change and deforestation in Indonesia and on how we study these processes using a variety of data sets. Section 2.3 examines how the splitting of districts affected deforestation, which we interpret in the light of a model of Cournot competition. In Section 2.4 we study the interaction between patterns of deforestation and the timing of elections. Section 2.5 investigates whether having access to alternative sources of public finance incentivises or disincentivises districts to engage in logging. Section 2.6 concludes.

2.2 Background and Data

Indonesia comprises an archipelago of islands in South-East Asia stretching from the Indian Ocean to the Pacific Ocean. It is a vast country. From tip-to-tip (from Sabang in Aceh to Merauke in Papua), Indonesia is 3,250 miles across; this is the same as the distance from Tampa, Florida to Juneau, Alaska. The conditions in Indonesia are ideal for the growth of forests and without the involvement of humans, Indonesia would be largely covered in forest.

In this section we first trace out the dramatic political changes that Indonesia has experienced in its recent past, and document how these change have resulted in a tug of war over the control of the forest sector. We then outline how we monitor forest loss using satellite data, and discuss how we capture political changes in our data. This section thus prepares the ground for the analysis of the political economy of deforestation which ensues in the subsequent three sections.

2.2.1 Background

2.2.1.1 Decentralisation in Post-Soeharto Indonesia The East Asian crisis brought to an end the thirty-two regime of President Soeharto on May 21st, 1998. He and his family had governed Indonesia as a personal fiefdom since 1967, and particularly in later years his New Order regime had become synonymous with the Soeharto family extracting rents from all key sources of economic activity in the country (Fisman, 2001).

Soeharto's departure ushered in one of the most radical reconfigurations of a modern state (Bertrand, 2008), combining a democratic transition with a radical decentralisation of power. Amidst fears that the multi-ethnic country would break apart, substantial administrative and fiscal authority was devolved to the approximately 300 district governments.[35] Off-Java regions which were rich in natural resources like forests, and oil and gas were particularly strident in their demands and wanted systems of control over these resources to be revised and for more of the revenue from their extraction to accrue to them (Cohen, 1998; Tadjoeddin et al., 2001; WB, 2003; Hofman and Kaiser, 2004; Wulan et al., 2004). The decentralisation laws (ROI, 1999a,b), which were passed in 1999 and took effect in 2001, devolved approximately 25% of the

[35] Unusually, Indonesian decentralisation transferred power to the approximately 300 district governments rather than the approximately 30 provincial governments, since districts, unlike provinces, were perceived to be too small for separatist tendencies (Hull, 1999; Niessen, 1999).

national budget to the districts in the form of block grants and dramatically increased their authority over almost all sectors of government. Local governments also received a substantial share of the natural resource royalties originating from their district.[36] Districts were administered by *Bupatis* (district heads), who were in turn indirectly selected by local legislatures.

The allure of self-government where districts could enjoy significant new political and fiscal powers led to a significant amount of district splitting. The total number of districts increased from 292 in 1998 to 483 in 2008. In contrast, the number of districts in Indonesia had remained largely unchanged during the New Order regime (1967-1999) (BPS, 2007). District splits thus represented a significant mechanism for the further decentralisation of power in the country (Cohen, 2003; Fitrani et al., 2005). What they also did, however, was to introduce a certain amount of disorganization as many districts lacked the human resources, technical capacities, and institutional structures to take on these new administrative powers (Tambunan, 2000).

Soon after decentralisation took effect, pressure mounted for a new reform, since it was felt that the 1999 Regional Governance law (ROI, 1999a) gave too much control to the local parliament and, thus, made the system susceptible to corruption (Mietzner, 2007) and elite capture (Erb and Sulistiyanto, 2009). Consequently, in 2004 a revised decentralisation law (ROI, 2004) considerably increased accountability by introducing direct election of the district head. Direct elections were to be held after the incumbent district head selected by the previous system had served their full tenure. This tenure, in turn, was dependent on when the terms of district heads appointed under Soeharto had to come to an end. This introduces asynchronicity in district elections.[37] Since the timing was driven by idiosyncratic factors from previous decades, it can be viewed as plausibly exogenous with respect to forest loss; indeed Skoufias et al. (2010) demonstrate that the timing of district elections is uncorrelated to virtually all pre-existing socioeconomic or geographic characteristics.

2.2.1.2 Implications for the Forest Sector During the Soeharto regime, the 1967 Basic Forestry Law (ROI, 1967) gave the national government the exclusive right of forest exploitation in the so-called Forest Estate (*Kawasan Hutan*); an area of 143 million hectares equivalent to three-quarters of the nation's territory (Barber

[36]In particular, an oil-producing district receives 6% of oil royalties and 12% of natural gas royalties; a further 6% (oil) and 12% (gas) is shared equally among all other districts in the same province. Districts are also allocated a large share of the one-off license fee for large-scale timber concessions (*Iuran Hak Pengusahaan Hutan* or *IHPH*) and of two types of volume-based royalties, the Reforestation Fund (*Dana Reboisasi* or *DR*) and Forest Resource Rent Provision (*Provisi Sumber Daya Hutan* or *PSDH*). Specifically, the producing and non-producing districts are each allocated 32% of the royalties. Furthermore, the district that contains the concession can keep 64% of the one-off fee with the rest going to the central government. Exceptions were made for the separatist provinces of Aceh (Special Autonomy Law 18 of 2001) and Papua (Special Autonomy Law 21 of 2001), who receive substantially larger shares. For a detailed discussion of Indonesia's transfer system refer to Brodjonegoro and Martinez-Vazquez (2002).

[37]For instance, only one-third of all districts (434) held direct elections in June 2005. By 2007, about 30% of all districts still had a district head that had not been elected directly.

and Churchill, 1987; Barber, 1990). The entire Forest Estate was managed by the central Ministry of Forestry, based in Jakarta. The Ministry in turn awarded a small group of forestry conglomerates (with close links to the regime's senior leadership) most of the timber extraction concessions in the Forest Estate, amounting to an area of about 69 million hectares inside the area designated as Production Forest (CIFOR, 2004). These exploitation rights were non-transferable, issued for up to 20 years and required the logging companies to manage the forest sustainably through selective logging (ROI, 1970). The second category inside the Forest Estate was the Conversion Forest, in which the largest wood producers could use Wood Utilization Permits (*Izin Pemanfaatan Kayu* or *IPK*) to clear-cut the forest and set up plantations for industrial timber, oil palm or other estate crops. Logging was prohibited in the remaining zones of the Forest Estate, which were designated for watershed (the Protection Forest) and biodiversity protection (the Conservation Forest).

The control over these forest zones, both in terms of issuing logging permits and enforcement, changed with the passing of the Regional Autonomy Laws in 1999. During the period from 1999-2002, district governments were legally allowed to issue a variety of small-scale, short-term forestry permits.[38] These licenses, both for the Production and Conversion Forest, often directly overlapped with the large-scale logging concessions and sometimes even the boundaries of national parks and protected areas.[39] In 2002, under pressure from the main forest concession holders, the national government revoked the right of district governments to issue these small-scale permits (ROI, 2002).

Even though after 2002 district governments were no longer able to issue logging permits, they still continued to play an important *de facto* role in forest management. In particular, the 1999 laws delegated enforcement of forestry laws to the districts, with some supervision from the central government, and these enforcement powers were not revoked in 2002. This means that, in most cases, the local forest office – the people who enforce logging quotas and who issue the log transport documents that accompany each log from source to destination – are directly or indirectly answerable to the head of the district.

Anecdotal evidence suggests that this move has drastically curtailed the enforcement capacity and has led to the sharp rise in illegal logging since the collapse of

[38]Government Regulation 62/1998 (ROI, 1998) gave the districts the authority to manage the so-called 'Privately Owned' and 'Community Forest' outside of the Forest Estate. Furthermore, they were given the authority to allocate small-scale Forest Product Extraction Licenses (*Hak Pemungutan Hasil Hutan* or *HPHH*) inside the Forest Estate. Ministerial Decree 310/1999 stipulated that districts governments could issue HPHHs inside the Conversion or Production Forest, but outside the boundaries of the large-scale concessions. However, Ministerial Decree 317/1999 established a procedure through with customary (*adat*) communities could obtain HPHHs for inside the boundaries of large-scale concessions as well.

[39]CIFOR and partners have conducted a series of district-level case studies to show the widespread proliferation of small-scale logging and forest conversion licenses. Prominent examples include Barr et al. (2001), Casson (2001b), McCarthy (2001), Obidzinski and Barr (2003), Samsu et al. (2004), and Yasmi et al. (2005).

the Soeharto regime (Casson and Obidzinski, 2002; Smith et al., 2003; Soetarto et al., 2003). District governments are believed to play an important role, particularly through the issuance of transport documents. Since even illegally harvested logs need a transport document (these are required at processing and export), and since these are usually issued by the district forest office, district governments have become in many ways the main gatekeepers for illegal logging. By controlling the amount of these transport documents issued, and by deciding when to turn a blind eye to enforcement and when to crack down, district governments can play a substantial *de facto* role in the rate of deforestation, even if their *de jure* role is limited. Estimates suggest that illegal logging makes up as much as 60-80% of total logging in Indonesia, making it a US$1 billion a year market (CIFOR, 2004). It is thus likely that these forces play a substantial role in determining the total amount of deforestation.

2.2.2 Data

2.2.2.1 Constructing the satellite dataset Given the prevalence of illegal logging, it is crucial to develop a measure of deforestation that encompasses both legal and illegal logging. To do so, we use data from the MODerate Resolution Imaging Spectroradiometer (MODIS) satellites to construct an annual measure of forest change for each year from 2001-2008. The resulting dataset traces, at 250m by 250m resolution, the patterns of deforestation across the entire country over time. This section describes how the forest change dataset is constructed from the raw satellite images.

There are two main challenges in constructing satellite-based images of deforestation. First, humid tropical regions like Indonesia have persistent cloud cover that shrouds the region year round. This makes it impossible to use high-spatial resolution sensors, like Landsat, which are usually used to measure forest cover change (Asner, 2001; Ju and Roy, 2008). Since these satellites typically only revisit the same area once every 1-2 weeks, cloud-free images are rarely recorded. Instead, it is necessary to draw on moderate-resolution sensors, such as the MODIS that pass over the same spot every 1-2 days. This considerably increases the likelihood of obtaining some good quality images, but at the cost of 250m by 250m resolution instead of the approximately 40m by 40m resolution available via Landsat. We start with the basic thirty-two day composites of the MODIS Land Surface Reflectance bands (Vermote et al., 2002) and the MODIS Land Surface Temperature Product (Wan et al., 2002) available on the NASA website, which aggregate daily images into monthly images to reduce cloud effects. We further aggregate them into annual composites to produce a cloud-free image of each pixel.

Second, one needs to take the composite of MODIS images and build a computer algorithm to discriminate between forest and non-forest. For each pixel, the MODIS satellite collects 36 "bands", each of which measures the strength of electromagnetic radiation in a particular part of the spectrum. That is, each pixel is essentially a

36-dimensional representation of the average electromagnetic radiation coming from a particular 250m by 250m spot. By contrast, the human eye, with its three types of cones, measures only three "bands", which correspond roughly to the blue, green, and red areas of the visual spectrum. The raw MODIS data is thus considerably richer than just a visual image at comparable resolution.

The key idea of remote sensing is developing an algorithm that identifies what signatures or set of signatures – i.e., what combinations of means and correlations among various parts of the 36-dimensions of spectrum that MODIS sees – best discriminate between forest and non-forest. For example, plants absorb electromagnetic radiation in the red visual range for use in photosynthesis, but reflect or scatter radiation in the near-infrared range. One common metric therefore examines the so-called NDVI (Normalized Difference Vegetation Index), which captures the difference in intensity between light in the red range and in the near-infrared range, and therefore identifies one signature for plant life (Gausman, 1977; Tucker, 1979; Curran, 1980).

In practice, one can do much better than using NDVI by exploiting additional dimensions of the data (see Wulder (1998) for a literature review). For example, forests tend to be cooler than surrounding areas, so bands that measure temperature can also be used (Gholz, 1982). Moreover, trees have different spectral signatures than other types of crops and plants (Curran, 1980). To take maximal advantage of the richness of the MODIS data, we use a statistical learning procedure known as a "tree bagging algorithm" to determine which spectral signatures best correspond to forest (Breiman et al., 1984; Breiman, 1996).

Specifically, we start with much higher resolution "training" images. For each of these images, experts have manually examined the image and coded each cell into forest, non-forest, or forest change (deforestation). We then apply the statistical tree-bagging algorithm to automatically group the MODIS data into naturally occurring groups that share common electromagnetic signatures, and then determine which of these sets of signatures corresponds to the manually-coded forest, non-forest, or forest change cells in the training dataset. This is akin to a regression, except that it allows for complex correlations between bands to be used in the prediction, rather than just means, and allows very flexible functional forms.

One then can extrapolate over the entire MODIS dataset to predict, for each year, the probability that a given pixel was deforested. We code a pixel as deforested if the probability exceeds 90% in any year; once it is deforested, we consider it deforested forever. The reason for this is that, especially in a humid tropical environment like Indonesia, once the original forest is cleared other crops or scrub brush emerges quickly; since the forest takes at least several decades to regrow, this regrowth is not actual tree cover. Deforestation thus is often represented by a pixel that is "green" one year, "brown" the next year, and then "green" again. Given this, Hansen et al. (2009) have shown that the key to detecting true forest change is the high probability of being deforested in a single year, rather than appearing "brown" year after year.[40]

The final outputs are annual forest change estimates for 2001-2008 for each of the 34.6 million pixels that make up Indonesia. Note that these estimates will provide a lower bound for forest change, as a 250m by 250m pixel is only coded as deforested if the majority of the area represented by the pixel is felled. This will reliably pick up clear-cutting, but will not necessarily capture selective logging if the forest canopy remains largely intact, and therefore may under-estimate total logging. It will also capture deforestation due to large-scale burns, which can be either intentional (for land clearing) or unintentional.[41] This cell-level data is then summed by district and forest zone (i.e., the four forest categories in the Forest Estate: the Production, Conversion, Protection and Conservation Forest). This yields our final left-hand-side variable $deforest_{dzt}$, which counts the number of cells deforested in district d in forest zone z and year t.

Figure 2.1 gives an idea of what our underlying forest cover data looks like. To do this we zoom in onto a small area, since the detailed nature of this dataset makes it impossible to visualize the 34.6 million pixels that make up Indonesia in a single map. It focuses on one of the main hotspots of deforestation during this time period (Hansen et al., 2009), namely the province of Riau on the island of Sumatra. The deforested cells are indicated in red, forest cover is shown in green and non-forest cover in yellow. The map clearly shows that substantial amounts of forest have been deforested during the period from 2001 to 2008. Furthermore, forest clearing seems to spread out from initial areas of logging, as access will be easier from already logged plots.

In addition to the satellite data, we also examine official logging statistics from the annual 'Statistics of Forest and Concession Estate' (*Statistik Perusahaan Hak Pengusahaan Hutan*), published by the Indonesian Central Bureau of Statistics for 1994-2007. These statistics report the quantity of logs cut at the province level and the associated price by wood type, for 114 different types of wood.[42] Because they are derived from production data, they include both clear-felling as well as selective logging; on the other hand, they capture only logging that was officially reported by the forest concessions, and so likely miss most illegal logging. Since they report the wood cut from the legal part of the Forest Estate, they should be compared to the satellite data from the Production Forest. This data also includes data on the price of woods; since market prices are determined by both legal and illegal logging, these prices will reflect the market equilibrium for both types. We use this second dataset as a consistency check for our satellite data and to examine impacts on prices, as

[40]Hansen et al. (2009) have compared the 90% thresholds to other thresholds and found that, in verification, 90% is reasonable.

[41]However, we show in Section 2.3 below that we obtain remarkably similar results in the Production Forest for the satellite-based deforestation measure and official logging statistics, suggesting that much of what we are picking up is, indeed, logging.

[42]We drop the 'other' (*Lainnya)* and 'mixed wood' (*Rimba Campuran)* category, since their composition varies considerably across provinces and over time.

described in further detail in Section 2.3 below.

2.2.2.2 Descriptive statistics of forest change Figure 2.2a illustrates the distribution of district-level logging across Indonesia over time. In particular, it shows the number of cells cleared at the district level in 2001 and 2008. We focus our analysis on the main forest islands of Indonesia: moving from West to East, these are Sumatra, Kalimantan, Sulawesi and Papua. The remaining islands (Java, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, and Maluku), shown in white, have negligible forest cover in the baseline period and are not included in our sample. In this map, low levels of logging are shaded in green, whereas high levels of forest clearing are indicated in orange and red. The figures suggest that most of the deforestation occurs in Kalimantan and in the lowlands of Sumatra along its eastern coast. From 2001 to 2008, there is a shift in deforestation in Kalimantan from the West to the East, and there is an intensification in deforestation in Sumatra, particularly in the provinces of Riau and Jambi in the east-centre of the island. There is also some intensive logging in the Southern part of Papua in 2001, but high clearing rates are not maintained in this area over time.

Table 2.1 reports the trends in forest cover over time in more detail, and Table 2.2 displays the summary statistics for our main measure of deforestation. The data in both tables is reported for the entire Forest Estate, the subcategories of the Forest Estate where logging may be legal (Production/Conversion Forest) and where all logging is illegal (Conservation/Protection Forest) as well as the individual subcategories of the Forest Estate. Table 2.1 shows the changes in the forest area measured in 1,000 hectares for our forest island sample. Total deforestation between 2000 and 2008 amounts to 4.894 million hectares or 48,940 square kilometres; this is roughly twice the size of Vermont.

Most of this change occurs in the Production Forest, where about 3.02 million hectares were deforested. Much smaller changes are reported for the other forest zones: 1.12 million hectares were deforested in the Conversion Forest and only 0.73 million hectares were deforested in the Conservation and Protection Forest combined. However, this last estimate will only provide a lower bound of the actual changes on the ground, since logging is prohibited in these parts of the Forest Estate. To the extent illegal logging is selective and, thus, occurs on a much smaller scale, moderate resolution sensors like MODIS will underestimate these changes.

Table 2.2 shows the summary statistics of our main left-hand side variable, $deforest_{dzt}$, which counts the number of cells cleared for district d in forest zone z and year t. To facilitate comparison with Table 2.1, the measurement of this variable has been converted into 1000 hectares. On average, 704 hectares are deforested annually at the district level. However, the variance of the 2,901 hectares suggests that there is a lot of variability in deforestation both across years and districts. The pattern of the results mimics the previous findings, i.e. most of the changes occur in the Production

Forest, where on average 1,451 hectares are deforested in each district and year.

2.2.2.3 Political Economy Data To capture increasing competition in the wood market, we take advantage of the extensive partitioning of districts following the collapse of the New Order regime. Figure 2.3 illustrates the distribution of district splits in our forest island sample. It displays the total number of districts that the original 1990 district partitioned into by 2008. High numbers of splits (3-7) are denoted by orange and red in the figure, whereas low numbers of splits (0-2) are denoted by blue and green. It is evident from this map that district splits happen all over the country. Most districts split at least once or twice, so that very few of the 1990 districts remain intact. In addition, the map suggests that the largest districts in 1990 split into more new administrative units.

We construct two sets of variables for the districts and provinces using the official publications on regency and municipality codes of Statistics Indonesia (*Badan Pusat Statistik* or *BPS*).[43] Note that we use the 1990 boundaries as a reference point, because 17 new districts were formed between 1990 and 1999 (BPS, 2007).[44] For the province-level data, we simply calculate the total number districts and municipalities within the 1990 boundaries of province p on island i in year t, $NumDistrictsInProv_{pit}$.[45] In addition, we construct two more variables at the district level. Firstly, we count into how many districts and municipalities the original 1990 district d on island i split in a year t, $NumOwnDistricts_{dit}$. Secondly, we sum across all the other districts within the same province, $NumOtherDistricts_{dit}$.

We also obtain other district-level covariates as follows. To examine the impact of political election cycles, we obtain district-level election schedules obtained from the Centre for Electoral Reform (CETRO)[46], and use them to construct a dummy for the year the election for district head was held, $Election_{dit}$. To examine the impact of other sources of rents available to district governments, we examine oil and gas revenues per capita at the district level, $PCOilandGas_{dt}$.[47] We obtained the revenue data from the Indonesian Ministry of Finance (*Menteri Keuangan*) webpage (`http://www.djpk.depkeu.go.id/datadjpk/57/`) and the population data for 2008, which is published by the Indonesian Central Bureau of Statistics. It is important to note that new districts often do not record their own share of revenue for the first few

[43]The most up-to-date lists of regency and municipality codes is available on the BPS webpage at `http://dds.bps.go.id/eng/aboutus.php?mstkab=1`.

[44]During the Soeharto regime, only 3 new *kabupaten* or *kota* were created outside of Jakarta prior to 1990: Kota Ambon (PPRI No. 13 Thn. 1979), Kota Batam (PPRI No. 34. Thn. 1983), and Kab. Aceh Tenggara (UURI NO. 4 Thn. 1984). Jakarta itself was split into 5 city parts in 1978.

[45]Each province is located on only one of the four islands – Sumatra, Kalimantan, Sulawesi, and Papua. We use the island subscript, i, as we will allow for differential time trends by island in the empirical analysis below.

[46]CETRO is an Indonesian NGO (`http://www.cetro.or.id/newweb/index.php`). We use the most up-to-date district-level election schedule available, which provides election dates up to 2011.

[47]Oil and gas is by far the largest source of natural resource rents for districts. For instance, in 2008 the average district-level revenue from oil and gas was 114.515 billion rupiah, whereas the corresponding figure for forestry was 5.302 billion rupiah.

years after the split, as the district is not fully functioning yet. We therefore allocate each new district the revenue share of its originating district until it reports its own share of revenue for the first time.

Figure 2.4 displays oil and gas revenue per capita in 2008 at the district-level. These natural resources are much more spatially concentrated than forest, so that most districts receive none or very little revenue shown as blue and green respectively. The districts that receive the largest share of revenue from oil and gas extraction are located in Eastern Kalimantan and in the province of Riau on Sumatra. Moreover, the map shows that there is some heterogeneity across districts within each province, where provinces are delineated with thick black borders. These differences are due to the revised revenue sharing rules, where the producing and non-producing districts each receive the same percentage of oil and gas revenue, which is then split evenly between the districts in each category (ROI, 1999b). Since the non-producing districts are usually larger in number, their final share of revenue will be smaller.

2.3 Cournot competition between districts

2.3.1 Theory

Although there is a large literature on optimal forest management, the forestry literature tends to consider how an optimal central planner should manage forest resources, trading off the growth rate of trees with discounting (e.g., Samuelson (1976), Dasgupta (1982); see Brown (2000) for a survey).[48] In this section, we consider what happens if, rather than a central planner making optimal forest extraction decisions, forest decisions are made by individual actors – in our case, district governments – who interact with one another through a common product market.

For simplicity, in this section we abstract away from issues involved in tree regrowth and instead treat forests as an exhaustible natural resource. This is consistent with substantial *de facto* logging practice in many tropical forests, including those in Indonesia, where virgin forests are heavily logged, and then either left in a degraded state or converted to a non-forest use, such as palm plantations. This type of non-sustainable clear-cutting and land conversion is also the type of forestry we will primarily be able to observe in the satellite data.[49]

We suppose that each period, district governments choose the quantity of forest to extract. As discussed above, this can occur in a variety of ways: by determining how many illegal log transport permits to issue, how many conversion permits to issue, etc. Once they determine quantities, prices are determined through the market. We

[48]The other strand of the literature considers multiple actors with competing property rights over the same forest (e.g. Larson and Bromley (1990), Ligon and Narain (1999)). However, to the best of our knowledge none consider the type of oligopolistic competition we study here, where each actor has full control rights over its own forest and strategic interactions occur through the product market.

[49]One could generalize the model to allow forests to regrow at some slow rate; we speculate that this would not substantially affect the qualitative predictions we consider here, which concern the strategic interactions between districts.

assume that transport costs across different parts of Indonesia, the need to process logs locally before export (Indonesia bans the export of raw, unprocessed logs), and capacity constraints at local sawmills combine to generate local downward-sloping demand curves for logs in each market; this assumption is discussed in more detail below.

The problem districts face is thus that of oligopolistic competition in a non-renewable natural resource. Lewis and Schmalensee (1980) show that many of the standard, static Cournot results generalize to this setting. In particular, they show that a greater number of actors in a market – in our case, more districts – leads to lower prices and greater resource extraction.[50] We will test this implication in the empirical section below.

Lewis and Schmalensee (1980) also discuss what happens if inventories or marginal costs of extraction vary among producers. When marginal costs are unequal, they demonstrate that it is possible that the low-cost firm will dominate extraction first, with the high-cost firm increasing its extraction later when it faces less competition in the product market from the low-cost firm. In our empirical setting, when districts split, the marginal costs of issuing illegal logging permits may be higher initially, as the new district government needs to spend time organizing a new district government, implement new regulations, learning how to evade the national monitoring system, etc. Only after several years when the district government is fully operational will the marginal cost of extraction return to where it was before. Under such a framework, the Lewis and Schmalensee model suggests that the other districts in the market will raise production immediately, knowing that they face more competition from other districts in the future. Prices will fall somewhat initially, with a greater fall in prices later as the marginal costs for the new district declines. We examine these predictions in the empirical work below.

2.3.2 Empirical Tests

To examine the impact of Cournot competition between districts, we will take advantage of the fact that the number of districts has increased dramatically over the period we study. As discussed above, across all of Indonesia, the number of districts increased from 292 prior to decentralisation to 483 in 2008. The increase is even more dramatic in the forest islands (Sumatra, Kalimantan, Sulawesi, and Papua) that are the focus of this study – from 146 districts prior to decentralisation to 311 districts in 2008, an increase of 213%. We exploit the staggered timing of these changes in administrative boundaries to identify the relationship between the number of administrative units and logging and to test the theoretical model outlined above.

[50] Because the resource is subsequently depleted more quickly with more actors, they also show that the price rises more quickly with higher N than with lower N as the resource moves more quickly towards exhaustion. In our case, since the rate of extraction is small relative to the reserves (e.g., about 0.5% per year), the increase in prices may happen too slowly to be observed in our data.

As analyzed in detail in Fitrani et al. (2005), the splitting of districts was driven by three principal factors: geographic area, ethnic clustering, and the size of the government sector.[51] From the perspective of this paper, the key question is not whether a district splits, but rather the timing of the split. Several idiosyncratic factors appear to influence the timing. First, the process of splitting a district is quite cumbersome, involving a number of preliminary steps (e.g., formal agreement of the district legislature, the district head, the provincial governor, and the provincial legislature; documentation of the new districts' ability to meet fiscal requirements; documenting a reason for the split (ROI, 2004), and, ultimately, the passage of a special law by the national parliament for each split that will take place). The amount of time each of these steps take varies, which in turn influences the total amount of time required. Moreover, there was a national moratorium on splits from 2004 (when the criteria for splits were revised) through 2007. This moratorium also creates plausibly exogenous delays in timing of splits, as many districts that may have been close to completing the process in 2004 had their split postponed by three years due to the moratorium.[52] In the empirical analysis below, we test for whether the timing of these splits is associated with trends in deforestation, though *a priori* there is little reason to believe they would be.

To test the predictions of the theory, a key question is what definition we should use for the "market" of wood products. While wood and wood products are traded on international markets (and, hence, one would expect the market to be global), there are several factors that make wood markets in Indonesia more local. In particular, since 2001 Indonesia has banned the export of raw logs. Instead, all timber felled in Indonesia must first be transported (either by river, when possible, or by road) to local saw mills, plywood mills, and paper mills, where it is processed before export. Capacity constraints at these mills thus determine local wood prices, and non-trivial transportation costs for raw logs imply that prices may differ across regions. We focus on the province as the key definition of a market, since provincial boundaries are coincident with the major river watersheds used for transporting logs.

We will examine several empirical predictions of the Cournot theory outlined above. First, taking a province as a measure of the market, we use panel data to test whether the number of districts in the province affects the prices and quantity

[51]Specifically, the Soeharto era districts were often quite large, so naturally they find that districts that were larger geographically are more likely to split to make administration easier. Second, there are often ethnic tensions in Indonesia, particularly off Java. Those districts where the different ethnic groups were clustered geographically were more likely to split. Finally, the block grant fiscal transfer had a fixed-component per district. While this gives all districts an incentive to split, they find that it is particularly likely in those districts with a large wage bill, which presumably are in greater need of the revenue. They find little consistent relationship between natural resources and splitting, with positive coefficients in the 1998-2000 period and negative coefficients in the 2001-2003 period, implying a zero effect on average. Details of these regressions can be found Fitrani et al. (2005).

[52]Unfortunately, we do not observe when the district began the process of filing for a split, as we only observe the date the final split law was passed by the Parliament, so we cannot exploit this three-year moratorium directly as an instrument.

of wood felled in the province. For this purpose, we will use our two complementary sources of forestry data; namely, the price and quantity records from the official government publications and the new satellite-based forestry dataset. For the price and quantity data, we will run an OLS fixed effects regression, as follows:

$$\log(y_{wipt}) = \beta NumDistrictsInProv_{pit} + \mu_{wpi} + \eta_{wit} + \varepsilon_{wipt}, \tag{4}$$

where y_{wipt} is the price or the quantity of wood type w harvested in province p (located on island i) and year t, and $NumDistrictsInProv_{pit}$ counts the total number of districts in province p in year t.[53] The regression also controls for wood-type-by-province and wood-type-by-island-by-year fixed effects, μ_{wp} and η_{wit} respectively. The robust standard errors are clustered at the 1990 province boundaries. Since there is a substantial variation in quantity of wood across wood species and provinces – the 5th percentile of the quantity variable is 42 m^3, whereas the 95th percentile of the quantity variable is 204,804 m^3 – this regression is weighted by the volume of production of wood type w in province p in the first year that we have data.

For the satellite-based forestry data, we will run a fixed-effects Poisson Quasi-Maximum Likelihood regression (Hausman et al. (1984), Wooldridge (1999); see also Wooldridge (2002)), with robust standard errors clustered at the 1990 province boundaries. Specifically, this estimates, by MLE, equations such that

$$\mathbf{E}\left(deforest_{pit}\right) = \mu_{pi} \exp\left(\beta NumDistrictsInProv_{pit} + \eta_{it}\right) \tag{5}$$

where, as in equation (4), μ_{pi} is a province fixed-effect, and η_{it} is an island-by-year fixed effect. In estimating (5), the Poisson QMLE model effectively weights each observation by $deforest_{pit}$. Thus, if one takes logs of equation (5), the coefficient β in equation (5) is directly comparable to the coefficient β in equation (4); both represent the semi-elasticity of deforestation with respect to the number of districts in the province, where observations are weighted by quantities.[54] The reason we use the Poisson specification for the satellite data, rather than estimate a log dependent variable with OLS, is that we use the satellite data's identification of the precise location of logging to separately estimate equation (5) for each land use zone. This allows us to differentiate between impacts in land use zones where logging may be legal as well as illegal (the Production and Conversion Forest) vs. the land use where all logging is illegal (the Conservation and Protection Forest). When we split the forestry data up by zones, we have many observations where the dependent variable is 0, so the Poisson model is more appropriate.

[53] As discussed above, there are four islands in our sample: Sumatra, Kalimantan, Sulawesi, and Papua. Each province is located on only one island.

[54] The only difference is that equation (4) is weighted by initial volumes in production ($logging_{wp0}$), whereas the Poisson model implicitly uses contemporaneous volumes for weights ($logging_{wpt}$) (see VerHoef and Boveng (2007)). We show below that using contemporaneous weights when estimating equation (4) produces virtually identical results.

Second, we will examine the impact of splits at the district level. In particular, we will test whether splits affect deforestation in the district that splits vs. how it affects deforestation in the remainder of the province. We estimate via Poisson QMLE a model such that:

$$\mathbf{E}\left(deforest_{dit}\right) = \mu_{di} \exp(\beta NumOwnDistricts_{dit} + \gamma NumOtherDistricts_{dit} + \eta_{it}) \tag{6}$$

where $deforest_{dit}$ is the number of cells cleared in district d (located on island i) between year $t-1$ and t, $NumOwnDistricts_{dit}$ counts into how many districts the original 1990 district d split into by year t, and $NumOtherDistricts_{dit}$ counts into how many other districts there are within the same province in year t. It also includes district fixed effects μ_{di} and island-by-year fixed effects η_{it}. The robust standard errors are now clustered at the 1990 district boundaries. The Poisson QMLE is again estimated separately by land use zones.

2.3.3 Results using official production statistics

2.3.3.1 Results Table 2.3 begins by estimating equation (4), using the data on prices and quantities from the official forest concession reports. Columns 1 and 2 provide the estimates for our main specification, which includes all wood types and covers the period 2001-2007.[55] Columns 3 and 4 show the results for the same sample period, but restrict attention to a balanced panel of wood types, where we observe production of the wood type in all years for a given province. Columns 5 and 6 present the results for all wood types for a longer time horizon that also includes the years of the pre-decentralisation period for which the official logging publications were also available, i.e. for 1994-2007. Panel A displays the estimates for the contemporaneous effect, and Panel B estimates the medium-run impact by including 3 lags of the number of districts variable in equation (4). Columns 1, 3, and 5 present equations where the natural log of prices are the dependent variables, and columns 2, 4, and 6 present equations where the natural log of quantities are the dependent variables.

Consistent with the theory, the main results in columns 1 and 2 of Panel A show that adding one additional district in a province decreases prices by 1.7% and increases the quantity of logs felled by 8.9%. Similar results are obtained for the alternative samples shown in columns 3 through 6, which are not statistically significantly different from the main specification.

Since increasing the number of districts is essentially a supply shock, one can infer the slope of the demand curve from the ratio of $dLnQuantity$ to $dLnPrice$. Combining the estimates from columns 1 and 2 implies a demand elasticity of -5.24. Given that markets are separated only by transportation costs, we would expect that demand for forest products should be quite elastic, consistent with the high elasticity we find in the

[55]Data is not yet available for 2008, so this is the most comparable time period to that used in the satellite data analysis below.

data. Moreover, the Cournot theory suggests that the equilibrium elasticity is related to the number of market participants, so with an average of 13 districts per province by the end of our sample we would also expect a high elasticity in equilibrium.[56]

Panel B estimates the medium-run impact of the number of districts on prices and quantities by including 3 lags of the $NumDistrictsInProv_{pit}$ variable.[57] The medium-run impact estimated by calculating the sum of the immediate effect and all 3 lags is even larger than in the main specification: at the end of 3 years prices have fallen by 3.29% and quantities increased by 13.1%. The estimated coefficients also become even more precisely estimated. The demand elasticity drops slightly to -3.89. As in Panel A, the estimates for the balanced panel and the longer time horizon are again similar in magnitude and significance. The higher long-run response suggests that it takes some time for the market players to fully adapt to the new, higher level of logging; we will explore more precisely the reasons for this in the district-level analysis below.

2.3.3.2 Alternative specifications Table 2.4 examines a series of alternate specifications. Panel A of Table 2.4 reestimates Panel A of Table 2.3, but instead of weighting by the quantity in the first year, we instead weight by current quantities. This weighting is most similar to the one applied by the Poisson Quasi-Maximum Likelihood estimator to be used in the analysis of the satellite data below (see footnote 54). The results do not change substantially and are similar in magnitude to Panel A in Table 2.3. The standard errors become slightly larger, which marginally lowers the significance of the results.

Panel B and C replicate the results in Panel A of Table 2.3 using different measures to count the number of districts within the province. In Panel B, we exclude from the district count *kotamadya* (major cities), which do not control any forest and hence should not affect logging. The results are virtually identical to the main results in Panel A of Table 2.3. Conversely, Panel C offers a falsification exercise by *only* counting *kotamadya* within a province in calculating the number of districts. As one would expect (since *kotamadya* do not control forests), we essentially only find noise, as all of the coefficients apart from column 6 are insignificant.

Finally, to test for differential trends in provinces with more splitting, Panel D includes three leads of the district variable as well as three lags. The observation count drops slightly, since we lose 2007 (since we do not yet know the number of districts in 2010, and hence can't calculate the third lead for 2007). The medium-run impact of district splits on prices and quantities is robust to the inclusion of leads and is similar in magnitude and significance to Table 2.3. For our main specification

[56]Specifically, a very simple Cournot model with zero marginal cost predicts that the elasticity of demand in equilibrium should be equal to $-N$, where N is the number of districts in the province. Since the number of districts with forests averages 13 per province, one would also expect from the theory that the demand elasticity should be quite high in equilibrium.

[57]The results do not change substantially if we use five lags instead.

(columns 1 and 2), both the sum of the leads and the p-value from a joint F-test of all three leads together are statistically insignificant, indicating that there are no pre-trends. The only case where we find the sum of the leads to be significant is in column 5, which suggests that there may be differential trends in prices in the pre-decentralisation period (i.e., outside the main period we study). The p-value from the F-test of all of the leads is also jointly significant at 10% for both the quantity and price data in the balanced panel (columns 3 and 4). While there is scattered evidence of significant effects on the leads in alternate specifications, in the main time period and specification we examine – 2001 through 2007 – we find no evidence of significant pre-period differential trends.

2.3.4 Results using the satellite data at the province level

The results in the previous section were from official forestry statistics, and thus may not fully capture illegal logging.[58] To gain a fuller measure of the impact on logging, we turn now to the satellite-based deforestation dataset, described in Section 2.2.2.1.

The main results are displayed in Table 2.5, which reports the findings separately for each subcategory of the Forest Estate. Column 1 presents all categories of the Forest Estate together, column 2 presents results for the zones where legal logging can take place (i.e., the Production and Conversion Forest), and column 3 presents results for the zones where no legal logging can take place (i.e., the Conservation and Protection Forest).[59] Columns 4-7 report the estimates for each zone individually.

The total estimated impact of district splits on deforestation is shown in column 1 of Panel A. We find that deforestation increases by 3.61% if an additional district is formed within a province.

The results reported in columns 2 and 5 show that the overall effect is largely driven by logging in the Production Forest. In fact, deforestation increases by 5.33% for each additional district in the province. In addition, it is important to notice that the quantities data from official forest statistics analyzed in Section 2.3.3 above refers to logs harvested inside the Production Forest, so in theory the coefficients in Table 2.4 (Panel A, column 2) and Table 2.5 (Panel A, column 5) should be comparable. In fact, when comparing the estimated coefficients, it becomes clear that they are not statistically significantly different from each other. This finding enhances our confidence in the satellite-based forestry data, as it yields results that closely match those from an independently collected data source.

[58] While one would expect that the official statistics would not capture any illegal logging whatsoever, this may not be the case. Each log needs a transportation permit (*Surat Keterangan Sahnya Hasil Hutan* or *SKSHH*) to be processed, which is signed at the origin and destination of the log. If concession wanted to transport illegally felled timber, they can obtain excess (illegal) transportation permits to do so (Casson et al., 2006). Once it obtains a transportation permit, the log is quasi-legal, so concessions may have included these logs in their official logging statistics.

[59] As discussed above, since the Poisson model weights each observation by the quantity, when we combine observations from multiple zones we obtain the correct weighted average effect.

One great advantage of the new remotely sensed forestry data is that it provides a comprehensive picture of the changes on the ground beyond the legal production areas. This is clearly illustrated by the finding in column 6, which suggests that district splits do not only impact legal logging in production areas, but also activity in the rest of the Forest Estate. In particular, one more district increases deforestation in the Conservation Forest – i.e., national parks – by 7.86%. Given that illegal logging is often selective and therefore difficult to capture with moderate resolution sensors like MODIS, this will only present a lower bound estimate for the actual extent of deforestation in the Conservation Forest.

Panel B reports the estimates of the medium-run impact of district splits. As in the official logging statistics analyzed above, the estimated coefficients for the sum of the contemporaneous effect and the lags are generally larger and more significant. For instance, a district split increase deforestation in the entire Forest Estate by 7.89%. The estimates for deforestation in legal and illegal logging zones, reported in columns 2 and 3 respectively, are now both significant and of similar magnitude. Logging in the Production Forest still seems to be one of the main drivers of the overall effect, as one additional district increases deforestation in this zone by 7.93% at the end of 3 years. Deforestation in the Conservation Forest increases in the medium-run to 12.5%.

Table 2.6 test for the presence of differential trends in the data by including three leads of the $NumDistrictsInProv_{pit}$ variable. We find that the our main results are robust to the inclusion of leads. Furthermore, and most importantly, the sum of the leads is insignificant across all forest zones. In addition, the p-value of the joint significance test for the leads is large and statistically insignificant for all samples. This is reassuring, as it is suggests that the results are indeed picking up the causal impact of district splits on both legal and illegal logging in the Forest Estate and are not being driven by unobserved trends.

2.3.5 Results for the satellite data at the district level

Using the satellite data, we can disaggregate logging by district as well as forest zone. This allows us separately to estimate the direct effect of a district splitting – i.e., the impact in the district that splits itself – from the indirect effect of the split – i.e., the impact on logging of other districts in the same province. While the direct effects may capture a variety of phenomena beyond the Cournot market forces described above, such as changes in enforcement regimes as the district forestry office is reorganized, the capital city moved, and new forestry offices are constituted, the indirect effect of a split on other districts in the province occurs only through market forces.

The results from estimating equation (6) are shown in Table 2.7, and paint a starkly different picture for direct and indirect effects of district splits. For direct effects the overall impact effect shown in Panel A is negative (though insignificant). This is driven by substantial decline in deforestation in the Production Forest – a decline of 21.1%. On the other hand, there appears to be an increase in illegal logging when the district splits – deforestation in the Conservation Forest (i.e., national parks) increases by 13.6%. Combined, these estimates suggest that in the year of the split, there may be turmoil in the local forestry offices, leading to a decline in legal logging (e.g., it may become temporarily harder to obtain log transport and cutting permits) and an increase in illegal logging (e.g., enforcement capacity declines).

Panel B shows, however, that the pattern of these direct effects begins to change over time. By the time the district has been in existence for three years, deforestation in legal logging zones begins to increase, partially offsetting the initial declines, so that the third lag on the number of district splits is positive and statistically significant. While the net effect (the sum of the lags) is not distinguishable from zero, the p-value on a joint test of the contemporary effect and all 3 lags in the legal logging zones (column 2) is < 0.01, suggesting that the pattern we observe – a decline in deforestation initially, followed by an increase – is indeed highly statistically significant. Meanwhile, deforestation in illegal logging zones continues to intensify, so that the net effect in illegal logging zones is an increase of 25.1% (Panel A, column 3, sum of lags), driven by a 37.0% increase in the Conservation Forest (column 6) and a 13.3% increase in the Protection Forest (column 7). On net, the total increase in deforestation after 3 years (shown in column 1) is 3.23%, though this is not statistically significant.

For indirect effects, by contrast, the impact on deforestation is positive and immediate, and is concentrated in the legal logging zones. The impact effect of a district splitting is to increase overall logging by 7.01% in all other districts in the province (Panel A, column 1); the medium-run impact is 9.51% (Panel B, column 1, sum of lags). There are no statistically significant impacts in illegal logging zones outside of the district that splits.

Table 2.8 includes three leads of the $NumOwnDistricts_{dit}$ and $NumOtherDistricts_{dit}$ variable in the regression. The main results are robust to the inclusion of the leads and we do not find a significant sum of leads for the $NumOtherDistricts_{dit}$ variable. In almost all specifications in Table 2.8, we do not find statistically significant effects on either the sum of the leads, or on the joint test of significance of all leads. The only exceptions are the sum of the leads for own splits in the Conservation Forest (column 6) and for other splits in the Conversion Forest, but given that we find significance in only 3 out of the 28 lead tests we consider it is likely that these are just noise, rather than true differential trends.[60]

The difference between the direct and indirect effects of a new district forming suggests a consistent explanation for the results in this section along the following lines. When a district splits, the initial disorganization disrupts legal logging activities but leads to an increase in illegal deforestation. Other districts within the same province increase logging immediately, as they know that prices will fall once the new districts

[60]There are 28 tests as follows: for each of the 7 columns, we examine both the sum of the leads and a joint F-test of significance of all leads for both own splits and other splits.

are fully established and begin to log more. Combined, this leads to a relatively small fall in prices and increase in quantities immediately – the 1.70% price drop and 3.61% logging increase shown in Panel A of Tables 2.3.3 and 2.5 – and a larger fall in prices and increase in quantities in the medium run – the 3.29% price drop and 7.89% logging increase shown in Panel B of Tables 2.3.3 and 2.5.

2.4 Political logging cycles

2.4.1 Empirical tests

The literature on political business cycles suggests that politicians tend to increase expenditures and postpone tax increases in the years leading up to elections, both at the national level (e.g., Nordhaus (1975), MacRae (1977), Alesina (1987), Rogoff and Sibert (1988), Akhmedov and Zhuravskaya (2004)) and at the local level (e.g., Poterba (1994), Besley and Case (1995), Levitt (1997), Finkelstein (2009)). This section examines whether political cycles affect not only the legal actions by the state, but the state's permissiveness towards illegal activity. In particular, we examine whether logging in general, and illegal logging in particular, increases in the years leading up to a district election.

To do so, we take advantage of the fact that the timing of district-level elections in Indonesia varies from district to district. As discussed in Section 2.2.1.1, prior to 2005, the heads of districts (known as *Bupati*) were indirectly selected by the district parliament. Starting in 2005, *Bupatis* were to be directly elected by the population in special elections (ROI, 2004). Crucially, the direct elections of *Bupatis* were phased in as the prior *Bupati*'s term expired, so that some districts had their first direct elections as early as 2005 while others had them as late as 2010.[61] As documented in detail by Skoufias et al. (2010), the timing of these direct elections was determined exclusively by when a *Bupati*'s term expired, which was in turn driven by idiosyncratic factors, such as retirements and appointments of existing *Bupatis* to other posts, that determined *Bupati* appointments under the pre-1998 Soeharto regime (Emmerson, 1999). Skoufias et al. (2010) examine this empirically and verify that the resulting timing of local elections is uncorrelated with a host of economic, social, and geographic characteristics.[62]

[61] No direct elections for *Bupatis* were held in 2009, as national Presidential elections were held that year. Those *Bupatis* whose term was ending in 2009 were extended on an interim basis and direct elections were held in 2010 instead.

[62] Specifically, Skoufias et al. (2010) run a regression of the probability of holding a direct election by 2007 and regress it on the end date of the previous *Bupati*'s term and the following variables: unemployment rate, log real per capita district GDP, log real per capita district GDP without oil and gas, share of minerals in the district GDP, share of energy in the district GDP, dummy for a district having oil and gas, share of population that is urban, share of asphalt roads in the district, share of rock roads in the district, access to telephones, distance to the provincial capital, dummy for being a split district, share of mountainous areas in the district, share of coastal areas for the district, share of valley areas in the district, a city dummy, and 5 island dummies. Other than the end date of the previous *Bupati*'s term, only 1 of the 21 variables they consider (a Sulawesi island dummy) is statistically significant at the 10% level. See Table A-1 of Skoufias et al. (2010).

To estimate the impact of elections on logging, we use the satellite data and estimate fixed-effects Poisson QMLE models on the various subcategories of the Forest Estate that estimates the following equation:

$$\mathbf{E}\left(deforest_{dit}\right) = \mu_{di} \exp\left(\sum_{j=t-2}^{t+2} \beta_j Election_{dij} + \eta_{it}\right) \tag{7}$$

where j indexes leads and lags of the *Election* variable, which is a dummy for a *Bupati* election taking place. As in equation (6) above, we include district fixed effects and island-by-year fixed effects, and cluster standard errors at the 1990 district level. We include up to 2 leads and 2 lags of the *Election* variable to fully capture the 5 year election cycle.[63] Note that since the official forestry statistics are only at the province level, whereas our variation is in the timing of elections within provinces, we cannot use the official forestry statistics dataset for this purpose.

2.4.2 Results

The results from estimating equation (7) are shown in Table 2.9. Panel A shows the impact effect of elections (i.e., no leads and lags); Panel B presents the results with 2 leads and lags of the *Election* variable. As before, we present results for the entire Forest Estate, as well as broken down by land use zone.

The results show clear evidence of a political logging cycle in the illegal forest zones. Focusing on column 3 of Panel B, which shows the impact on the Conservation and Protection Forest where no legal logging is allowed, we find that illegal logging increases dramatically in the years leading up to an election: by 29.4% 2 years prior to the election and by 42.7% in the year before the election. Illegal logging then falls dramatically (by 36.4%) in the election year and does not resume thereafter. Looking zone-by-zone, we see that the pattern is strongest statistically in the Protection Forest (column 7), but that the point estimates suggest a very similar pattern in the Conservation Forest as well (column 6).

There are several possible explanations for the increase in illegal logging in the years leading up to the elections. One potential explanation is that logging was permitted or facilitated by district officials in return for campaign funds.[64] A second explanation is that district officials simply reduced enforcement of logging in the Conservation and Protection Forest in order to increase their popularity and win votes. Since these two stories are observationally equivalent in terms of the predicted impact on deforestation, it is not possible to tease them apart empirically.

Turning at the zones where logging may be legal or illegal (the Conversion and Production Forest), we see a different pattern. In the Conversion Forest, we find

[63] The omitted category is therefore the years prior to 2 years before the first direct election.

[64] Although we know of no direct qualitative evidence for this link, Greenpeace Indonesia has asserted that political parties amassed campaign funds for the 2009 general election through facilitating illegal logging (Greenpeace, 2009).

a 40.5% increase in logging in the year of the election and a 57.9% increase in the year following the election. We find no impact on the Production Forest. According to Barr et al. (2006), many district governments have redirected their interest towards the development of oil palm plantations and other agroindustrial estates in recent years. It is possible that the observed increase in clear-cutting in the Conversion Forest after the election is a repayment for favours during the election campaign. Alternatively, it could be an attempt to grab first rents upon being elected. Once again, these stories are observationally equivalent, so it is not possible to tease them apart empirically with the existing data. Since the effects in the Conservation/Protection Forest and the Production/Conversion Forest have different patterns, column 1 shows little impact overall.

2.5 Substitutes or complements? Logging vs. other sources of rents

2.5.1 Empirical implementation

An important question in the economics of corruption is how corrupt officials with multiple opportunities for rent extraction respond if one type of corruption becomes harder or easier. If corrupt officials behave like profit maximizing firms, and there are no spillovers from one type of corrupt activity to the other, then they would optimize separately on each dimension, and there would be no impact of a change in one type of corruption opportunity on the other type of corruption.

More generally, however, one could imagine effects going in either direction. If corrupt officials worry about being detected, and if being detected means the official loses both types of corruption opportunities, then the two types of corruption will appear to be substitutes, and increasing corruption opportunities on one dimension will lower them on the other dimension. On the other hand, if there are fixed costs of being corrupt (for example, those with a low disutility from being corrupt selecting into the civil service), multiple corruption opportunities could be complements. The two existing studies that have examined this question empirically (Olken (2007) and Niehaus and Sukhtankar (2009)) have both found evidence that alternative forms of corruption appear to be substitutes.

In this section, we examine this question by examining how logging responds to changes in another source of local rents for district governments: oil and gas revenues. Under Indonesia's Fiscal Balancing Law (ROI, 1999b), a fraction of all oil and gas royalties received by the central government is rebated back to districts, with half of the rebate going to the district that produces the oil and gas and the other half being shared equally among all other districts in the same province. This can amount to a substantial amount of revenue – as much as US$729.63 per capita in the highest district – which can in turn be a tempting source of rents for district officials.[65] Moreover, the

[65] District government officials have recently been exposed in a wide variety of strategies to capture rents from the oil and gas revenue sharing fund. In Kabupaten Kutai Kartanegara, for example, the

precise amount of oil and gas revenue allocated to each district varies substantially over time as oil and gas production fluctuates, oil and gas prices change, and district boundaries change.

A key distinction between our context and the existing literature is that while the latter (Olken (2007) and Niehaus and Sukhtankar (2009)) studies short-run substitution from one type of corruption to another, our setting allows us to examine both the short and medium run. If the fixed costs of corruption are important, adjustment may take time, and the short- and medium-run effects could be quite different.

To examine the short-run impact of oil and gas rents on illegal logging, we estimate a version of equation (6). Since district splits influence oil and gas revenue through the sharing formula, we control for district splits directly, and estimate the following equation:

$$\mathbf{E}\left(deforest_{dit}\right) = \mu_{di} \exp\left(\beta PCOilandGas_{dit} + \gamma Numdistricts_{dit} + \eta_{it}\right) \tag{8}$$

where $PCOilandGas_{dit}$ is the per-capita oil and gas revenue received by the district (in US$). Note that in computing $Numdistricts_{dit}$ when estimating (8), we count a district as having split only when it reports receiving its own oil and gas revenue.[66] As above, μ_{di} is a district fixed effect, η_{it} is an island-by-year fixed effect, and robust standard errors are reported, adjusted for clustering at the 1990 district boundaries. Since district oil and gas sharing revenue is, on average, 20 times larger than that generated by the forestry sector, one would not expect forestry decisions to influence oil and gas choices. Oil and gas revenue should thus be exogenous with respect to deforestation. To examine the medium-run impacts of oil and gas rents on illegal logging, we estimate (8) as above, but include 3 lags of $PCOilandGas_{dit}$.

2.5.2 Results

The results from estimating equation (8) are shown in Table 2.10. Panel A, which shows the immediate impact of oil and gas revenue on logging, confirms evidence of short-run substitution between deforestation and oil and gas rents. Specifically, each US$1 of per-capita oil and gas rents received by the district reduces logging by 0.32%. These effects are found in both the legal logging zones (0.28% in Pro-

national Anti-Corruption Commission recently documented that in 2001 the *Bupati* simply issued a decree giving himself, top district government officials, and district parliamentarians an official monthly stipend equal to 3% of the amount the government received in oil and gas revenue, amounting to over US$9 million over a 4-year period (KaltimPost, 2009b,a). In Kabupaten Natuna, Sumatra, a former *Bupati* was arrested in 2009 by the Anti-Corruption Commission for allegedly embezzling US$8 million in oil and gas revenue funds, by appropriating the funds to a fake committee that he never set up (Kompas, 2009). In Kabupaten Karawang, West Java, in 2004 the *Bupati* allegedly deposited US$600,000 in oil and gas revenue sharing funds into his personal account rather than the district treasury (KoranTempo, 2006).

[66] As described above, the *de facto* establishment of a district takes 1-3 years after the official *de jure* implementation. Since we care about district splits in this case because they affect the oil and gas allocation formula, it is important to control here for the *de facto* date the district split took effect, as that is the date the oil and gas formula would be affected.

duction/Conversion Forest; column 2) and in the illegal logging zones (0.59% in the Conservation/Protection Forest). To interpret the magnitudes, note that the standard deviation of $PCOilandGas_{dit}$ after removing district fixed effects is 23.7; so a one-standard deviation change in $PCOilandGas_{dit}$ decreases deforestation by 6.64% in the Production/Conversion Forest and by 13.98% in the Conservation/Protection Forest.

Panel B shows, however, that the short-run and medium-run effects are quite different. While the immediate effect of oil and gas revenue on logging is still negative (0.49% per US$1, Panel B, column 1), the sum of the lags is now positive. That is, after three years, the total medium-run effect of US$1 of per-capita oil and gas rents is to increase logging by 0.63%. Once again, this shift occurs equally in the legal logging zones (0.64%, column 2) and illegal logging zones (0.54%, column 3), although the illegal logging zone effect is not statistically significant. Once again, these magnitudes are substantial: a one-standard deviation increase in $PCOilandGas_{dit}$ leads to a medium-run increase in deforestation of 14.93%. This suggests that the short- and medium-run impacts are different, and in the medium run, oil and gas rents and rents from logging appear to be complements.

2.6 Conclusions

This paper has demonstrated how the incentives faced by local politicians and bureaucrats play an important role in determining the rate of deforestation in Indonesia. Using a novel MODIS satellite-based dataset that tracks deforestation on an annual basis across the whole of Indonesia, we have shown that the rates of deforestation are influenced by the return local officials face in the market for logs, by their short-term electoral needs, and by the availability of alternative sources of rent extraction.

More broadly, the results in the paper demonstrate that the pattern of forest cutting in Indonesia is not the result of some optimal forest management model implemented by the Ministry of Forestry. The decisions that matter for whether trees are cut down or not take place not just in ministerial meeting rooms. They are a result of the push and shove of local politics.

To the extent one wants to slow the rate of tropical deforestation, these results matter. They suggest that unless the incentives of local politicians and bureaucrats are taken into account, then central or donor-driven policies to counter deforestation may be ineffective. In particular, the recently ratified Reducing Emissions from Deforestation and Forest Degradation (REDD) initiative provides countries with a source of funding to reduce deforestation that is likely to grow considerably over time. Indeed, Indonesia is to be the first recipient of REDD funds, having signed a US$1 billion REDD agreement with Norway in 2009. However, unless REDD programs are designed to consider those local actors who currently derive considerable benefits from legal and illegal logging, they are unlikely to be effective.

Though instructive in terms of revealing how political economy plays a central role in the deforestation process, our paper very much leaves open the central question of how to counter this process. Nonetheless, the results imply that taking the financial and electoral pressures upon local politicians and bureaucrats seriously is a crucial step to sustainable tropical forestry.

B Data appendix

This appendix provides information complementary to Section 2.2 on the variables used in this paper. These variables will be constructed for all provinces and districts on the islands of Sumatra, Kalimantan, Sulawesi and Papua. We use three different samples. The first sample uses the 1990 province boundaries and comprises 21 provinces.[67] In addition, we have two district-level samples. The first sample uses the 1990 boundaries and includes 146 districts and municipalities and the second uses the 2008 boundaries for a total 311 districts and municipalities.

B.1 Government forestry data

The government forestry dataset has been compiled from the annual 'Statistics of Forest and Concession Estate' (*Statistik Perusahaan Hak Pengusahaan Hutan*) published by the Indonesian Central Bureau of Statistics (*Badan Pusat Statistik*). This data is available at the province level for 1994-2007. It reports the volume of logs cut in cubic metres and the value of production in 1,000 *rupiahs* for 114 wood types. We drop the 'other' (*Lainnya)* and 'mixed wood' (*Rimba Campuran)* category, since their composition varies considerably across provinces and over time.

We calculate the prices for the remaining 112 wood types by dividing the value of production with the volume produced. The price data is then converted into US dollars and deflated by the US Consumer Producer Index for 2000. Subsequently, we sum the volume and deflated price data for each wood type within the 1990 province boundaries. By taking logs we obtain our two dependent variables used in Section 2.3.3. Note that the associated regressions will be weighted by the first volume of production reported for each wood type.

B.2 District split data

We collect data on district and province boundaries from the publications on regency and municipality codes of the Central Bureau of Statistics. The most up-to-date list is available at `http://dds.bps.go.id/eng/aboutus.php?mstkab=1`. Since the coding of districts has been revised several times between 1990 and 2008, we create a master district dataset that only uses 2008 identifiers for all years and follows the original 1990 districts as they split over time. Note that we use the 1990 boundaries as a reference point, because 17 new *kabupaten* and *kota* were formed between 1990 and 1999.[68]

[67]These provinces are Nanggroe Aceh Darussalam, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Lampung, Kepulauan Bangka Belitung, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Papua Barat and Papua

[68]During the Soeharto regime, only 3 new *kabupaten* or *kota* were created outside of Jakarta prior to 1990: Kota Ambon (PPRI No. 13 Thn. 1979), Kota Batam (PPRI No. 34. Thn. 1983), and Kab. Aceh Tenggara (UURI NO. 4 Thn. 1984). Jakarta itself was split into 5 city parts in 1978.

For our Cournot analysis we construct the following set of variables. For the province-level data, we simply count the total number districts and municipalities within the 1990 province boundaries. We also separately sum the number of *kabupaten* and *kota* for each province and year, which will be used as a robustness check. In addition, we construct two more variables at the district-level. Firstly, we count into how many districts and municipalities the original 1990 district split into in a given year. Secondly, we sum across all the remaining administrative units within the same province. Finally, we construct three leads and lags for each variable in turn.

B.3 District election data

The district-level election schedules are obtained from the Centre for Electoral Reform (CETRO), which is an Indonesian NGO (`http://www.cetro.or.id/newweb/index.php`). We use the most up-to-date dataset on district head elections, which reports the dates of all of the direct *Bupati* and *Walikota* elections held since 2005. It also provides information on elections that are scheduled up to 2011. From this dataset we construct a dummy, which is equal to one in the year of the election. We also calculate two leads and lags of this indicator variable.

B.4 District public finance data

The dataset on oil and gas revenue has been compiled from the webpage of the Indonesian Ministry of Finance (*Menteri Keuangan*), which is available at `http://www.djpk.depkeu.go.id/datadjpk/57/`. The revenue figures are reported for each district for 2001-2008. It is important to note that new districts often do not record their own share of revenue for the first few years after the split, as they are still setting up their basic government functions. We therefore allocate each new district the revenue share of its originating district until it reports its own share for the first time.

The revenue data has then been converted into US dollars and divided by the district-level population figure for 2008. This yields the Oil and Gas Revenue per capita variable used in our analysis. The population data for 2008 has been downloaded from the Indonesian Central Bureau of Statistics.

Figure 2.1: Forest cover change in the province of Riau, 2001-2008

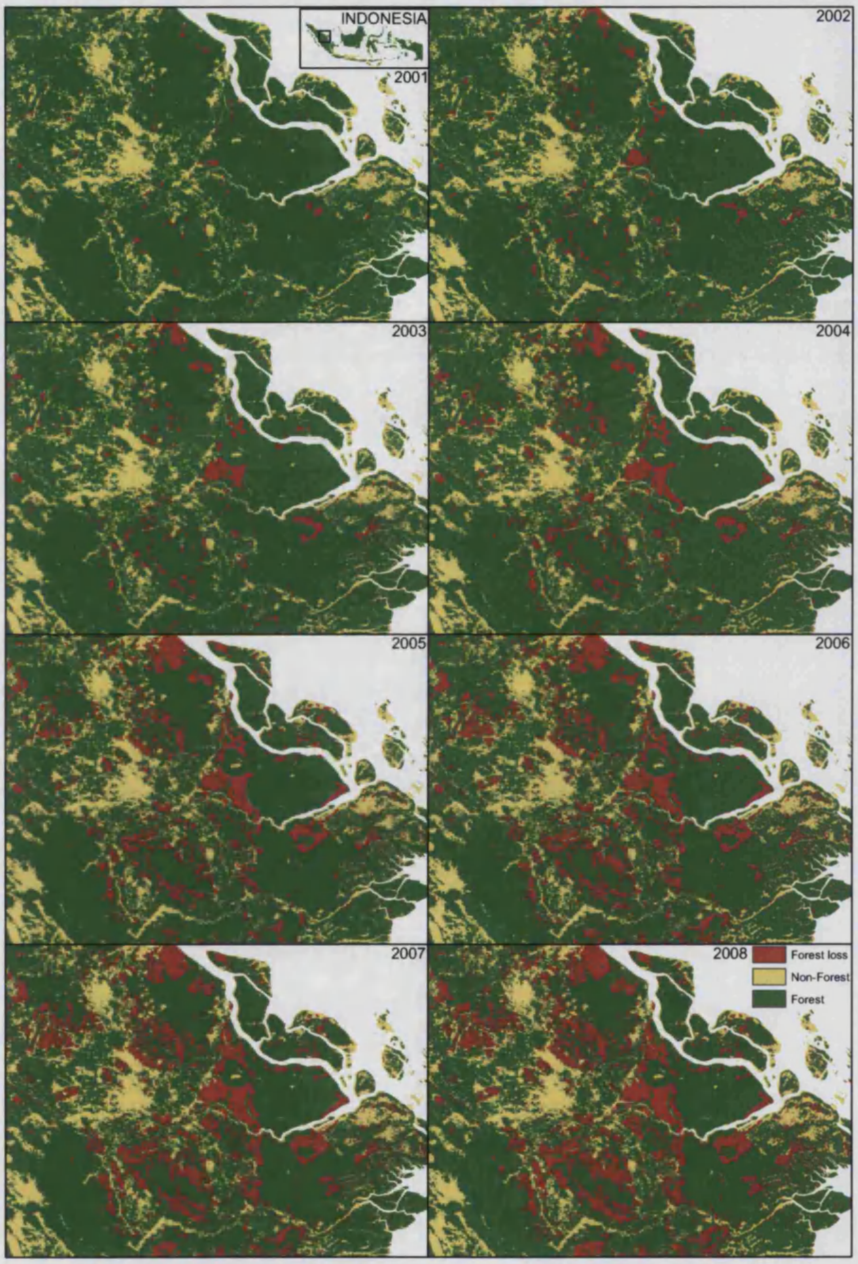


Figure 2.2: District-level logging in Indonesia using the 2008 district boundaries

(a) 2001

Figure 2.2: District-level logging in Indonesia using the 2008 district boundaries

(b) 2008

Figure 2.3: Total number of district splits using the 1990 district boundaries, 1990-2008

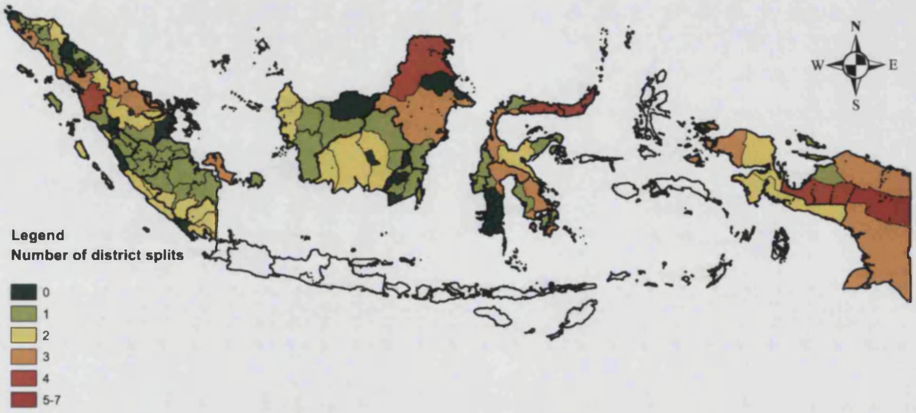


Figure 2.4: Oil and gas revenue per capita using the 2008 district boundaries, 2008

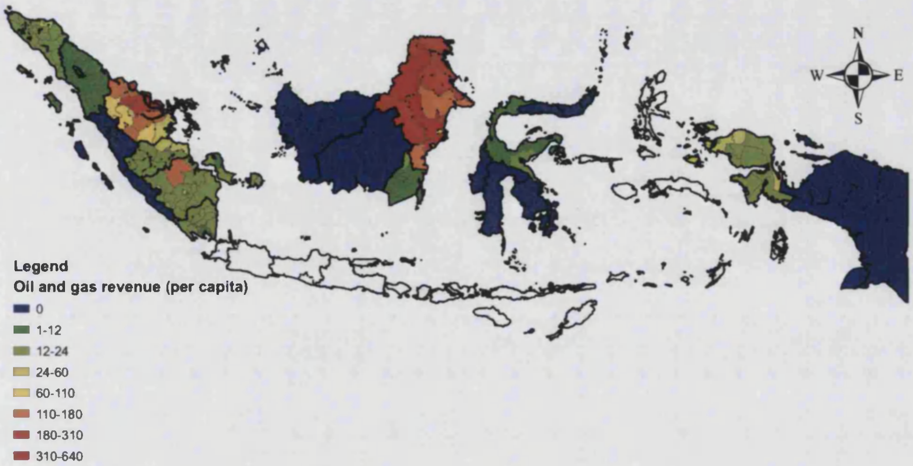


Table 2.1: Forest area in 1000 HA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Year	Total land area	2000	2001	2002	2003	2004	2005	2006	2007	2008	Change 2000-08
All Forest	118,664	109,795	109,335	108,459	108,047	107,499	106,970	105,916	105,349	104,901	-4,894
Legal Forest	74,339	67,908	67,521	66,858	66,502	66,049	65,579	64,677	64,154	63,745	-4,163
Illegal Forest	44,325	41,886	41,814	41,601	41,545	41,449	41,391	41,238	41,195	41,156	-731
Conversion	19,363	16,576	16,460	16,294	16,197	16,065	15,912	15,704	15,566	15,455	-1,121
Production	54,977	51,332	51,061	50,564	50,305	49,984	49,667	48,973	48,588	48,290	-3,042
Conservation	17,074	15,720	15,692	15,564	15,537	15,490	15,472	15,381	15,357	15,343	-377
Protection	27,250	26,166	26,122	26,037	26,007	25,959	25,919	25,857	25,838	25,812	-354
Change all forest			-459	-877	-412	-548	-529	-1,054	-567	-448	-4,894

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres.

Table 2.2: Summary statistics of forest area cleared in 1000HA by districtXyear

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Logging	All Forest	Production/ Conversion	Conservation/ Protection	Conversion	Production	Conservation	Protection
Mean	0.704	1.269	0.199	0.947	1.451	0.248	0.165
Std. deviation	2.901	4.008	1.022	2.644	4.594	1.38	0.661
Observations	6952	3280	3672	1184	2096	1520	2152

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. The variable shown here is the key dependent variable analyzed in Sections 2.3-2.5.

Table 2.3: Impact of Splits on Prices and Quantities: Legal Logging Data

	(1)	(2)	(3)	(4)	(5)	(6)
	2001-2007		2001-2007		1994-2007	
	All wood obs		Balanced panel wood obs		All wood obs	
VARIABLES	Log Price	Log Quantity	Log Price	Log Quantity	Log Price	Log Quantity
	(1)	(2)	(3)	(4)	(5)	(6)
	2001-2007		2001-2007		1994-2007	
	All wood obs		Balanced panel wood obs		All wood obs	
VARIABLES	Log Price	Log Quantity	Log Price	Log Quantity	Log Price	Log Quantity
Panel A						
Number of districts	-0.017*	0.089**	-0.019*	0.106**	-0.023**	0.081***
in province	(0.009)	(0.041)	(0.010)	(0.036)	(0.009)	(0.016)
Observations	1003	1003	532	532	2355	2355
Panel B: Lags						
Number of districts	-0.025**	0.098	-0.027**	0.126	-0.029***	0.071***
in province	(0.010)	(0.074)	(0.012)	(0.078)	(0.008)	(0.023)
Lag 1	0.010**	-0.041	0.009	-0.035	0.010**	-0.001
	(0.004)	(0.036)	(0.005)	(0.041)	(0.004)	(0.035)
Lag 2	-0.001	0.041	-0.001	0.018	0	0.017
	(0.008)	(0.045)	(0.009)	(0.021)	(0.004)	(0.027)
Lag 3	-0.017**	0.033	-0.017**	0.043	-0.015*	0.029
	(0.006)	(0.044)	(0.007)	(0.040)	(0.008)	(0.037)
Observations	1003	1003	532	532	1960	1960
Joint p	0.00271	0.000533	0.00756	0.000583	0.000109	0.00645
Sum of lags	-0.0329***	0.131**	-0.0361**	0.153**	-0.0339**	0.117***
	(0.0103)	(0.0527)	(0.0116)	(0.0505)	(0.0131)	(0.0363)

Notes: The log price and log quantity data has been compiled from the 'Statistics of Forest and Concession Estate'. The *Number of districts in province* variable counts the number of *kabupaten* and *kota* within each province. The regression also includes wood-type-by-province and wood-type-by-island-by-year fixed effects and are weighted by the first volume reported by wood type and province. The robust standard errors are clustered at the 1990 province boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.4: Impact of Splits on Prices and Quantities: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	2001-2007		2001-2007		1994-2007	
	All wood obs		Balanced panel wood obs		All wood obs	
VARIABLES	Log Price	Log Quantity	Log Price	Log Quantity	Log Price	Log Quantity
Panel A: Contemporaneous weights						
Number of districts	-0.017	0.096*	-0.021*	0.108**	-0.024***	0.056**
in province	(0.010)	(0.045)	(0.010)	(0.045)	(0.007)	(0.019)
Observations	1003	1003	532	532	2357	2357
Panel B: No cities						
Number of districts	-0.018*	0.088*	-0.020*	0.106**	-0.026***	0.073***
in province	(0.008)	(0.042)	(0.009)	(0.036)	(0.008)	(0.020)
Observations	1003	1003	532	532	2355	2355
Panel C: Cities only (falsification)						
Number of districts	0.098	0.420	0.157	0.494	-0.001	0.184**
in province	(0.131)	(0.353)	(0.130)	(0.319)	(0.043)	(0.084)
Observations	1003	1003	532	532	2355	2355

Table 2.4: Impact of Splits on Prices and Quantities: Alternative Specifications (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel D: Leads						
Number of districts	-0.020***	0.111	-0.019**	0.151**	-0.019*	0.098*
in province	(0.006)	(0.074)	(0.007)	(0.061)	(0.009)	(0.053)
Lag 1	0.007	-0.043	0.006	-0.038	0.006	-0.010
	(0.006)	(0.037)	(0.008)	(0.044)	(0.006)	(0.041)
Lag 2	-0.003	0.036	-0.004	0.002	-0.005	0.013
	(0.004)	(0.055)	(0.005)	(0.018)	(0.004)	(0.032)
Lag 3	-0.016**	0.040	-0.017*	0.048	-0.021**	0.037
	(0.007)	(0.050)	(0.008)	(0.044)	(0.010)	(0.040)
Lead 1	0.023	0.036	0.050*	0.236**	-0.008	-0.007
	(0.014)	(0.121)	(0.025)	(0.088)	(0.010)	(0.022)
Lead 2	-0.022	0.039	-0.032	0.040	-0.009*	-0.035
	(0.049)	(0.249)	(0.057)	(0.270)	(0.004)	(0.059)
Lead 3	0.002	-0.088	0.000	-0.207	-0.018*	0.020
	(0.029)	(0.218)	(0.042)	(0.227)	(0.009)	(0.041)
Observations	865	865	456	456	1822	1822
Joint p	2.28e-06	8.16e-05	5.66e-08	7.09e-06	9.51e-05	0.000173
Sum of lags	-0.0330***	0.144**	-0.0337***	0.163**	-0.0386**	0.138**
	(0.0107)	(0.0593)	(0.00904)	(0.0646)	(0.0162)	(0.0524)
Sum of leads	0.00272	-0.0129	0.0184	0.0684	-0.0354***	-0.0217
	(0.0456)	(0.211)	(0.0376)	(0.287)	(0.00960)	(0.0407)
Joint p leads	0.379	0.980	0.0639	0.0532	0.00614	0.907

Notes: The log price and log quantity data has been compiled from the 'Statistics of Forest and Concession Estate'. The *Number of districts in province* variable counts the number of *kabupaten* and *kota* within each province. The regression also includes wood-type-by-province and wood-type-by-island-by-year fixed effects and are weighted by the first volume reported by wood type and province. The robust standard errors are clustered at the 1990 province boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.5: Satellite data on impact of splits, province level

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Panel A							
Number of districts	0.0361**	0.0424**	0.0391	0.0283	0.0533***	0.0786*	0.00645
in province	(0.0160)	(0.0180)	(0.0317)	(0.0333)	(0.0199)	(0.0415)	(0.0322)
Observations	672	336	336	128	168	144	168
Panel B: Lags							
Number of districts	0.0370	0.0435	0.0833***	0.0447	0.0523	0.0959**	0.0657*
in province	(0.0284)	(0.0332)	(0.0299)	(0.0420)	(0.0350)	(0.0417)	(0.0377)
Lag 1	0.0405	0.0434	-0.129**	0.00823	0.0419	-0.170	-0.0732
	(0.0446)	(0.0461)	(0.0651)	(0.0641)	(0.0434)	(0.130)	(0.0623)
Lag 2	-0.0717***	-0.0740***	0.0186	-0.0883**	-0.0625**	0.111	-0.0851
	(0.0265)	(0.0250)	(0.0762)	(0.0346)	(0.0257)	(0.153)	(0.0679)
Lag 3	0.0731*	0.0654	0.117*	0.107	0.0476	0.0889	0.141**
	(0.0397)	(0.0399)	(0.0610)	(0.0880)	(0.0357)	(0.0614)	(0.0610)
Observations	672	336	336	128	168	144	168
Joint p	4.75e-06	6.95e-08	0.0235	0.0428	0.000923	0.0486	0.0665
Sum of lags	0.0789***	0.0783***	0.0900**	0.0712	0.0793***	0.125**	0.0484
	(0.0200)	(0.0190)	(0.0400)	(0.0616)	(0.0214)	(0.0611)	(0.0357)

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *Number of districts in province* variable counts the number of *kabupaten* and *kota* within each province. The regression also includes wood-type-by-province and wood-type-by-island-by-year fixed effects. The robust standard errors are clustered at the 1990 province boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.6: Satellite data on impact of splits with leads, province level

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Number of districts	0.0406	0.0444	0.0882**	-0.0105	0.0637	0.138***	0.00976
in province	(0.0396)	(0.0460)	(0.0352)	(0.0304)	(0.0491)	(0.0490)	(0.0614)
Lag 1	0.0244	0.0202	-0.105	-0.0126	0.0166	-0.124	-0.0517
	(0.0480)	(0.0511)	(0.0692)	(0.0834)	(0.0473)	(0.104)	(0.0773)
Lag 2	-0.0603	-0.0547	-0.00237	-0.0712	-0.0395	0.0400	-0.0822
	(0.0385)	(0.0362)	(0.0853)	(0.0588)	(0.0336)	(0.122)	(0.0819)
Lag 3	0.0856*	0.0755	0.135	0.148	0.0584	0.156*	0.134
	(0.0518)	(0.0494)	(0.0884)	(0.123)	(0.0413)	(0.0924)	(0.0947)
Lead 1	0.0879	0.0925	0.0498	0.324*	0.0370	0.172	0.0444
	(0.114)	(0.120)	(0.136)	(0.173)	(0.110)	(0.138)	(0.136)
Lead 2	-0.118	-0.156	-0.0141	-0.257	-0.149	0.132	-0.0897
	(0.137)	(0.136)	(0.163)	(0.180)	(0.131)	(0.180)	(0.170)
Lead 3	0.0364	0.0635	-0.0432	0.117	0.0689	-0.157	0.0180
	(0.107)	(0.104)	(0.103)	(0.130)	(0.103)	(0.115)	(0.112)
Observations	504	252	252	96	126	108	126
Joint p	0.000251	0	0.0129	0	0	2.72e-09	0.0817
Sum of lags	0.0903***	0.0854***	0.116*	0.0536	0.0992***	0.210**	0.00992
	(0.0281)	(0.0239)	(0.0663)	(0.0677)	(0.0223)	(0.0944)	(0.0774)
Sum of leads	0.00586	-5.16e-05	-0.00758	0.184	-0.0432	0.147	-0.0274
	(0.0663)	(0.0587)	(0.0976)	(0.132)	(0.0566)	(0.148)	(0.0889)
Joint p leads	0.714	0.660	0.608	0.201	0.296	0.430	0.550

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *Number of districts in province* variable counts the number of *kabupaten* and *kota* within each province. The regression also includes province and island-by-year fixed effects. The robust standard errors are clustered at the 1990 province boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.7: Satellite data on impact of splits, district level

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Panel A							
Number of districts in	-0.102	-0.172*	0.0663	-0.0174	-0.211**	0.136*	-0.0284
original district boundaries	(0.0778)	(0.0913)	(0.0519)	(0.150)	(0.0864)	(0.0767)	(0.0839)
Number of districts	0.0701**	0.0967***	0.0336	0.0380	0.122***	0.0677	0.0138
elsewhere in province	(0.0275)	(0.0311)	(0.0308)	(0.0486)	(0.0326)	(0.0452)	(0.0315)
Observations	6872	3136	3544	1112	1952	1360	2040
Panel B: Lags							
Number of districts in	-0.0627	-0.0984	0.107**	0.0139	-0.133	0.151*	0.0421
original district boundaries	(0.0830)	(0.103)	(0.0542)	(0.154)	(0.0969)	(0.0874)	(0.0576)
Lag 1	-0.0185	-0.0780	-0.0739	0.207	-0.140	-0.0828	-0.0259
	(0.130)	(0.159)	(0.103)	(0.239)	(0.141)	(0.138)	(0.0806)
Lag 2	-0.0767	-0.129	0.0252	-0.438	-0.0625	0.153	-0.143
	(0.115)	(0.151)	(0.0956)	(0.287)	(0.133)	(0.161)	(0.103)
Lag 3	0.190***	0.218***	0.193**	0.154	0.243***	0.148	0.261**
	(0.0669)	(0.0737)	(0.0794)	(0.138)	(0.0789)	(0.0940)	(0.106)

Table 2.7: Satellite data on impact of splits, district level (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of districts	0.0702*	0.0901**	0.0883***	0.0356	0.116***	0.105**	0.0771**
elsewhere in province	(0.0371)	(0.0434)	(0.0316)	(0.0599)	(0.0384)	(0.0436)	(0.0357)
Lag 1	0.0582	0.0802	-0.140**	-0.0296	0.0946	-0.194*	-0.0803
	(0.0584)	(0.0643)	(0.0572)	(0.0872)	(0.0619)	(0.100)	(0.0556)
Lag 2	-0.0656	-0.0535	0.0207	-0.00584	-0.0517	0.101	-0.0573
	(0.0477)	(0.0520)	(0.0780)	(0.0668)	(0.0543)	(0.119)	(0.0973)
Lag 3	0.0322	0.0111	0.0935	0.0932	-0.0238	0.0732	0.0972
	(0.0396)	(0.0426)	(0.0584)	(0.0776)	(0.0445)	(0.0527)	(0.0629)
Observations	6872	3136	3544	1112	1952	1360	2040
Joint p original	0.0632	0.00753	0.0555	0.119	0.00465	0.212	0.0120
Sum of lags original	0.0323	-0.0867	0.251***	-0.0623	-0.0929	0.370**	0.133**
	(0.114)	(0.115)	(0.0964)	(0.193)	(0.115)	(0.176)	(0.0680)
Joint p elsewhere	0.0100	0.00331	0.0265	0.589	0.00118	0.00983	0.130
Sum of lags elsewhere	0.0951**	0.128***	0.0622	0.0934	0.135***	0.0851	0.0367
	(0.0390)	(0.0432)	(0.0385)	(0.0586)	(0.0480)	(0.0654)	(0.0311)

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *Number of districts in original district boundaries* variable counts the number of *kabupaten* and *kota* the district split into and the *Number of districts elsewhere in province* variable counts the number of *kabupaten* and *kota* all other districts within the same province split into. The regression also includes district and island-by-year fixed effects. The robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.8: Satellite data on impact of splits with leads, district level

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Number of districts in	-0.0291	0.0379	-0.00473	0.00271	0.0477	0.101	-0.0564
original district boundaries	(0.110)	(0.178)	(0.0500)	(0.309)	(0.168)	(0.0688)	(0.0909)
Lag 1	-0.0660	-0.205	-0.0172	0.0381	-0.279	0.00135	0.0145
	(0.152)	(0.195)	(0.0953)	(0.323)	(0.179)	(0.112)	(0.0801)
Lag 2	-0.0613	-0.107	-0.0248	-0.255	-0.0728	0.0493	-0.144
	(0.139)	(0.196)	(0.0972)	(0.328)	(0.175)	(0.121)	(0.0949)
Lag 3	0.244***	0.286***	0.164	0.298**	0.305***	0.200*	0.242*
	(0.0720)	(0.0900)	(0.105)	(0.148)	(0.0921)	(0.110)	(0.128)
Lead 1	0.136	0.0859	0.395	0.800*	0.0237	0.907**	0.250
	(0.243)	(0.247)	(0.281)	(0.449)	(0.234)	(0.418)	(0.179)
Lead 2	-0.306	-0.396	-0.0744	-0.624	-0.371	0.291*	-0.301
	(0.247)	(0.296)	(0.258)	(0.401)	(0.274)	(0.175)	(0.341)
Lead 3	0.0933	0.117	-0.154	0.405*	0.0268	-0.503**	0.224
	(0.125)	(0.122)	(0.222)	(0.213)	(0.141)	(0.231)	(0.318)
Number of districts	0.0673	0.0563	0.109***	-0.00830	0.0813	0.148***	0.0297
elsewhere in province	(0.0485)	(0.0657)	(0.0372)	(0.0921)	(0.0645)	(0.0419)	(0.0603)
Lag 1	0.0541	0.0936	-0.115**	-0.0157	0.115*	-0.135**	-0.0537
	(0.0602)	(0.0638)	(0.0526)	(0.0997)	(0.0650)	(0.0680)	(0.0609)
Lag 2	-0.0559	-0.0364	-0.00989	-0.00795	-0.0293	0.0209	-0.0455
	(0.0559)	(0.0636)	(0.0799)	(0.0922)	(0.0646)	(0.0891)	(0.0910)
Lag 3	0.0285	-0.00160	0.130*	0.103	-0.0338	0.153**	0.0827
	(0.0372)	(0.0402)	(0.0675)	(0.0972)	(0.0423)	(0.0740)	(0.0686)

Table 2.8: Satellite data on impact of splits with leads, district level (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lead 1	0.0989	0.120	-0.0228	0.257**	0.0758	0.110	-0.0148
	(0.0857)	(0.100)	(0.110)	(0.124)	(0.0986)	(0.132)	(0.103)
Lead 2	-0.0457	-0.0485	-0.00274	-0.235	-0.0293	0.0572	-0.0128
	(0.103)	(0.109)	(0.109)	(0.168)	(0.110)	(0.150)	(0.128)
Lead 3	-0.00966	-0.0213	-0.00960	0.0814	-0.0185	-0.0654	-0.0398
	(0.0855)	(0.0897)	(0.0946)	(0.151)	(0.0906)	(0.123)	(0.119)
Observations	5148	2352	2646	810	1464	990	1512
Joint p original	0.0756	0.000177	0.000515	0.0769	0.000632	8.78e-05	0.527
Sum of lags original	0.0874	0.0115	0.117	0.0830	0.00154	0.352**	0.0555
	(0.127)	(0.154)	(0.107)	(0.247)	(0.154)	(0.161)	(0.101)
Sum of leads original	-0.0771	-0.193	0.167	0.581	-0.321	0.695*	0.173
	(0.262)	(0.292)	(0.311)	(0.406)	(0.296)	(0.416)	(0.174)
Joint p leads original	0.530	0.606	0.0655	0.154	0.559	0.0946	0.530
Joint p elsewhere	0.323	0.222	0.00978	0.0558	0.137	0.001000	0.609
Sum of lags elsewhere	0.0940**	0.112**	0.114**	0.0711	0.133**	0.187**	0.0132
	(0.0477)	(0.0570)	(0.0579)	(0.0948)	(0.0602)	(0.0787)	(0.0663)
Sum of leads elsewhere	0.0435	0.0503	-0.0351	0.103	0.0280	0.102	-0.0674
	(0.0660)	(0.0832)	(0.0724)	(0.124)	(0.0833)	(0.119)	(0.0768)
Joint p leads elsewhere	0.669	0.654	0.953	0.0485	0.881	0.854	0.663

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *Number of districts in original district boundaries* variable counts the number of *kabupaten* and *kota* the district split into and the *Number of districts elsewhere in province* variable counts the number of *kabupaten* and *kota* all other districts within the same province split into. The regression also includes district and island-by-year fixed effects. The robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.9: Elections

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Panel A							
ElectionYear	-0.133	-0.0732	-0.593***	0.124	-0.128	-0.398***	-0.658***
	(0.0959)	(0.112)	(0.155)	(0.156)	(0.107)	(0.117)	(0.214)
Observations	6872	3136	3544	1112	1952	1360	2040
Panel B: Leads and Lags							
ElectionYear	0.0277	0.0804	-0.364**	0.405*	-0.00920	-0.125	-0.493***
	(0.142)	(0.155)	(0.152)	(0.241)	(0.151)	(0.187)	(0.183)
Lead 1	0.200	0.173	0.427**	0.242	0.134	0.244	0.501**
	(0.130)	(0.140)	(0.216)	(0.226)	(0.146)	(0.171)	(0.220)
Lead 2	0.131	0.120	0.294**	0.295	0.0869	0.223	0.300**
	(0.166)	(0.185)	(0.130)	(0.223)	(0.184)	(0.149)	(0.134)
Lag 1	0.282*	0.305*	0.140	0.579**	0.235	0.352	-0.111
	(0.155)	(0.170)	(0.217)	(0.236)	(0.186)	(0.282)	(0.201)
Lag 2	-0.0427	-0.0463	0.0180	0.0896	-0.0671	0.0892	-0.103
	(0.173)	(0.193)	(0.266)	(0.302)	(0.205)	(0.339)	(0.236)
Observations	6872	3136	3544	1112	1952	1360	2040
Lags Joint p	0.00305	0.00447	0.000358	1.61e-06	0.0383	0.0695	0.0257
Sum of lags	0.267	0.339	-0.206	1.074	0.158	0.315	-0.708
	(0.429)	(0.470)	(0.547)	(0.733)	(0.489)	(0.664)	(0.500)
Leads Joint p	0.291	0.458	0.0598	0.413	0.641	0.252	0.0418
Sum of leads	0.331	0.293	0.721**	0.536	0.221	0.468*	0.801**
	(0.270)	(0.295)	(0.314)	(0.418)	(0.302)	(0.283)	(0.320)

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *ElectionYear* variable is a dummy equal to 1 if the district holds a *Bupati* or *Walikota* election that year. The regression also includes district and island-by-year fixed effects. The robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

Table 2.10: The Income Elasticity of Corruption

VARIABLES	(1) All Forest	(2) Production/ Conversion	(3) Conservation/ Protection	(4) Conversion	(5) Production	(6) Conservation	(7) Protection
Panel A							
Oil and Gas	-0.00316**	-0.00284*	-0.00596**	-0.00912***	-0.00220	-0.00474**	-0.00986***
Revenue pc	(0.00160)	(0.00164)	(0.00252)	(0.00165)	(0.00146)	(0.00218)	(0.00148)
Observations	6872	3136	3544	1112	1952	1360	2040
Panel B: Lags							
Oil and Gas	-0.00494**	-0.00429**	-0.0118***	-0.0115***	-0.00370*	-0.0112***	-0.0119***
Revenue pc	(0.00205)	(0.00205)	(0.00311)	(0.00212)	(0.00191)	(0.00436)	(0.00215)
Lag 1	0.000621	-1.07e-05	0.00702***	0.00462*	0.000240	0.00929***	-0.000758
	(0.00129)	(0.00155)	(0.00121)	(0.00273)	(0.00134)	(0.00180)	(0.00238)
Lag 2	0.00195	0.00216	0.000913	-0.000364	0.00236	0.00284*	-0.00103
	(0.00142)	(0.00160)	(0.00158)	(0.00266)	(0.00167)	(0.00159)	(0.00248)
Lag 3	0.00863***	0.00851***	0.00926***	0.0143***	0.00680***	0.00585	0.0129***
	(0.00169)	(0.00184)	(0.00304)	(0.00354)	(0.00163)	(0.00473)	(0.00483)
Observations	5136	2322	2634	822	1428	972	1494
Joint p	0	0	0	0	0	0	1.17e-09
Sum of lags	0.00626**	0.00636**	0.00536	0.00707*	0.00571*	0.00674	-0.000821
	(0.00304)	(0.00308)	(0.00458)	(0.00365)	(0.00305)	(0.00491)	(0.00466)

Notes: The forest dataset has been constructed from MODIS satellite images, as described in Section 2.2.2.1. It counts the total number of forest cells by year and forest zone. Note that 1000HA = 10 square kilometres. *Oil and Gas Revenue per capita* variable reports the value of per capita revenue from oil and gas extraction at the district-level in US dollars. The regression also includes district and island-by-year fixed effects. The regression also includes the *Number of districts in the original district boundaries* variable. A district split is recorded once the district reports its first own share of oil and gas revenue. The robust standard errors are clustered at the 1990 district boundaries and reported in parentheses. *** 0.01, ** 0.05, * 0.1

3 An Amenity-based Theory of Squatter Settlements

3.1 Introduction

There is a growing global concern about the living conditions of the urban poor, as 924 million people were living in slums in 2001 (UN-Habitat, 2003). That is, about one third of the world's urban population is currently facing poor housing quality, overcrowding, insecure residential status and inadequate access to safe water and sanitation.[69] The majority of these slum dwellers reside in developing countries and often account for a large proportion of their urban population.[70] This situation is set to worsen, as anthropogenic climate change will destroy rural livelihoods in many developing countries (IPCC, 2007a), which will intensify the already rapid urbanization process. Given the limited scope of their formal construction programs (Simha, 2006; WB, 2002), the number of slum inhabitants is thus expected to rise by about 2 billion over the next 30 years.

The main policies suggested to tackle this looming housing crisis involve slum upgrading programs and the provision of greater security of tenure (UN-Habitat (2003), chapter 7). Notwithstanding the beneficial impacts of these policies on market outcomes[71], their consequences can only be fully understood if the analysis is integrated into a spatial context. This aspect has been neglected so far by the academic literature, despite the well-known fact that slum locations are not coincidental. For instance, squats are often intentionally set up on 'low quality' land to lower the risk of eviction.[72] This is a major oversight, as any policy that changes the value of the slum location itself will not only affect the socioeconomic behaviour of its inhabitants, but also the equilibrium in both the formal and informal housing market. The following shows that the resulting distortion casts doubt on the economic efficiency of these policies.

This paper attempts to fill this gap by developing a new theoretical framework, which introduces both an informal housing market and heterogeneous space into the

[69] This is the operational definition of UN-Habitat (UN-Habitat, 2002a,b).

[70] The largest proportion of slum dwellers were recorded in Sub-Saharan Africa and South-Central Asia, where they make up 72% and 58% of the urban population respectively. Moreover, about one third of all urban residents live in slums in the rest of Asia, Northern Africa, Latin America and the Caribbean (UN-Habitat, 2003).

[71] The rationale for the up-scaling of slums through investment in infrastructure, roads and waste management is immediate. In addition, the development literature on property rights makes a strong case for increasing the security of tenure, because of its beneficial impact on health outcomes, labour market participation and access to credit. See Besley (1995b) for a seminal microeconomics paper and Carter and Wiebe (1994), Galal and Razzaz (2001), Carter and Olinto (2003), and Field (2007) for an analysis of informal or *de facto* property rights.

[72] A prominent example are the squatters in Rio de Janeiro, who occupy the steep slopes surrounding the city and are thus frequently exposed to flooding and landslides (Bartone et al., 1994).

standard urban Muth-Mills model.[73] To keep the analysis tractable, I focus on squatter settlements, which are the most important component of the slum housing stock in developing countries.[74] More specifically, I build on the seminal work of Jiminez (1984) by modelling the informal housing market as an integral part of the urban residential equilibrium. In this setup, risk-averse households can either rent in the formal housing market or squat at a lower price. The latter will compensate them for the risk of eviction and make them indifferent between the two housing markets.

However in contrast to Jiminez (1984), I introduce a spatial dimension into the analysis, which is important for two reasons. First of all, it changes the squatter's optimization problem, since the household now maximizes expected utility over two different states *and* locations; in the non-evicted state the household lives in the squat, whereas it has to move to a *different* location in the formal housing market if evicted. This lowers the slope of the squatter's bid rent curve relative to the formal housing market, as she attaches less weight to the characteristics of the squat knowing that she has to relocate with a positive probability. The bid rent curve of the informal housing market, therefore, does not only lie below its formal market counterpart, but it is also flatter.

Secondly, by adding a spatial dimension to the model, I am able to endogenize the location of the squat. Note that squatters can only occupy land that is deemed uninhabitable by the formal housing market, as the latter always outbids them for every lot. Given that squatter settlements are usually established on 'low quality' land, this can be achieved by introducing heterogeneous space into the theoretical framework.[75] In particular, I follow Brueckner et al. (1999) and allow for an exogenous, distance-varying amenity level that is meant to capture the geographical features of the parcel and the availability of basic infrastructure. A steep hillside or a swamp can then be represented by a significant drop in the amenity level over a predetermined range within the city boundaries.

If households derive utility from this amenity level, their bid rent curve will reflect the same discontinuity. Consequently, it is possible that the formal housing market bids so little for the 'low quality' parcel that landowners prefer to leave it vacant for speculative purposes. This outside option is motivated by the observation that inadequate regulation and rent control in developing countries (Jiminez, 1984; Firman, 1997) as well as uncertain land rents (Bar-Ilan and Strange, 1996) limit investment

[73]This theory has originated from the pioneering work of Alonoso (1964), Mills (1967), Muth (1969), and Wheaton (1974).

[74]The other categories are inner city slums, slum estates, and illegal subdivisions (UN-Habitat (2003), chapter 5).

[75]The urban economics literature has provided other explanations for the presence of open space within the city boundaries. For instance, Turner (2005) has provided a preference-based theory of urban sprawl, whereas Katz and Rosen (1981) and Katz and Rosen (1987) have argued that exclusionary municipal zoning may lead to undeveloped land. Both lines of reasoning assume that the land is homogenous *ex ante* and only changes its value due to constraints imposed by preferences or institutions. However, these approaches are not suitable in this context, since our objective is to model open space of intrinsic low value.

opportunities and can, therefore, make it profitable to postpone development.[76] The vacant lot is then available for squatting[77], so that the formal and informal housing market coexist in the residential equilibrium.

The set of residential equilibria becomes more complex, once income heterogeneity is introduced into the analysis. In the standard Muth-Mills model, the poor usually live in the city centre and the rich in the suburbs. Nonetheless, this pattern can be reversed, if the marginal valuation of the amenity level strictly increases with disposable income. If this is the case, the model replicates the residential equilibria found in many developing countries (UN-Habitat (2003), chapter 5). That is, the rich reside in the high amenity core, which is followed by a middle-income neighbourhood. The latter contains the 'low quality' plot and is thus interrupted by a squat. The informal settlement is largely inhabited by low-income households, who care least about the low amenity level. The remaining poor live on the outskirts of the city.

This new theoretical framework can then be used to analyse the impact of the two policy interventions mentioned above. I find that a slum-upgrading program can lead to conflict over the improved lot, as the bid rent increases in both the formal and informal housing market. It is, therefore, possible that a profit-maximizing landowner will then find it optimal to evict the squatters after the program and rent to more affluent, formal market residents. This can be prevented via a titling initiative, which will, however, be inefficient, if the rich value the land more. Consequently, it is important to assess on a case-by-case basis, if a slum-upgrading program is efficient or if resources should alternatively be channelled towards public investment in citywide infrastructure and the provision of housing for the displaced communities.

Before proceeding to the analysis, it is important to discuss briefly the related literature. As has already been mentioned above, no theoretical framework exists that would allows us to study squatter settlements in a spatial context. Most of the urban and regional economics literature has so far focused on developed economies.[78] It has only recently begun to study the effects of rapid urbanization in developing countries on regional and urban development (Venables, 2005; Henderson and Wang, 2005; Henderson and Venables, 2009) or on city growth (Au and Henderson, 2006a,b). Yet none of these papers makes any reference to the informal housing market.

There does exist a small literature on squatting that started with the insightful analysis of Jiminez (1984), which is used as a starting point for this analysis. He was the first to develop a theoretical model of squatting and also provided empirical evidence that tenure insecurity affects the price of the informal housing market; a

[76] The literature on dynamic urban models (Ohls and Pines, 1975; Mills, 1981; Fujita, 1983; Brueckner and Rabenau, 1981; Turnbull, 1988; Braid, 1991) offers a closely related explanation. Their main argument is that if investment decisions are irreversible, it might be efficient to postpone development of a given parcel to ensure that it will better suit future needs.

[77] The landlord tolerates the squat until the profitable economic opportunity arises, since the eviction of the squatters is costly.

[78] As a case in point, there is not a single article on developing countries in the most recent volume of the Handbook of Urban and Regional Economics (Henderson and Thisse, eds, 2004).

result that has been reconfirmed by Friedman et al. (1988) and, more recently, by Field (2007). A series of other papers studies the impact of eviction uncertainty on squatter behaviour (Jiminez, 1985) or endogenized the eviction probability (Hoy and Jiminez, 1991; Turnbull, 2004). All of these papers ignore the spatial context and assume that the existence of the squat is determined exogenously.

Even the most recent contribution by Brueckner and Selod (2008) does not deal with these concerns. They instead focus on the interaction between the community organizers of the squat and formal market residents, as they compete for land within the city. This competition leads to a 'squeezing' of the formal housing market, which is, however, limited by the community organizers, so that eviction is absent in equilibrium. In their model, formalization will lead to a pareto-improvement, because the gains of existing formal market residents are large enough to compensate the losses of the squatters. Since their analysis applies to study long-established squats for which the risk of eviction is absent (Flood, 2006), it should be seen as complementary to this analysis.

The remainder of the paper is structured as follows. Section 3.2 describes the model set up with homogeneous agents in both the formal and informal housing market and derives the associated bid rent curves. The subsequent section introduces income heterogeneity into the model, which allows me to analyse a particularly rich set of equilibrium configurations. Section 3.4 turns to the policy analysis. The final section concludes.

3.2 Basic model

3.2.1 The setup

In the stylised city of the Muth-Mills model, each agent is assumed to work in the central business district (CBD) and commutes from her residential location at x to the CBD.[79] In the context of informal settlements, on might argue that this assumption is not valid, since squats can provide employment opportunities in terms of servicing and public goods provision (Neuwirth, 2005; Davis, 2006). However, even if this is the case for some members of the community, squatter households still spend on average up to 30% of their income on transport (UN-Habitat, 2003). I thus keep this assumption for simplicity.

Furthermore, suppose that there exists an exogenously given function of the amenity level $a(x)$, which declines with the distance x from the CBD. This assumption is motivated by the observation that most cities in developing countries have a well-serviced central area, after which the amenity level falls considerably. This is true for both the 'colonial-style' cities in Africa and South Asia as well as for the 'multi-nucleated' urban areas in Southeast Asia and Latin America (UN-Habitat (2003), chapter 5).

[79] For a comprehensive exposition of the model see, for example, Brueckner (1987) or Fujita (1989).

This downward-sloping amenity function is thus meant to capture the 'quality' of the land at location x, e.g. its geographical features and access to basic infrastructure.

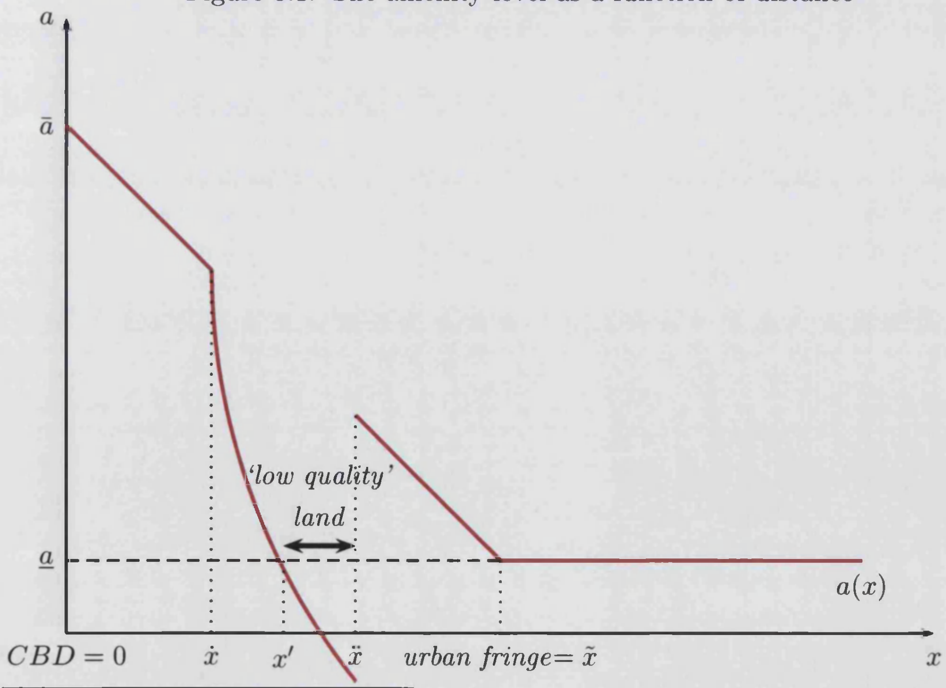
The key assumption is that the amenity level falls significantly over a predetermined range, $\dot{x} \leq x \leq \ddot{x}$, to model 'low quality' land, such as a steep hillside, a floodplain or a garbage dump. At $\ddot{x}$ there is a discontinuous jump in $a(x)$, as the amenity level reverts back to a higher level and then continues its previous downward trend. In fact, there exists a lot of anecdotal evidence that basic infrastructure often runs through squatter settlements to connect the neighbouring areas without servicing the slum itself (UN-Habitat (2003), chapter 2 and 4). At the urban fringe $\tilde{x}$, the amenity level has reached $\underline{a}$ and remains constant thereafter.

For simplicity, let's assume that $a(x)$ can be summarized by equation (9)

$$a(x) = \begin{cases} \bar{a} - bx & \text{if } 0 \leq x \leq \dot{x} \\ \bar{a} - bx - cx^2 & \text{if } \dot{x} \leq x \leq \ddot{x} \\ \bar{a} - bx & \text{if } \ddot{x} \leq x \leq \tilde{x} \\ \underline{a} & \text{if } x \geq \tilde{x}, \end{cases} \tag{9}$$

where $\bar{a} > \underline{a} > 0$, $b > 0$ and $c > 0$. Note that 'low quality' land is defined by $a(x) \leq \underline{a}$.[80] This is achieved if b and c are sufficiently large, i.e. if $c \geq \frac{(\bar{a} - \underline{a}) - b\ddot{x}}{\ddot{x}^2} > 0$ and $b < \frac{\bar{a} - \underline{a}}{\tilde{x}}$. Figure 3.1 below shows an illustration of the amenity level $a(x)$ as it declines with distance from the CBD, where 'low quality' land ranges from x' to $\ddot{x}$.

Figure 3.1: The amenity-level as a function of distance



[80]The amenity level $a(x)$ is allowed to be negative to model polluted or hazardous land. Examples include land along railways (Patel et al., 2002) or garbage dumps (Vliet, 2002).

Finally, it is assumed that households are homogenous and risk averse. Their well-behaved preferences can be represented by $u\big(c, h(x), a(x)\big)$, where c is the numeraire composite good, and $a(x)$ and $h(x)$ are the location-specific amenity level and housing services respectively. Households derive utility from all three components, such that $u_i > 0$ and $u_{ii} < 0$ for $i = c$, $a(x)$, and $h(x)$, where subscripts denote partial derivatives. Following the literature, I also assume that $h'(x) > 0$ and $h''(x) < 0$. The household then maximizes its utility subject to the budget constraints of either the formal or informal housing markets.

Sections 3.2.2 and 3.2.3 now analyse the constrained utility maximization problem in both housing markets in turn.

3.2.2 The formal housing market

This section analyses the household's utility maximization problem in the formal housing market and the associated bid rent curve. Since the following discussion is general, Appendix C.1 provides a mathematical example by solving the problem for the Cobb-Douglas utility function.

The analysis mainly draws on Brueckner et al. (1999), who were the first to introduce an exogenous, distance-varying amenity level into the standard urban model. In this setup, households choose the optimal consumption and housing bundle $\big(c, h(x)\big)$ to maximize $u\big(c, h(x), a(x)\big)$ subject to their budget constraint

$$c + p^F(x)h(x) = y - tx, \tag{10}$$

where $p^F(x)$ is the unit price in the formal housing market, y is income and tx is the linear commuting cost. In addition, it is assumed that all agents enjoy the same utility level $\bar{u}$ at every location within the city boundaries.[81] Since disposable income falls as the household moves further away from the CBD, this implies that the bid rent $p^F(x)$ must decline with x.

The household then determines the optimal bid rent schedule $p^F(x)$ and level of housing services $h(x)$ by solving the uniform utility condition

$$\max_h u\big(y - tx - p^F(x)h(x), h(x), a(x)\big) = \bar{u}. \tag{11}$$

To obtain the slope of the bid rent function in the formal housing market, equation (11) has to be differentiated with respect to distance x and rearranged to yield

$$p^{F\prime}(x) = \frac{u_h h'(x)}{u_c h(x)} - \frac{t}{h(x)} - \frac{p^F(x)h'(x)}{h(x)} + \frac{u_a a'(x)}{u_c h(x)} = -\frac{t}{h(x)} + \frac{v_a a'(x)}{h(x)}, \tag{12}$$

where the second equality uses the first order conditions of the constrained maximiza-

[81] Note that in this context it is not important whether we deal with an open or closed city model, i.e. whether $\bar{u}$ is determined endogenously or exogenously.

tion problem, $u_h = u_c p^F(x)$, and the fact that the marginal valuation of the amenity level (after an optimal adjustment for consumption) is equal to the marginal rate of substitution, i.e. $v_a = \frac{u_a}{u_c}$.

Equation (12) shows that the bid rent curve in the formal housing market is downward sloping for all values of x, if $a'(x) < 0$ and $v_a > 0$ as has been assumed above. Moreover, it is steeper than in the standard urban model, where agents do not derive utility from the amenity level $a(x)$. Brueckner et al. (1999) provide an intuition for this result by arguing that households located at x now have to be compensated for the greater commuting cost to the CBD as well as for the lower amenity level.

Therefore, it is possible that the bid rent offered by the formal housing market is less than the landowner's outside option, if the amenity level $a(x)$ is very low. The landowner, who for simplicity is assumed to be absentee, then prefers to leave the land vacant for speculative purposes.[82] This outside option seems to be particularly pertinent in a developing country context, as imperfect capital markets and inadequate regulation limit investment opportunities (Jiminez, 1984; Firman, 1997) and can thus make it profitable to postpone development, especially if land rents are uncertain (Bar-Ilan and Strange, 1996). Assuming that the present discounted value of the speculation is equal to p_A, location x is left unoccupied for all $x' \leq x \leq \ddot{x}$, where x' is determined by $p^F(x') = p_A$ and $x' < \tilde{x}$[83] (see Figure 3.2 on the next page.[84])

Note that the discontinuity in the bid rent function creates a 'paradox', where the bid rent on the 'low quality' plot can be lower than on locations further away from the CBD. This result is driven by the fact that the household's disutility is larger, the steeper the amenity function $a(x)$. This is in turn reflected in the slope of the bid rent curve, as $\frac{\partial p^{F'}(x)}{\partial a'(x)} = \frac{v_a}{h(x)} > 0$. It is, therefore, possible that the utility loss from the sharp drop in the amenity level is too large to offset the benefits of the lower commuting costs. The bid rent on the 'low quality' plot then has to drop accordingly to ensure that households enjoy the citywide utility level $\bar{u}$.

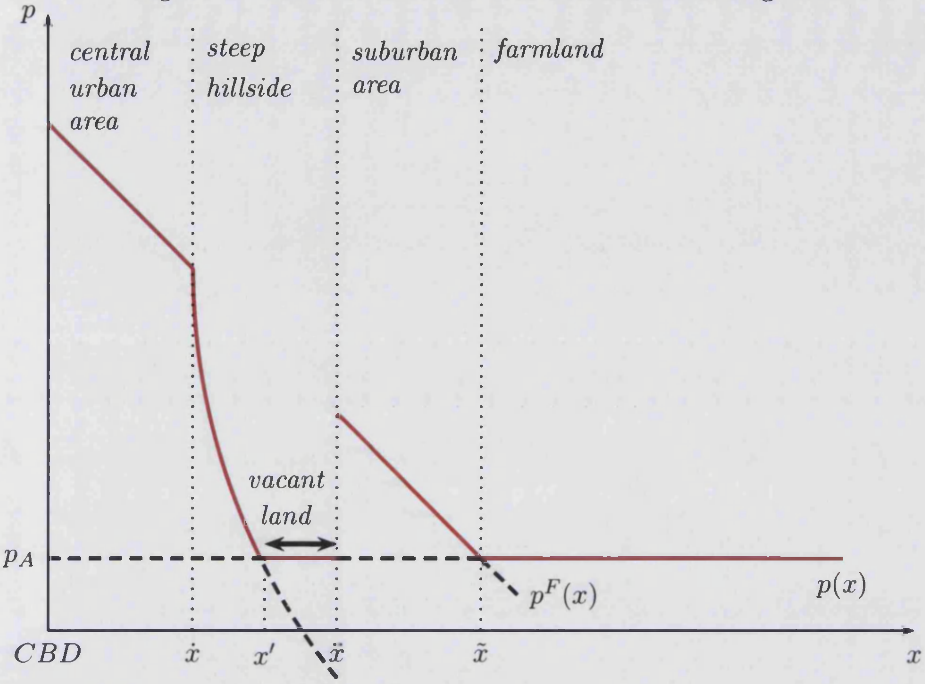
Finally, as in the standard Muth-Mills model, the urban fringe $\tilde{x}$ is determined by the second intersection of the bid rent curve with the outside option, i.e. $p^F(\tilde{x}) = p_A$ where $\tilde{x} > x'$. The equilibrium price schedule $p(x)$ is then the upper envelope of $p^F(x)$ and p_A.

[82] It is usually assumed that squatter settlements only form on public land. However, this is not generally the case (Buckley and Kalarickal, 2006; WB, 2007; Galiani, 2004; DiTella et al., 2007).

[83] Note that there is no vacant land if $p^F(x') = p_A$ for $x' \leq \dot{x}$ or $x' \geq \ddot{x}$. The intersection would then merely determine the urban fringe $\tilde{x}$, since $p^F(x) < p_A$ for all $x \geq x'$.

[84] Note that I have drawn $p^F(x)$ as a linear function for convenience. However, it is clear from equation (12) that this is not necessarily the case, as $p^{F'}(x)$ changes with x as long as $h'(x) \neq 0$ or $a''(x) \neq 0$.

Figure 3.2: The bid-rent curve in the formal housing market



3.2.3 The informal (squatting) housing market

This section now analyses the expected utility maximization problem of the informal housing market and discusses the associated bid rent schedule. However, this time I cannot provide an example in the appendix, since there is no closed-form solution for the Cobb-Douglas utility function or any other standard utility function used in economics. I am only able to solve the problem for the log utility, which is not suitable for this analysis, since it lacks a key property; namely that the marginal valuation of the amenity level changes with income. This will be crucial in determining the residential equilibrium later. Appendix C.2 thus only provides the technical details for the general derivation.

The basic set up is taken from Jiminez (1984), who has been the first to develop a theoretical framework of the squatter's choice problem. More specifically, I adopt two of his key assumptions that establish the links between the formal and informal housing market. First of all, it is assumed that households can choose between being a formal market resident or squatting. As households are assumed to be risk-averse, they are only willing to squat, if they are compensated for the tenure insecurity. Therefore, the formal and informal housing market only coexist in equilibrium, if the unit housing price in the squatter settlement is considerably lower than its formal market counterpart. The size of this risk premium is determined by the uniform utility condition, which has to bind for both housing markets in equilibrium.

However, Jiminez (1984) has ignored the main implication of this result; namely that squatters can only settle on land, which is deemed uninhabitable by the formal housing market, as the latter outbids them for every other location. This has been rectified by the current framework through the introduction of the distance-varying amenity level. The vacant lot, which has been derived in Section 3.2.2 above, now accommodates the squat. It is assumed that the landowner tolerates the informal settlement until a profitable economic opportunity arises, since an eviction is expensive.[85]

The second assumption borrowed from Jiminez (1984) is that squatters are able to find accommodation in the formal housing market if evicted. However, different to his analysis, I introduce a spatial component into the squatter's choice problem. That is, the evicted household has to move from her location x^S in the squat, where $x' \leq x^S \leq \ddot{x}$, to a *different* site x^F in the formal housing market. It should be noted that this second assumption is not crucial for the subsequent results.[86] Yet, segmented housing markets can only be motivated in this framework through rationing by income.[87] I hence postpone any further discussion of this issue until income heterogeneity is introduced in Section 3.3.

The squatter thus maximizes her expected utility over two separate states *and* locations

$$E(u^S) = \pi(x^S)u^E\left(c^E, h^E(x^F), a(x^F)\right) + \left(1 - \pi(x^S)\right)u^N\left(c^N, h^N(x^S), a(x^S)\right), \quad (13)$$

where $\left(c^E, h^E(x^F)\right)$ and $\left(c^N, h^N(x^S)\right)$ are the consumption and housing bundle of the evicted and not evicted state respectively, and $\pi(x^S)$ is the risk of eviction in the squat. It is assumed that the probability of eviction declines with distance from the CBD, as the authorities are more aware of slums closer to the centre (UN-Habitat, 2003). In addition, once income heterogeneity is introduced in the model, I show in Section 3.3.1 that the rich prefer the central locations because of the low commuting cost and high amenity level. The higher eviction probability closer to the CBD could thus also be explained by 'gentrification'.

The budget constraint in the non-evicted state can be described as

$$c^N + p^S(x^S)h^N(x^S) = y - tx^S, \quad (14)$$

where $p^S(x^S)$ is the unit price in the informal housing market. Since the squatter settlement is illegal, $p^S(x^S)$ is not an actual rental payment to a landlord. However, living in squat is not costless. For instance, it could represent a protection fee paid to 'local people of influence', such as the police or the mafia (Lanjouw and Levy, 2002;

[85] As is explained in more detail below, the squatters have to be compensated when they are evicted.

[86] This is true as long as long as the household still faces a downward sloping bid rent curve when evicted.

[87] Alternatively, market segmentation can be introduced, if squatting carries a strong enough stigma. For instance, if formal employers 'redline' workers from undesirable areas in the city, squatting will be unattractive for formal market residents (Boccard and Zenou, 2000; Zenou, 2002). This assumption effectively removes the uniform utility condition and introduce differences in preferences for this kind of stigma across households. The same effect can be modelled more simply by allowing for rationing in income without substantially changing the model.

WB, 2007) or could reflect the cost of contributing to the organization and patrolling of the squat (Jiminez, 1985; Brueckner and Selod, 2008). In addition, Field (2007) has shown that squatters devote a substantial amount of time to securing their tenancy through formal and informal channels. Hence, $p^S(x^S)$ could also be interpreted as an opportunity cost in terms of foregone earnings.

If the squatters are evicted, their ownership cost $p^S(x^S)h^N(x^S)$ is lost, as the housing stock is demolished and they have to move to the new location x^F. The budget constraint in the evicted state can then be written as,

$$c^E + p^S(x^S)h^N(x^S) + p^F(x^F)h^E(x^F) = y - tx^F + F, \tag{15}$$

where $F > 0$ is a lump-sum compensation the household receives upon its eviction. These type of compensations were uncommon during the 1970s and 1980s, when forced evictions were the usual policy response to squatter settlements (Cohen, 1983; Badcock, 1984; Murphy, 1990).[88] However, nowadays evictions are usually more participatory and often involve some form of compensation or resettlement to a formal housing block on the urban fringe (UN-Habitat, 2003). The latter can be captured by this framework, if we assume that households choose the new location x^F, such that their disposable income is the same in both states, i.e. $y - tx^F + F = y - tx^S$. This implies that evicted households essentially choose an accommodation further away from the CBD, as $x^F = x^S + \frac{F}{t} > x^S$. This seems to be a reasonable assumption, also because housing prices and population density are higher closer to the city centre.[89]

The household then determines the optimal bid rent schedule $p^S(x^S)$ and level of housing services $h^N(x^S)$ and $h^E(x^F)$ by solving the uniform utility condition

$$\begin{aligned}\max_{h^E,h^N} \pi(x^S)u^E\big(y - tx^S - p^S(x^S)h^N(x^S) - p^F(x^F)h^E(x^F), h^E(x^F), a(x^F)\big)+ \\ \big(1-\pi(x^S)\big)u^N\big(y - tx^S - p^S(x^S)h^N(x^S), h^N(x^S), a(x^S)\big) = \bar{u},\end{aligned} \tag{16}$$

where $x^F = x^S + \frac{F}{t}$.

Similar to the formal housing market, the slope of the squatter's bid rent function is obtained by totally differentiating equation (16) with respect to x^S. Appendix C.2

[88] In fact, Jiminez (1984) adds a fixed cost to the budget constraint in the evicted state, because households incur a penalty and moving costs.

[89] This assumption can be further rationalized if we argue that households are of a fixed size. Thus, it is possible that the bid price in the centre is so high that the evicted household might not be able to afford sufficiently large accommodation, even if they use the compensation and take advantage of the lower commuting costs. This argument gains more weight once we introduce income heterogeneity and find that the rich outbid the poor in the centre and that the poor are those that live in the squat.

provides the detailed steps, which yield

$$\begin{aligned} p^{S\prime}(x^S) = &-\frac{t}{h^N(x^S)} + \frac{a'(x^S)}{h^N(x^S)}\left(\frac{\pi(x^S)u_a^E + \big(1-\pi(x^S)\big)u_a^N}{\pi(x^S)u_c^E + \big(1-\pi(x^S)\big)u_c^N}\right) \\ &- p^{F\prime}(x^F)\frac{h^E(x^F)}{h^N(x^S)}\frac{\pi(x^S)u_c^E}{\pi(x^S)u_c^E + \big(1-\pi(x^S)\big)u_c^N} \\ &+ \frac{\pi'(x^S)}{h^N(x^S)}\frac{u^E - u^N}{\pi(x^S)u_c^E + \big(1-\pi(x^S)\big)u_c^N} \end{aligned} \tag{17}$$

The interpretation of the first two terms is analogous to what has been argued in Section 3.2.2. That is, households have to be compensated for the higher commuting cost and lower amenity level, the further they move away from the CBD.[90] However, it is important to note that both terms are smaller in absolute value than their corresponding expressions in equation (12). This is due to the fact that unit housing prices in the squat are lower than in the formal housing market for every x within the city boundaries. As a consequence, households in the non-evicted state have more disposable income than a formal market resident, who would bid for the same lot x^S. Therefore, the squatter consumes more housing services if she is not evicted, because both the numeraire composite good and housing services are assumed to be normal goods. This implies that the first term in equation (17) is smaller in absolute value that its formal market counterpart.

The same is true for the first part of the second term. To determine the size of the numerator of the second ratio, it is helpful to notice that formal market residents enjoy the same marginal valuation of the amenity level u_a at x^S as a non-evicted squatter. However, once the latter is evicted, she moves to a higher level amenity plot x^F, such that $u_a^E < u_a^N = u_a$ because of diminishing marginal returns and, thus, $\pi(x^S)u_a^E + \big(1-\pi(x^S)\big)u_a^N < u_a$. The overall size of the denominator is less likely to change; non-evicted households consume more of the composite good relative to a formal market resident at x^S because of lower unit housing prices. However, consumption is lower in the evicted state, since they have lost their investment in the squat. The second term in equation (17) should thus also be smaller in absolute magnitude.

Consequently, the first two terms yield a flatter slope than in the formal housing market. This effect is reinforced by two new terms, which both enter positively into the expression. The third term is positive, because $p^{F\prime}(x^F) < 0$ for all x. It captures the fact that squatter households are tied less to a given location; they know that they have to move to x^F with a positive probability. The squatters thus weigh the characteristics of the squat relative to the scenario they encounter when evicted. For instance, if the bid rent curve in the formal market gets steeper, the unit housing price is lower in the evicted state. To maintain the uniform utility level $\bar{u}$, the unit housing

[90] As before, this assumes that $a'(x) < 0$, $u_a^E > 0$, and $u_a^N > 0$.

price in the informal market hence falls less with distance.

The fourth and final term is also positive, since $\pi'(x^S) < 0$ for all x^S and $u^E < u^N$, because the household has to invest twice into the housing stock in the evicted state. The intuition of this result is that the decline in the risk of eviction with distance from the CBD partly offsets the utility loss due to the lower amenity level and higher commuting cost. This further lowers the slope of the bid rent curve. Furthermore, note that as the probability of eviction declines with distance from the centre, the bid rent curves of the formal and informal housing market eventually intersect once the threat is removed.

Given that the first two terms in equation (17) are negative and the last two positive, the overall sign of the slope of the squatter's bid rent curve is ambiguous. As has already been explained at the beginning of this section, I am, unfortunately, not able to sign this expression, as the common utility functions used in economics do not have a closed form solution. Nonetheless, it should be noted that the bid rent curve in the informal housing market is a 'shadow' bid rent curve, which is only relevant once a parcel is left vacant by the formal housing market. Since this is only the case for the low level amenity plot, where $a'(x) << 0$, it is likely that the bid rent curve slopes downwards. Summarizing yields:

> **Result 1.** *The bid rent curve of the informal housing market lies below its formal market counterpart and is flatter. The two curves intersect at the point where the risk of eviction is equal to zero.*

3.3 The extended model

This section introduces income heterogeneity into the model. I can thus analyse how the location by income of the standard Muth-Mills model changes in the presence of a discontinuous amenity function and an informal housing market.

I follow the literature by capturing income heterogeneity not only through differences in income y, but also in the value of time t. This assumption is based on the observation that skill-intensive jobs usually command a higher wage, so that richer households have a higher opportunity cost of leisure. I introduce a high- (H), a middle- (M) and a low-income (L) class, where $y_H > y_M > y_L$ and $t_H > t_M > t_L$. Furthermore, note that although the value of time t rises with y, disposable income can still be ranked, i.e. $y_H - t_H x > y_M - t_M x > y_L - t_L x$ for every x.

The above analysis has shown that each income group has a bid-rent curve for the formal and informal housing market respectively and that $p_j^F(x) > p_j^S(x)$ for all x and $j = H, M$, or L. To determine the residential equilibrium, I first have to study the location by income in the formal housing market (Section 3.3.1) and the associated residential equilibria (Section 3.3.2). If a plot of land is then left vacant in equilibrium, a squat can be established and the informal housing market will come into existence. I can then proceed to do a similar analysis for the squatters and derive the final residential equilibrium configuration in the urban housing market (Section 3.3.3).

3.3.1 The location by income in the formal housing market

In general, the location by income is determined by comparing the slopes of the bid rent functions of income group j and k at the intersection of the curves. Income group j outbids group k for all locations closer to the CBD, if $p_j^{F\prime}(\hat{x}) < p_k^{F\prime}(\hat{x})$ at the intersection $\hat{x}$, i.e. if income group j's bid rent curve is more negatively sloped. One complication that arises in my extended version of the standard urban model is that the bid rent schedules are not continuous. Therefore, several intersections are possible when comparing two income groups. However, if the bid rent curves can still be ranked in terms of their relative steepness, the number of possible configurations is reduced.

The following shows which assumptions are necessary to ensure such a ranking. Let us begin by analyzing the difference between the slopes of the high- and middle-income group, $\triangle_{HM}$, at the intersection $\hat{x}$ of their bid rent curves $p_H^F(x)$ and $p_M^F(x)$:

$$\begin{aligned}\triangle_{HM}(\hat{x}) &\equiv p_H^{F\prime}(\hat{x}) - p_M^{F\prime}(\hat{x}) \\ &= -\left[\frac{t_H}{h_H(\hat{x})} - \frac{t_M}{h_M(\hat{x})}\right] \\ &\quad + a'(\hat{x})\left[\frac{v_a\big(y_H - t_H\hat{x}, p_H^F(\hat{x}), a(\hat{x})\big)}{h_H(\hat{x})} - \frac{v_a\big(y_M - t_M\hat{x}, p_M^F(\hat{x}), a(\hat{x})\big)}{h_M(\hat{x})}\right]\end{aligned} \tag{18}$$

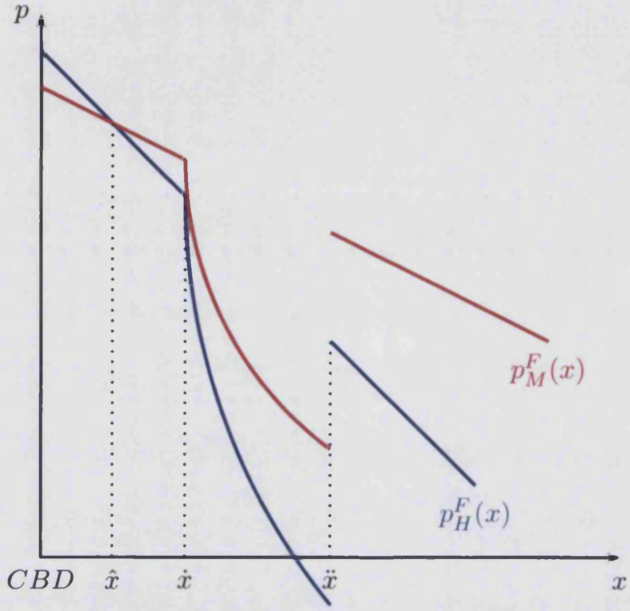
As in Brueckner et al. (1999), I assume that the conventional forces favour the location of the rich in the suburbs, i.e. $\triangle_{HM} > 0$ if $a'(\hat{x}) = 0$. This effect can be achieved if the time preference rate rises less with disposable income relative to the lot size, such that $\frac{t_H}{h_H(\hat{x})} - \frac{t_M}{h_M(\hat{x})} < 0$. However, as soon as we allow for $a'(\hat{x}) < 0$, $\triangle_{HM}$ can be negative. This is the case if the amenity advantage of the city centre is sufficiently large, as is the case in this model, and the rich care more about the amenity level than the poor. To ensure the latter, Brueckner et al. (1999) assume that the ratio of the marginal valuation of the amenity level to the lot size strictly increases with disposable income, such that $v_a\big(y_H - t_H\hat{x}, p_H^F(\hat{x}), a(\hat{x})\big) - v_a\big(y_M - t_M\hat{x}, p_M^F(\hat{x}), a(\hat{x})\big) > 0$. Appendix C.3 shows that this the case for the Cobb-Douglas utility function.

If these assumptions are satisfied, then $\triangle_{HM} < 0$. The number of possible intersections is then determined by where the two bid rent curves cross for the first time. That is, if the first intersection occurs on $0 \le \hat{x} \le \dot{x}$, no second crossing exists. This is due to the fact that the slope of the bid rent curve is increasing over the low level amenity plot and reverts back to the initial slope for all $x \ge \ddot{x}$ (Figure 3.3a on the next page). A less clear-cut prediction can be derived if the bid rent curves intersect for the first time on the steep hillside, i.e. if $\dot{x} \le \hat{x} \le \ddot{x}$. In this case, a second crossing

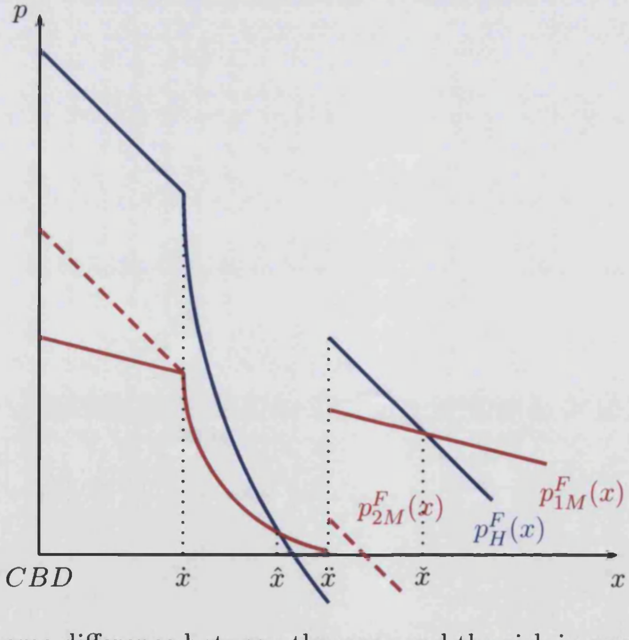
$\breve{x}$ is feasible, where $\ddot{x} \leq \breve{x}$ (Figure 3.3b below).[91]

Figure 3.3: Possible intersections of the bid-rent curve for the high- and middle-income household

(a) Only one possible intersection



(b) One or two possible intersections



Since the income difference between the poor and the rich is even larger, it is then also the case that $\triangle_{HL} < 0$ under these assumptions, where $\triangle_{HL}$ is the difference in slopes of the high- and low-income bid rent schedules at their intersection. That is, the rich outbid both the middle- and the low-income class for locations closer to the

[91]This depends on the slope and intercept of $p_H^F(x)$ and $p_M^F(x)$ and thus on the characteristics of the income groups and the assumed functional form of $a(x)$ and $u(c, h(x), a(x))$.

CBD, i.e $\triangle_{HL} < \triangle_{HM} < 0$. A similar argument implies that the middle-income class always offers a higher bid rent for locations closer to the city centre than the poor, as $\triangle_{ML} < 0$. To summarize:

Result 2. *Even in the presence of a discontinuous amenity function, the bid rent curves can be ranked by income as long as the marginal valuation of the amenity level increases more with disposable income than the lot size. Different to Brueckner et al. (1999), it is possible that bid rent curves intersect twice, if they cross for the first time on the low level amenity plot. If the amenity advantage of the city centre is sufficiently large, the rich always locate closer to the CBD after each intersection with the bid rent curve of the middle-income group, whereas the opposite is true for the poor.*

3.3.2 Possible residential equilibria in the formal housing market

This section discusses a set of possible residential equilibria in the formal housing market. Section 3.3.1 has just shown that the bid rent curves of the three income groups can intersect more than once in this theoretical framework. This depends on the slope and intercept of the bid rent schedules and, hence, on the characteristics of the income groups and the assumed functional forms of the amenity curve $a(x)$ and utility function $u(c, h(x), a(x))$.

To keep the analysis as general as possible, I make no additional assumptions. Instead I discuss one particularly rich residential equilibrium, which is consistent with the 'colonial-style' cities and 'multi-nucleated' urban areas found in many developing countries. Figure 3.4 on the next page shows the equilibrium without a vacant plot, whereas the configuration in Figure 3.5 has the potential for squat formation.

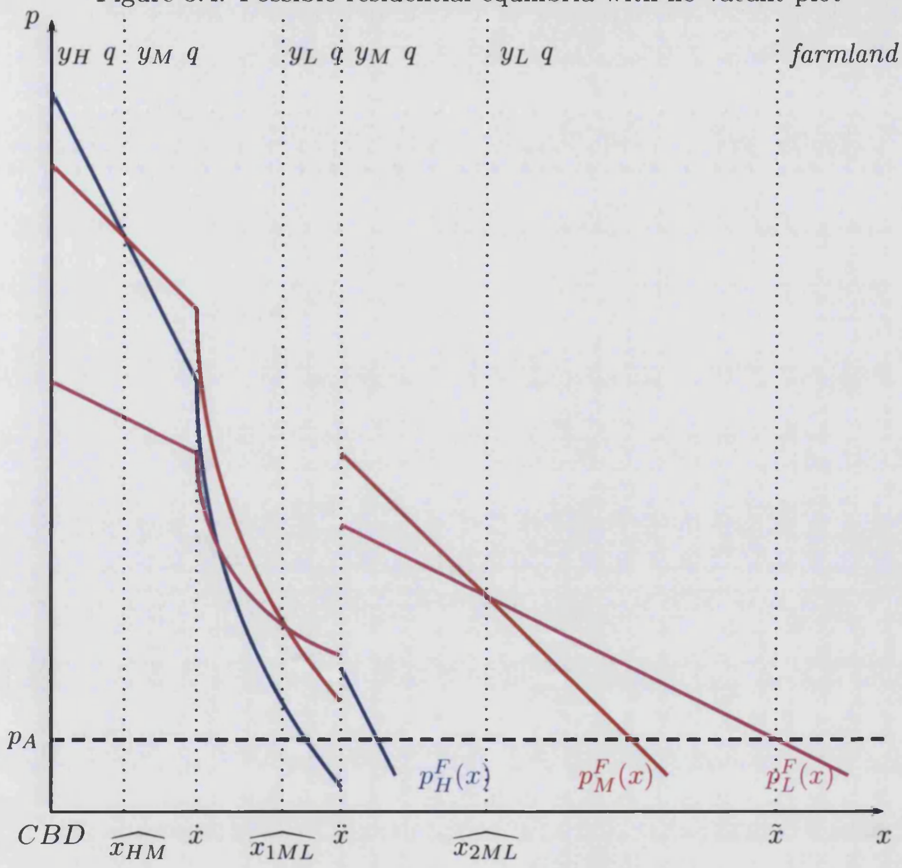
In both scenarios, the rich only occupy a small neighbourhood close to the CBD, y_H q, as the intersection of the high- and middle-income bid rent curves at $\widehat{x_{HM}}$ occurs on $0 \leq \widehat{x_{HM}} \leq \dot{x}$. We know from the analysis in Section 3.3.1 that there is no other intersection, since the $p_H^{F\prime}(x) < p_M^{F\prime}(x)$ for all $x > \widehat{x_{HM}}$. Such a small neighbourhood for the rich is sufficient, if the high-income class is much smaller in comparison to the other two income groups. This seems to be a realistic assumption, especially in a developing country context.

In addition, I allow for a large income difference between the middle- and low-income groups, i.e. $\triangle_{ML} << 0$. Consequently, their bid rent curves intersect twice at $\widehat{x_{1ML}}$ and $\widehat{x_{2ML}}$. We know from the above that the poor choose to locate further away from the CBD after each crossing. Finally, the urban fringe $\tilde{x}$ is again determined by the intersection of the outside option p_A with the bid rent curve of the low-income group $p_L^F(x)$.

The existence of a vacant plot of land is then determined by the outside option. For instance, Figure 3.4 on the next page illustrates a residential equilibrium, in which

p_A is so low that all the land within the city boundaries is used by the formal housing market. That is, the rich locate nearest to the CBD, the middle-class in the central urban area and the poor on the low-level amenity plot. After the discontinuous jump in amenity level at $\ddot{x}$, the middle-income class again outbids the poor and locate in the suburbs, whilst the remainder of the low-income group moves to the periphery.

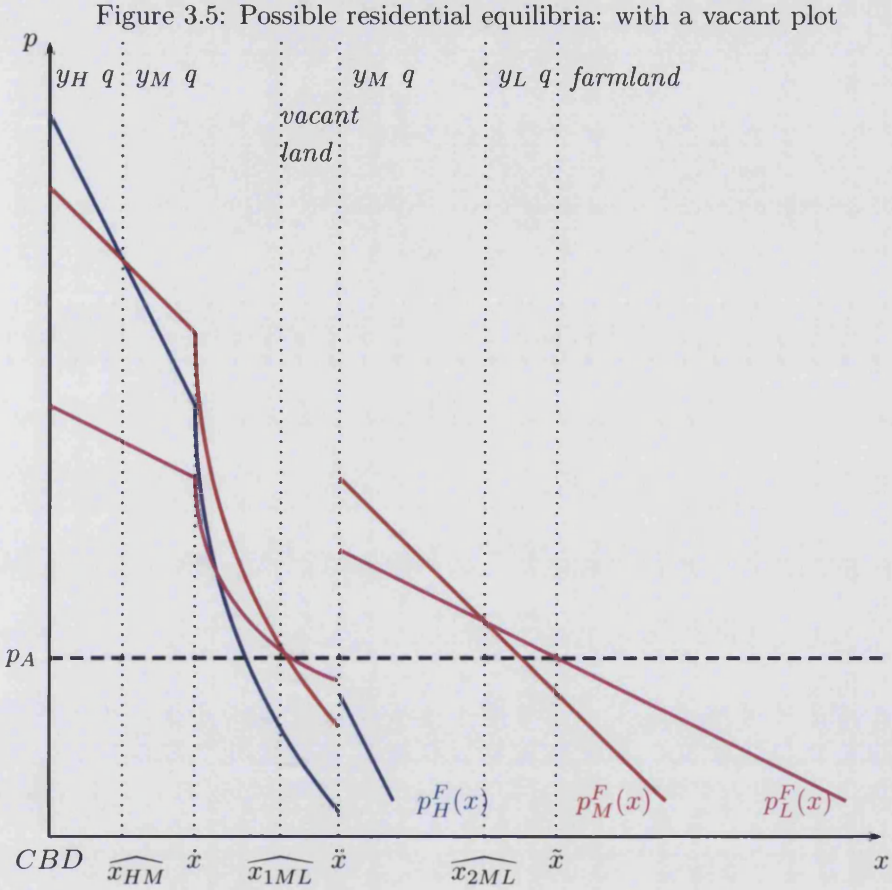
Figure 3.4: Possible residential equilibria with no vacant plot



As the present discounted value of the outside option rises, the landlord finds it less and less profitable to rent out the land to the poor on the steep hillside. For example, p_A has risen so high in Figure 3.5 on the next page that the poor solely live on the outskirts of the city. The 'low quality' plot is left vacant and thus becomes available for squatting.

If p_A increases even further such that $\tilde{x} \leq \widehat{x_{2ML}}$, the low-income class is crowded out of the formal housing market altogether. In that case, the entire low-income group has to find accommodation in the informal housing market. If the squatters are then evicted from the 'low quality' plot, they have to find other illegal accommodation. One possibility would be to allow for illegal subdivisions on the urban fringe, which is a phenomenon often observed in developing countries (UN-Habitat (2003), pp.83f). Since in this case the landowners rent out their farmland without having official permission, there is no longer a threat of eviction. They only have to pay a fine, which is essentially a mark-up on the rent.

Figure 3.5: Possible residential equilibria: with a vacant plot



Note that if the two housing markets are segmented, it has to be the case that the utility level of the squatters is lower than in the formal housing market. Otherwise the high- and middle-income class would have an incentive to squat. Therefore, the only difference introduced by market segmentation is that policy changes will affect the utility of the poor and the two upper classes separately; a point I discuss briefly in Section 3.4.1.1.

For now I continue to focus on the case of no market segmentation (Figure 3.5), because it allows me to study the location by income in the informal housing market to which I now turn.

3.3.3 The location by income in the informal housing market and the final residential equilibrium

To derive the residential equilibrium with squatters, I proceed as in Section 3.3.1 and first determine the location by income in the informal housing market. I then return to Figure 3.5 and discuss the final equilibrium configuration with squat formation.

I begin by signing the difference between the slopes of bid rent curve of the high- and middle-income class squatters, $p_H^S(x)$ and $p_M^S(x)$ respectively, at their intersection

$\widehat{x^S}$. Note that this intersection does not necessarily have to be on the land available for squatting, because the informal housing market can theoretically bid for every location within the city boundaries. That is, there is a 'shadow' bid rent curve of the informal housing market for all x, but we only observe it for $x' \leq x^S \leq \ddot{x}$, as formal market residents outbid the squatters for all other locations.

$$\begin{aligned}
\triangle_{HM}^{S} \equiv p_H^{S\prime}(\widehat{x^S}) - p_M^{S\prime}(\widehat{x^S}) = &-\left[\frac{t_H}{h_H^N} - \frac{t_M}{h_M^N}\right] \\
&+ \left[\frac{a'}{h_H^N} \frac{\pi c_{Ha}^E + (1-\pi) c_{Ha}^N}{(\pi u_{Hc}^E + (1-\pi) u_{Hc}^N)} - \frac{a'}{h_M^N} \frac{\pi c_{Ma}^E + (1-\pi) c_{Ma}^N}{\pi u_{Mc}^E + (1-\pi) u_{Mc}^N}\right] \\
&- \left[p_H^{F\prime} \frac{h_H^E}{h_H^N} \frac{\pi u_{Hc}^E}{\pi u_{Hc}^E + (1-\pi) u_{Hc}^N} - p_M^{F\prime} \frac{h_M^E}{h_M^N} \frac{\pi u_{Mc}^E}{\pi u_{Mc}^E + (1-\pi) u_{Mc}^N}\right] \\
&+ \left[\frac{\pi'}{h_H^N} \frac{u_H^E - u_H^N}{\pi u_{Hc}^E + (1-\pi) u_{Hc}^N} - \frac{\pi'}{h_M^N} \frac{u_M^E - u_M^N}{\pi u_{Mc}^E + (1-\pi) u_{Mc}^N}\right],
\end{aligned} \tag{19}$$

where the subscript $j = M, H$ indicates the income group, e.g. the marginal utility of consumption in the evicted state of the middle-income class is u_{Mc}^E. Note that the dependencies on $\widehat{x^S}$ are suppressed in the main expression for ease of notation.

The first two terms in this expression are analogous to what has been obtained for the formal housing market in equation (18). Therefore, under the maintained assumptions, they are both negative. This would suggest a location by income, which is similar to the formal housing market, where the rich locate in the city centre and the poor on the outskirts. The interpretation of the last two terms is more ambiguous. Nonetheless, I can still make some conjectures on the overall sign of the difference in slopes.

First of all, I am able to unambiguously sign the third expression, if the ratio of housing services in the two states and the share of marginal utilities are independent of income.[92] The third term then enters positively into the expression, as I have already shown in Section 3.3.1 that $\triangle_{HM} < 0$. If this is not the case, diminishing marginal returns should increase the marginal utility of consumption in the evicted state disproportionately more for the middle-income class relative to the rich. Similarly, we would expect that housing services are more evenly spread across states for more affluent households. Both of these effects would work in the opposite direction and render the third term less positive and, thus, more in line with the results of the formal housing market.

Secondly, the sign of the fourth and final expression is ambiguous. On the one hand, it could be argued that the overall expression is negative, because the term for the middle-income class is most likely to be larger in absolute value. First of all, it is assumed that housing services are a normal good, such that $\left|\frac{\pi'(\widehat{x^S})}{h_M^N(\widehat{x^S})}\right| > \left|\frac{\pi'(\widehat{x^S})}{h_H^N(\widehat{x^S})}\right|$. Secondly, the difference in utilities in the evicted and non-evicted state should be

[92]This is true for the log and the CES utility function.

larger for the middle-income households, since the utility function is concave. On the other, diminishing marginal returns imply that the expected marginal utility of consumption is lower for the high-income households. This in turn increases the size of the first term. The overall sign is hence ambiguous.

As a consequence, I am not able to unequivocally sign the difference in slopes in the informal housing market. However, I assume in the following that the ranking of the bid rent curves mimics the ordering in the non-squatting housing market. Otherwise, the location by income could be reversed, such that the rich would choose to live close to the CBD in the formal housing market, but prefer to squat in the periphery. Such a model feature is neither realistic nor consistent with monotonic preferences and is hence ruled out.

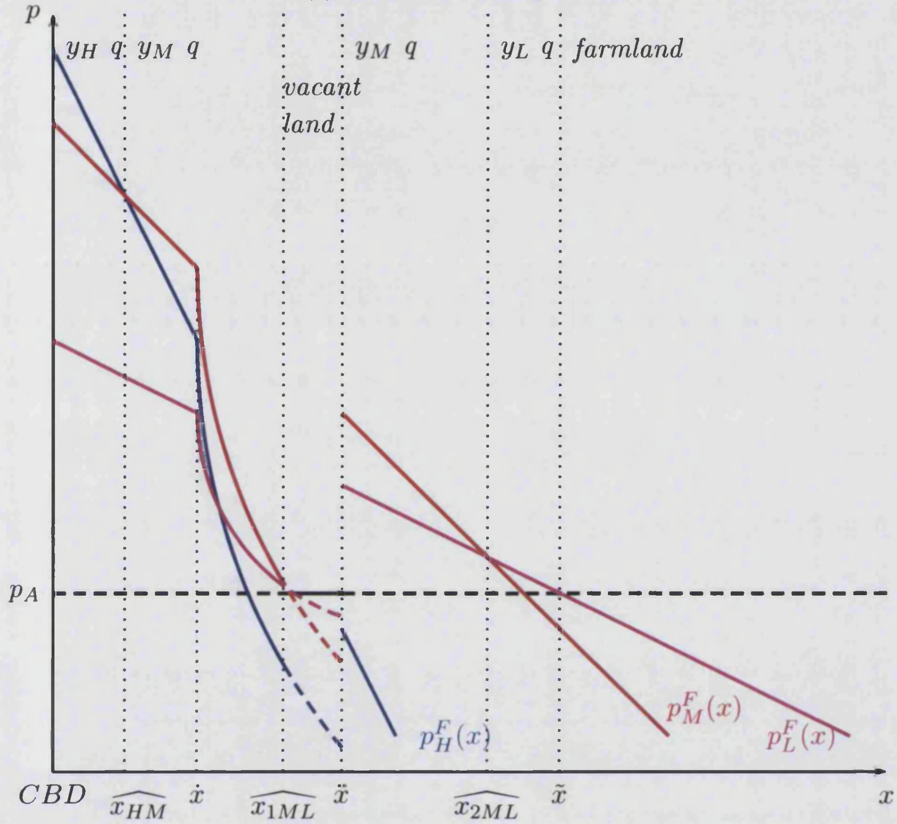
Let us now return to the residential equilibrium with squat formation, which has been reproduced in Figure 3.6a on the next page. Note that this figure shows the bid rent curves for the formal housing market as dashed lines over the 'low quality' plot, $\widehat{x_{1ML}} \leq x \leq \ddot{x}$, which is left vacant for speculative purposes. Figure 3.6b depicts the bid rent schedules for the informal housing market. They are only shown as solid lines over the vacant land, as they are outbid for every other location x by the formal housing market.

I would now like to analyse which income groups are most likely to become squatters on the 'low quality' plot left vacant by the formal housing market. First of all, it should be noted that the rich are least likely to join the squatter community, since they suffer the largest utility loss on the 'low quality' plot. In this particular example, their formal market bid rent curve $p_H^F(x^S)$ is already so steep over the potential squat, $\widehat{x_{1ML}} \leq x^S \leq \ddot{x}$, that their bid rent curve for the informal housing market $p_H^S(x^S)$ is largely negative for most of the interval. We can, hence, exclude the rich from the analysis, as the bid rent has to be non-negative by definition.

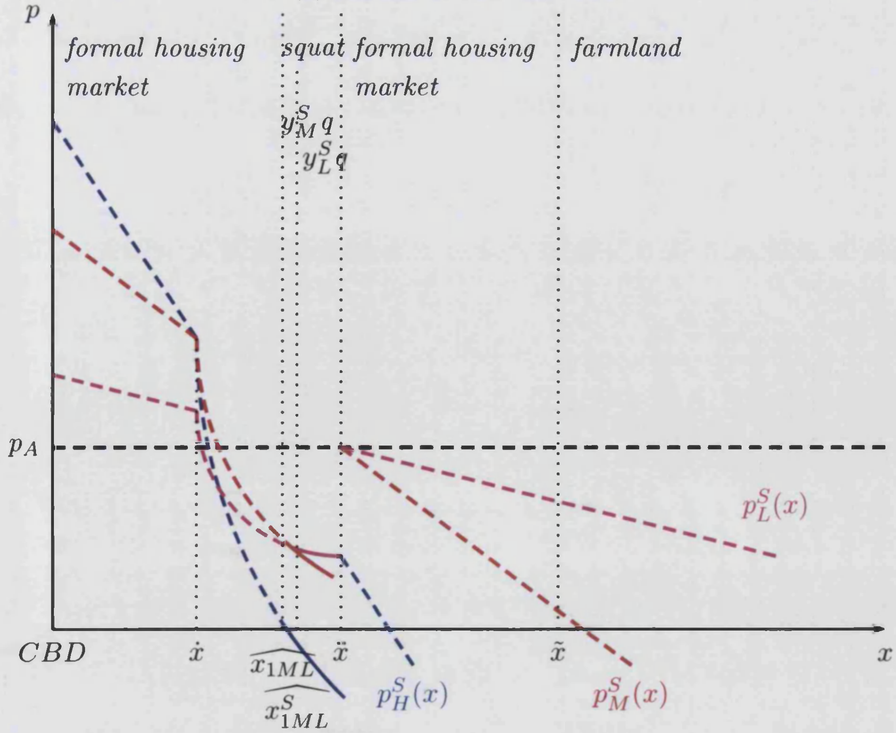
Therefore, it is sufficient to compare the squatter bid rent curves of the middle- and low-income group, $p_M^S(x^S)$ and $p_L^S(x^S)$ respectively. In this equilibrium, their formal market counterparts intersect for the first time on the steep hillside. It is then likely that the bid rent curves for the two squatter groups will also cross over the same range, since both have a flatter slope. It is thus possible that a small fraction of the vacant land will be occupied by squatters of the middle-income class. In this case, they occupy $\widehat{x_{1ML}} \leq x \leq \widehat{x_{1ML}^S}$. In fact, there is some anecdotal evidence that this is often the case in developing countries (UN-Habitat (2003), p.66f). Yet, if the income difference between the two groups is sufficiently large, most of the squat is occupied by the poor, who suffer least from the low amenity level and thus have the flattest bid rent curve. This is also shown in Figure 3.6b, where the poor live on $\widehat{x_{1ML}^S} \leq x \leq \ddot{x}$.

Figure 3.6: The final residential equilibrium with squat formation

(a) The formal housing market



(b) The informal housing market



3.4 Policy analysis

The main aim of this paper was to develop an analytical framework that explains under which conditions squatter settlements arise in the standard urban model. This set up can now be used to study the effects of the two main policies that are usually suggested to improve the livelihood of squatters, namely slum-upgrading programs and the provision of greater security of tenure (UN-Habitat (2003), chapter 7). The following analyses the impact of each policy on the residential equilibrium and discuss their economic efficiency.

3.4.1 The effect of a slum-upgrading program on the residential equilibrium

This section carries out a comparative statics exercise to study how a slum-upgrading program affects the level and slope of the bid rent curves in the formal and informal housing markets and, thus, the urban residential equilibrium. I first discuss how the policy is modelled in this framework and then analyse the effects on the formal and informal housing market in turn. Appendix C.4 provides the calculations for the formal housing market using the Cobb-Douglas utility function, whereas Appendix C.5 derives the general solution for the squatters.

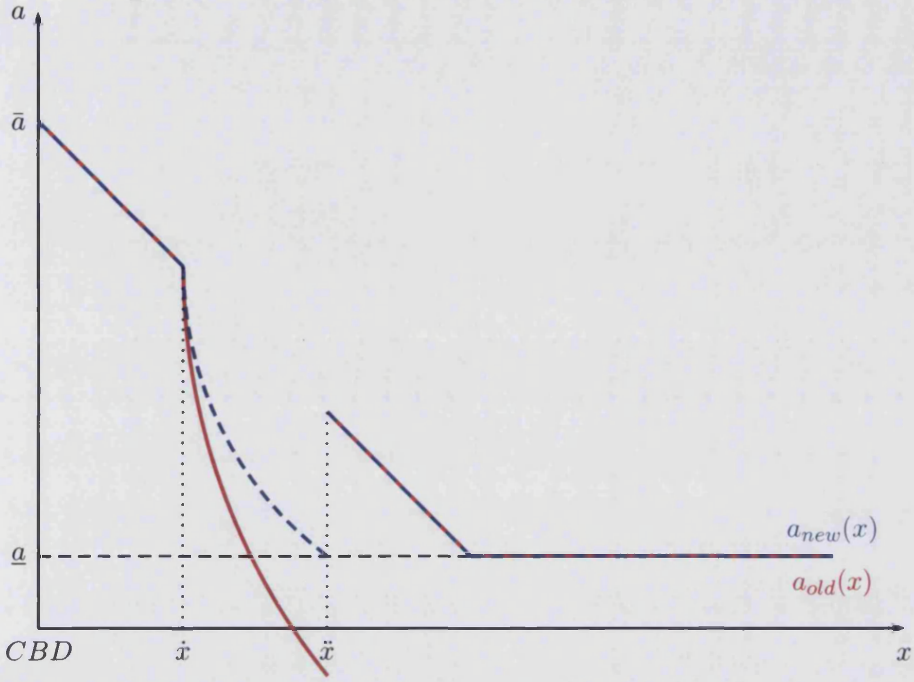
In the past, urban development programs have mainly concentrated on improvements of the housing stock and the physical environment. More recently, the focus has shifted towards *in situ* slum-upgrading programs that usually involve slum-level investment in water and sanitation infrastructure, and the provision of electricity, waste management and better access to roads. However, within this model the effect of both policies can be captured by the same increase in the amenity level on the steep hillside, i.e.

$$a_{new}(x) = \begin{cases} \bar{a} - bx & \text{if } 0 \le x \le \dot{x} \\ \bar{a} - b\dot{x} - c_{new}x^2 & \text{if } \dot{x} \le x \le \ddot{x} \\ \bar{a} - bx & \text{if } \ddot{x} \le \text{x} \le \tilde{x} \\ \underline{a} & \text{if } x \ge \tilde{x}, \end{cases} \tag{20}$$

where $c_{new} < c$ in equation (9). Note that the discontinuity in the amenity level is removed completely, if $c_{new} = 0$. Therefore, a slum-upgrading program increases the amenity level $a(x)$ available on the steep hillside, $\dot{x} \le x \le \ddot{x}$, *and* the slope of its curve over the same range. (Figure 3.7 on the next page).

I begin the analysis by studying how these two changes affect the bid rent curve of income group j in the formal housing market. As has already been shown in Section 3.2.2, the optimal bid rent is determined by the uniform utility condition $\max_{h_j} u\big(y_j - t_j x - p_j^F(x)h_j(x), h_j(x), a(x)\big) = \bar{u}$. By using the implicit function theorem, I can then determine the change in the bid rent curve in response to an increase in

Figure 3.7: An improvement in the amenity level for $\dot{x} \leq x \leq \ddot{x}$



the amenity level

$$\frac{\partial p_j^F(x)}{\partial a(x)} = -\frac{u_{aj}}{-u_{cj}h_j(x) + u_{hj}\frac{\partial h_j(x)}{\partial p_j^F(x)}} > 0, \tag{21}$$

since the own price effect is negative, $\frac{\partial h_j(x)}{\partial p_j^F(x)} < 0$, and the marginal utilities are positive, $u_{aj} > 0$ and $u_{cj} > 0$. This result is highly intuitive, because it implies that households are willing to bid more for locations with a higher amenity level, all other things being equal.

The slum-upgrading program also changes the slope of the bid rent curve in the formal housing market through an increase in the amenity level *and* a flatter slope of the amenity function $a_{new}(x)$. First of all, the increase in the amenity level itself lowers the slope of the bid rent curve, i.e. $\frac{\partial p_j^{F'}(x)}{\partial a(x)} = v_{aaj}\frac{a'(x)}{h_j(x)} > 0$ since the utility function exhibits diminishing returns, $v_{aaj} < 0$, and $a'(x) < 0$. The intuition for this result is that households have to be compensated more for a decline in $a(x)$, the lower its initial level. A slum-upgrading program thus leads to a flatter bid rent curve through increasing the overall quality of the land.

This effect is further reinforced, because changes in the slope of the amenity function directly translate into a change in the bid rent schedule, as $\frac{\partial p_j^{F'}(x)}{\partial a'(x)} = \frac{v_{aj}}{h_j(x)} > 0$. Again this is highly intuitive, since households no longer have to be compensated for a sharp decline in the amenity level by a lower rent. Finally, it is important to note that this effect is larger for richer households, because it has been assumed that $\frac{v_{aj}}{h_j(x)}$ is increasing in disposable income. As a consequence, the effect on the bid rent curve of the high-income group should be largest, whereas we should observe the smallest increase for the poor.

These results are summarized in Figures 3.8a and 3.8b on the next page, which show the residential equilibrium in the formal housing market before and after the policy change respectively. The response illustrated in Figure 3.8b is extreme, as the formal housing market would now want to occupy all of the land within the city boundaries. Nonetheless, it is generally the case that an improvement in the amenity level reduces the size of the plot, which is deemed uninhabitable by the formal housing market. That is to say, the program will lead to conflict over at least parts of the land occupied by the squatters.

The squatters themselves also value the land more after the program, as analogous results can be obtained for the informal housing market (for more details see Appendix C.5). This is not surprising, since households in both markets derive the same utility from an increase in the amenity level. Yet, even if the squatters are now willing to pay more for the improved parcel, we know from our earlier analysis that they are always outbid by the formal housing market, if the latter takes an interest in the same plot.

The key question thus becomes under which conditions the squatters are allowed to remain on the improved land. Moreover, we would like to know if such an equilibrium configuration is efficient. The answer ultimately depends on the institutional framework in which the investment has been made. Sections 3.4.1.1 and 3.4.1.2 now in turn discuss the consequences of a private and public investment in the amenity level on the residential equilibrium.

3.4.1.1 Scenario 1: private investment into the amenity level of the squat

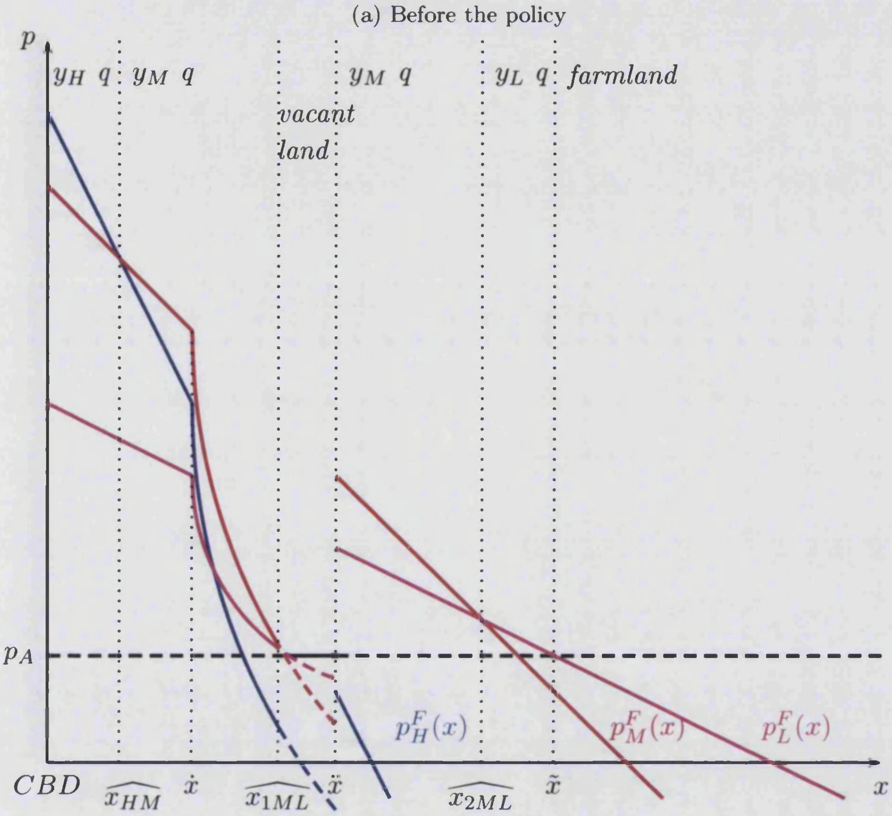
This section discusses how the residential equilibrium changes when a private investment is made into the amenity level of the squat. As has already been argued in Section 3.2.2, the landowner prefers to leave the land vacant for speculative purposes if the bid rent offered by the formal housing market is particularly low. A private landowner thus only improves the amenity level of the squat if a profitable economic opportunity arises, e.g. if there is an unexpected increase in the urban population or an industry has signalled an interest in the parcel. It is assumed for simplicity that the improved lot is sufficient to accommodate this new demand.

The entire squatting community is then evicted and has to move to the low-income neighbourhood on the outskirts. Note that the effect of this policy on the residential equilibrium is essentially identical to an increase in the population of the poor; previously only a fraction π of the squatters moved to the low-income neighbourhood, whereas now the entire community has to relocate there. Such a population increase has already been analysed in the standard urban model (see Fujita (1989), chapter 4 for an extensive analysis) and I thus only briefly discuss the main result here.

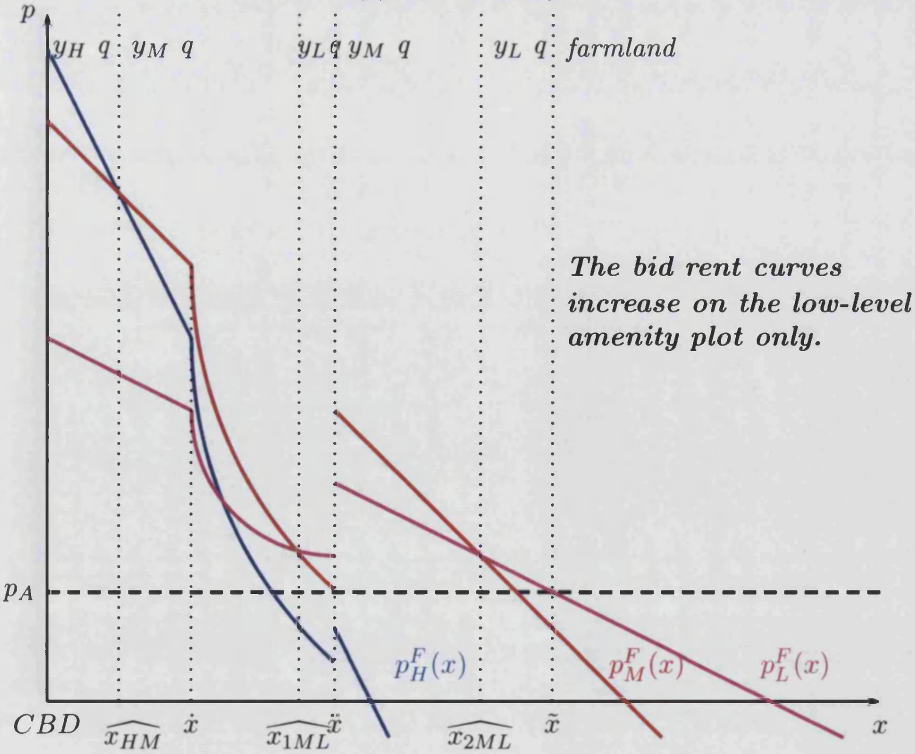
What will happen is that an increase in the housing demand of the poor will shift out their bid rent schedule and increase the rent in the low-income neighbourhood. In our set up, this parallel shift will push out the urban fringe and squeeze the neigh-

Figure 3.8: The effect of an investment in the amenity level of $\dot{x} \leq x \leq \ddot{x}$ on the bid rent curves in the formal housing market

(a) Before the policy



(b) After the policy



bouring middle-income quarter. The increased population density in the latter will then bid up their rents and start squeezing the adjacent central urban area, where the rich live. This process will then be repeated over and over again, until the bid rent in the entire city will have increased, making the total urban population worse off. Figure 3.9a on the next page shows the residential equilibrium before the increase in the housing demand, whereas Figure 3.9b depicts how it changes after the shock *and* a private investment into the amenity level of the squat.

Note that even the new urban residents now have to pay a higher rent. However, in their case this is not due to the squeezing process, but because the profit maximizing landowner wants to recover the fixed cost of the investment, which includes the eviction cost of the squatters. It is assumed that the private investor passes on the entirety of the cost to the new tenants in the terms of a premium to the agent's first bid for the plot (the latter was derived for the old utility level in Figure 3.8b above). Hence, the uniform utility condition again holds for each location within the city boundaries, but at a lower level than before the policy change.

It is important to note at this point that the increase in the housing demand has to be satisfied in any case. Hence, if the landlord had not made the investment into the 'low quality' plot, then the urban population would have been made even worse off. That is, without the investment the new residents would have needed accommodation in the neighbourhoods of their respective income groups and would have thus increased the population density and rent in the entire city. If the average new resident is richer than the average squatter, they would have demanded larger lot sizes, which would have increased the population density even more. Therefore, although the utility level is now lower than before the policy change, it will still be higher than under the alternative scenario of no investment.

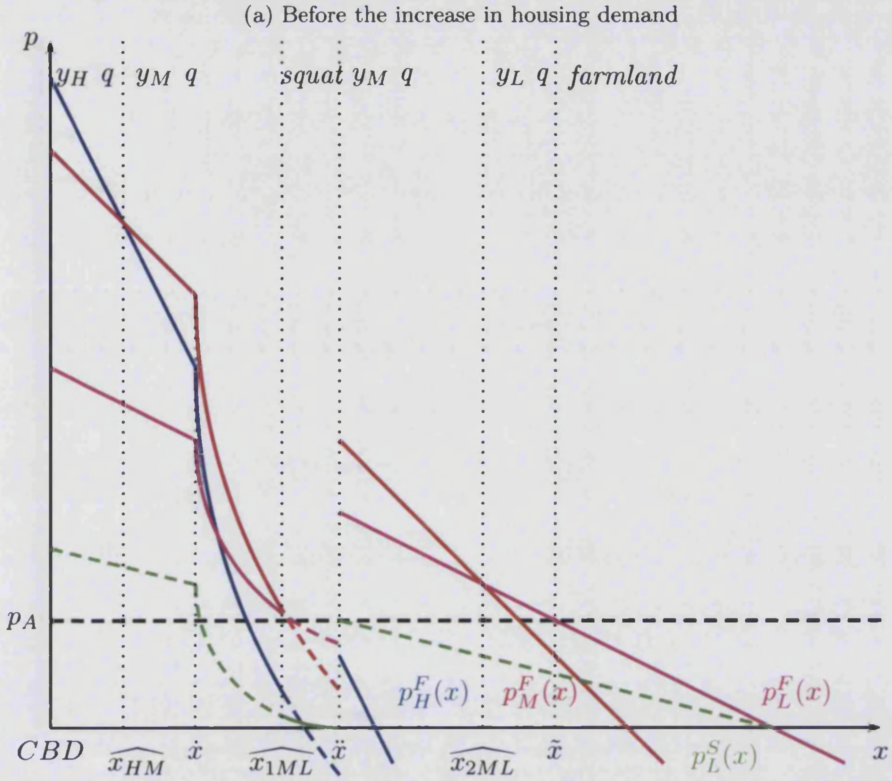
Finally, note that this result is subject to an important caveat. Up to now, it has been assumed that the poor are able to find accommodation in the formal housing market when evicted. However, if the two housing markets are segmented, the evicted squatters will be crowded into illegal subdivisions on the urban fringe. This will make the entire low-income class worse off without any effect on the formal market residents. In this case, a private investment into the squat's amenity level will enhance welfare inequalities between the income groups.

Summarizing yields

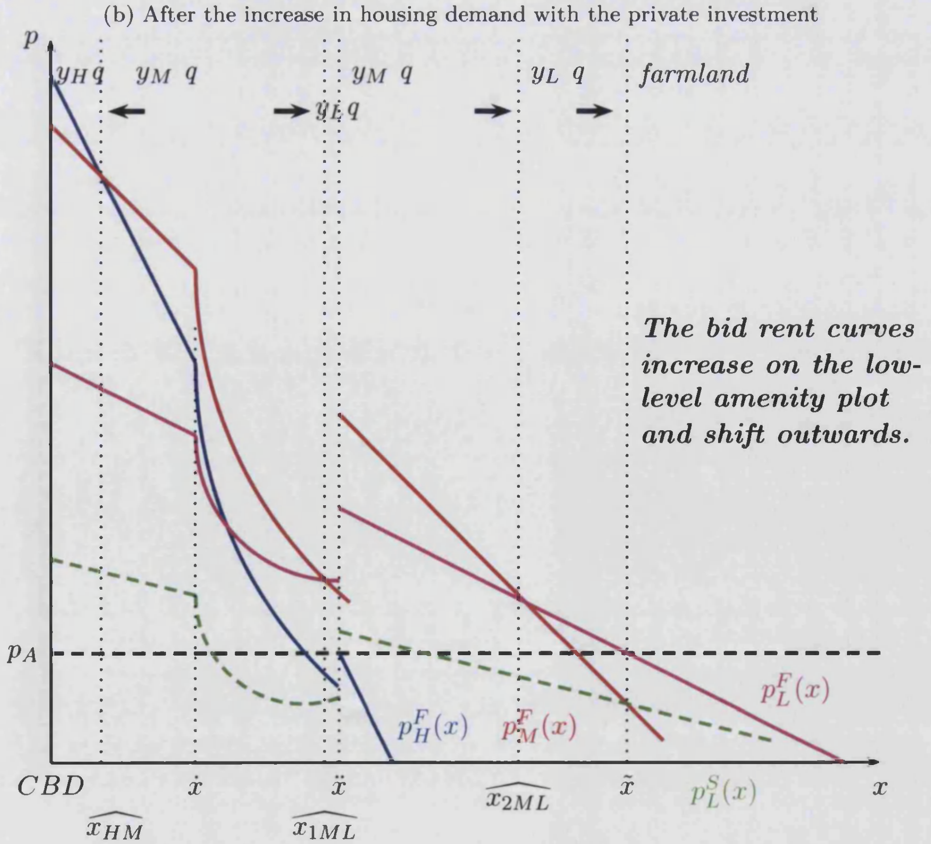
Result 3. *A private investment into the amenity level of the squat is efficient if the average new urban resident is richer than the average squatter. Since the bid rent for the improved lot increases in both the formal and informal housing markets, (most of) the squatters will be evicted and will have to relocate to the low-income neighbourhood on the urban fringe. The subsequent increase in the bid rent will make the total urban population worse off, but less so than if no private investment had been made. If*

Figure 3.9: The effect of a private investment in the amenity level of $\dot{x} \leq x \leq \ddot{x}$ on the residential equilibrium

(a) Before the increase in housing demand



(b) After the increase in housing demand with the private investment



housing markets are segmented, the poor bear the entire cost, which will increase welfare inequalities.

3.4.1.2 Scenario 2: public investment into the amenity level of the squat

This section discusses the second possible scenario under which the amenity level of the squat could be improved, namely a slum upgrading program. Different to the private investment example above, this policy is not undertaken in response to an exogenous increase in the housing demand. Instead these programs are often driven by egalitarian motives and are financed by an international aid agency or the government.

This implies that the improved lot is not rented out to the highest bidder after the program, which would be the middle-income class of the formal housing market in this example (the residential equilibrium before and after the policy is shown on the next page in Figures 3.10a and 3.10b respectively[93]). On the contrary, a commitment is made to the squatters that they can stay on the improved lot.[94] They will choose to hold on to the plot, even if they could sell it at a higher price, if the housing stock in the formal market is scarce (Simha, 2006; WB, 2002) or if there are positive externalities to living in a squat (Neuwirth, 2005; Davis, 2006). Naturally, this efficiency loss will be larger, the greater the improvement in the amenity level.

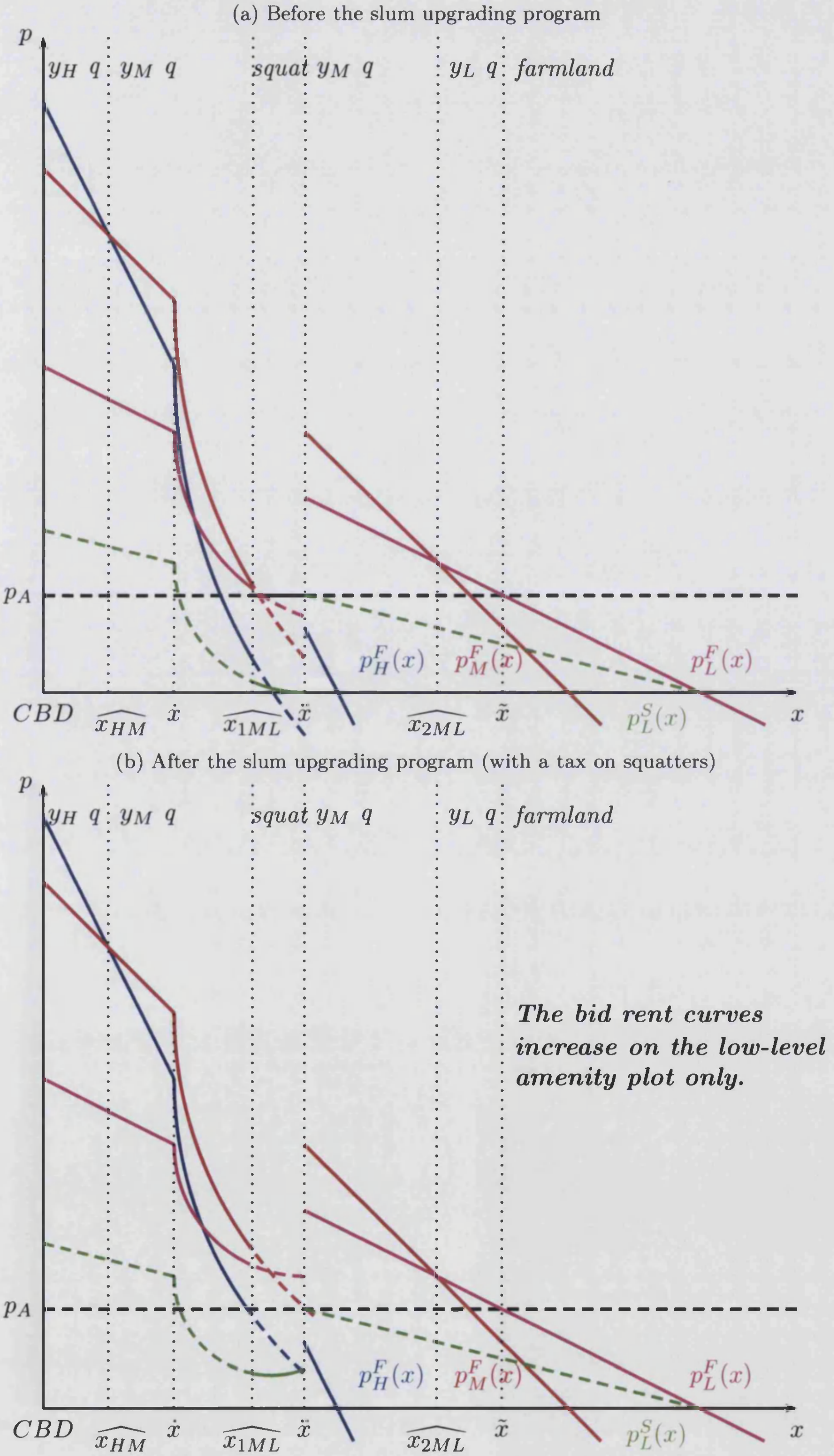
Due to this commitment, no relocations occur within the residential equilibrium as long as the uniform utility condition is satisfied. However, we know that the squatters are made better off by an improvement in the amenity level. We can, thus, only maintain the original residential equilibrium if the squatters are obliged to pay the difference between their old and new valuation of the squat, for instance, in the form of a tax. This contribution could then be used to cover at least part of the fixed investment cost. If this is the case, there is no welfare gain for the urban population, since they are all as well off as before the policy. Furthermore, it is even possible that the slum-upgrading program leads to a social welfare loss, if the contribution of the squatter community is not large enough to cover the fixed investment (Buckley and Kalarickal, 2006).

Alternatively, no tax could be imposed on the squatters. Since they then enjoy a higher utility level than their poor counterparts in the outskirts, the latter have an incentive to move to the informal settlement. The effect of this relocation is opposite to what has been argued in the Section 3.4.1.1. More specifically, the population density in the periphery will fall, which will lower the bid rent curve in the low-income neighbourhood. This effect will be transmitted throughout the city making all income groups better off. However, in this scenario the fixed cost of the investment will not be recovered at all and society will, consequently, incur a welfare loss. To summarize

[93]Note that the bid rent curves for the formal and informal housing markets are only drawn as solid lines when they apply. Otherwise they are represented as dashed lines.

[94]The effect of the commitment itself is studied in more detail in the next section.

Figure 3.10: The effect of a public investment in the amenity level of $\dot{x} \leq x \leq \ddot{x}$ on the residential equilibrium



Result 4. *A public investment into the amenity level of the squat is inefficient, as long as squatters are allowed to remain on the improved lot and richer formal market residents value the lot more. If attempts are made to recover the cost of investment via a tax on squatters, the household's utility remains unchanged. If such attempts are not made, the total urban population will be made better off, since some low-income households will relocate to the improved lot, which will lower the bid rent throughout the city. Yet, society will incur a social welfare loss, since the investment will not be recovered.*

Finally, it is important to note that a slum upgrading program is not merely an improvement in the amenity level. That is, if a commitment is made that the squatters cannot be evicted, this is equivalent to increasing their security of tenure. As a consequence, I am only be able to make a final statement on the size of the efficiency loss created after having studied how the bid rent curves in the informal market respond to a lower risk of eviction. The subsequent section gauges the size of this effect.

3.4.2 The effect of greater security of tenure on the residential equilibrium

This section analyses the effect of increasing the security of tenure on the residential equilibrium. To do this, I simply have to study how the squatter's bid rent curve and its slope respond to a reduction in the probability of eviction, $\pi(x^S)$, for each location x^S. In particular, it is interesting to see how it changes relative to its formal market counterpart, as $\pi(x^S)$ is reduced. We already know from Section 3.2.3 that it lies below the bid rent curve of the formal housing market, $p_j^F(x)$, and is flatter. Now I would like to determine how it approaches this curve when the threat of eviction is reduced.

Since the bid rent offered by squatter household type j has to satisfy the uniform utility condition in equation (16), I use the implicit function theorem to obtain

$$\frac{\partial p_j^S(x^S)}{\partial \pi(x^S)} = \frac{-(u_j^E - u_j^N)}{\pi(x^S)\big(u_{hj}^E \frac{\partial h_j^E}{\partial p_j^S} - u_{cj}^E h_j^N(x^S)\big) + \big(1 - \pi(x^S)\big)\big(u_{hj}^N \frac{\partial h_j^N}{\partial p_j^S} - u_{cj}^N h_j^N(x^S)\big)}. \quad (22)$$

This expression is negative, as the own price effects are negative, $\dfrac{\partial h_j^N}{\partial p_j^S(x^S)} < 0$ and $\dfrac{\partial h_j^E}{\partial p_j^S(x^S)} < 0$, and since households are made worse off in the evicted state, i.e. $u_j^E < u_j^N$ for all $j = L$, M or H. This result is highly intuitive; households attach a lower value to a plot of land on which they face a higher risk of eviction. Consequently, an improvement in the security of tenure increases the households' willingness to pay for a given location.

Furthermore, to find out how the slope of the squatter's bid rent function changes with the threat of eviction, I simply have to differentiate $p_j^{S\prime}(x^S)$ with respect to $\pi(x^S)$. After a few simplifications this yields

$$\begin{aligned} \frac{\partial p_j^{S\prime}(x^S)}{\partial \pi(x^S)} = & \frac{a'}{h_j^N(x^S)} \frac{u_{aj}^E u_{cj}^N - u_{aj}^N u_{cj}^E}{\left(\pi(x^S) u_{cj}^E + \left(1 - \pi(x^S)\right) u_{cj}^N\right)^2} \\ & - p_j^{F\prime}(x^F) \frac{h_j^E(x^F)}{h_j^N(x^S)} \frac{u_{cj}^E u_{cj}^N}{\left(\pi(x^S) u_{cj}^E + \left(1 - \pi(x^S)\right) u_{cj}^N\right)^2} \\ & - \frac{\pi'(x^S)}{h^N(x^S)} \frac{(u_j^E - u_j^N)(u_{cj}^E - u_{cj}^N)}{\left(\pi(x^S) u_{cj}^E + \left(1 - \pi(x^S)\right) u_{cj}^N\right)^2} \end{aligned} \qquad (23)$$

To sign this expression, I use two observations about the evicted and non-evicted state. Firstly, the amenity level in the evicted state is larger than in the non-evicted state. Section 3.2 has shown that the squat is only established on 'low quality' land that is deemed uninhabitable by the formal housing market. Therefore, the amenity level in the evicted state has to be higher, as the squatters find accommodation in the formal housing market. Diminishing marginal returns thus imply that the marginal utility of the amenity level is larger in the non-evicted state, i.e. $u_{aj}^N > u_{aj}^E$ for all $j = L$, M or H.

Secondly, it is important to note that households consume less in the evicted state despite having the same disposable income. This is due to the fact that their investment in the squat's housing stock is lost when they are evicted. Consequently, the marginal utility of consumption in the non-evicted state is lower, such that $u_{cj}^N < u_{cj}^E$ for all $j = L$, M or H. Moreover, this implies that their overall utility level is higher in the non-evicted state, i.e. $u_j^E < u_j^N$ for all $j = L$, M or H; a result that has already been used above.

Since the amenity function, the bid rent curve of the formal housing market, and the probability of eviction all decline with distance from the CBD, I am now able to sign the expressions in equation (23). The overall sign of the partial derivative is ambiguous, because the first two terms are positive and the third negative. Given that there are no closed form solutions for the conventional utility functions, I can again only make conjectures about how the slope of the squatter's bid rent curve will change, as the threat of eviction falls from 1 to 0.

Note that the expression is most likely to be positive on the low quality plot, where $a'(x)$ is large and negative. If this is indeed the case, a reduction in the risk of eviction makes the bid rent curve in the informal housing market steeper and more negative. This is intuitive, as the squatter attaches a greater value to location in the squat, knowing that she is less likely to move to the higher amenity parcel in the formal housing market. Therefore, she has to be compensated for the characteristics of the squat with a lower bid rent, i.e. the slope has steepened. As the risk of eviction falls further, the slope approaches its formal market counterpart from below.

Combining these results with the discussion of slum upgrading programs above yields

Result 5. *The efficiency loss of a slum upgrading program can be reduced if the security of tenure is enhanced for the squatters. Their bid rent curve then approaches its formal market counterpart from below, as the risk premium falls and the slope becomes steeper. However, if a richer income group exists that values the improved lot more, a pareto improvement is still possible, even if the risk of eviction is removed completely.*

3.5 Conclusion

This paper has offered a new theoretical framework to study the presence of squatter settlements in a spatial context; a dimension which has been ignored so far by the existing literature. More specifically, I have introduced both an informal housing market and heterogeneous space into the standard urban Muth-Mills model. This allows me to derive a residential equilibrium, where the formal and informal housing markets compete for land within the city boundaries. In addition, the existence of a squat is endogenously generated on 'low quality' land.

The main policy implication from this analysis is that slum upgrading programs are inefficient. That is, the improved lot does not only become more attractive for the squatters, but also for the formal market residents. Since the latter always outbid the informal housing market, the squatters will be evicted unless they are granted greater security of tenure. If this is indeed the case, they derive utility both from the higher amenity level and the lower risk of eviction. It is then possible to recover (at least part of) the investment through taxation.

Nonetheless, the equilibrium is still inefficient if the rich value the amenity level more than the poor. The squatters are then still outbid by the rich for the improved parcel, even if the threat of eviction is removed completely. However, the poor might be reluctant to sell for a higher price, if there is a housing deficit in the formal market or there are positive externalities to being located in a squat. There is anecdotal evidence that both of these factors are important and more efforts should be directed towards understanding their root causes.

In the context of this model, a relocation policy of the entire squat would be more economically efficient. However, these policies have had mixed success in the past (UN-Habitat, 2003). It is, therefore, crucial to determine the circumstances under which these programs can be successfully implemented. Alternatively, investment could be made in city-wide infrastructure to improve the amenity level for every location. Eventually, the appropriate policy will have to be determined on a case-by-case basis.

However, it should be borne in mind that this model is highly stylised. One possible extension would be to introduce a dynamic framework. This would allow for an explicit modelling of the landowner's outside option and of the shocks that would lead to a private investment in the parcel. Moreover, it would then be possible to study the progression of the neighbourhood after the policy change and how this is influenced by the socioeconomic characteristics of the squatter community, and possible congestion or spillover effects.

One additional step could then be to introduce urban population growth and rural-urban migration into the model. This is particularly important with a view to the climate change debate, as the loss of rural livelihoods through rising temperatures will increase the pressure on urban areas. Given that many developing countries are ill-equipped to expand their formal construction programs, most of these new urban residents will then have to locate in the informal housing market. To avert a major urban housing crisis, research in this field should be given high priority to help shape an effective policy agenda.

C Appendix

C.1 Derivation of the bid rent curve and its slope in the formal housing market; example: Cobb-Douglas utility

This section solves the constrained maximization problem in the formal housing market for the Cobb-Douglas utility function. It then proceed to derive the equilibrium bid rent schedule and its slope.

It is assumed that the utility function takes on the Cobb-Douglas functional form, such that

$$u\big(c, h(x), a(x)\big) = c^{\alpha} h(x)^{\beta} a(x)^{1-\alpha-\beta}, \tag{24}$$

where $\alpha > 0$, $\beta > 0$ and $\alpha + \beta < 1$.

Using the budget constraint of equation (10), I can then set up the Lagrangean L to solve for the optimal consumption and housing bundle

$$\max_{c,h(x)} L = c^{\alpha} h(x)^{\beta} a(x)^{1-\alpha-\beta} + \lambda\big(y - tx - c - p^F(x)h(x)\big) \tag{25}$$

The associated first order conditions are

$$\alpha c^{\alpha-1} h(x)^{\beta} a(x)^{1-\alpha-\beta} = \lambda \tag{26}$$

$$\beta c^{\alpha} h(x)^{\beta-1} a(x)^{1-\alpha-\beta} = \lambda p^F(x), \tag{27}$$

which can be solved for $h(x)$ to yield

$$h(x) = \frac{\beta}{\alpha p^F(x)} c \tag{28}$$

Equation (28) is then substituted into the budget constraint and rearranged to yield the optimal c and $h(x)$:

$$c = \frac{\alpha}{\alpha+\beta}(y - tx) \tag{29}$$

$$h(x) = \frac{\beta}{(\alpha+\beta)p^F(x)}(y - tx) \tag{30}$$

Since households enjoy the same utility level $\bar{u}$ in equilibrium, the bid rent has to decline with distance x to offset the rising commuting costs. To find the bid rent curve in the formal housing market, I therefore only have to express the uniform utility condition at the equilibrium consumption and housing bundle

$$\bar{u} = \frac{\alpha^{\alpha}\beta^{\beta}}{p^F(x)^{\beta}} \left(\frac{y - tx}{\alpha+\beta}\right)^{\alpha+\beta} a(x)^{1-\alpha-\beta} \tag{31}$$

and rearrange it for $p^F(x)$ to yield

$$p^F(x) = \left(\frac{\alpha^\alpha}{\bar{u}}\right)^{\frac{1}{\beta}} \beta \left(\frac{y - tx}{\alpha + \beta}\right)^{\frac{\alpha+\beta}{\beta}} a(x)^{\frac{1-\alpha-\beta}{\beta}}. \tag{32}$$

To obtain the slope of the bid rent curve, I simply differentiate equation (32).

$$p^{F\prime}(x) = \frac{p^F(x)}{\beta}\left(- \frac{(\alpha+\beta)^2 t}{y - tx} + (1 - \alpha - \beta)\frac{a'(x)}{a(x)}\right) \tag{33}$$

The bid rent curve is thus downward sloping, as $a'(x) < 0$.

C.2 Derivation of the slope of the bid rent curve in the informal housing market

This section explains the derivation of the slope of the bid rent curve in the informal housing market. Since the bid rent curve is determined for the optimal choices of the housing and consumption bundle, I first have to solve the constrained expected utility maximization problem of the informal housing market. That is, I maximize the Lagrangean L with respect to c^E, c^N, h^E and h^N, where

$$\begin{aligned} L =& (1 - \pi(x^S))\left[u^N(c^N, h^N(x^S), a(x^S)) + \lambda^N(y - tx^S - p^S(x^S)h^N(x^S) - c^N)\right] \\ &+ \pi(x^S)\left[u^E(c^E, h^E(x^F), a(x^F)) + \lambda^E(y - tx^F + F - p^S(x^S)h^N(x^S) \right. \\ &\left. - p^F(x^F)h^E(x^F) - c^E)\right]. \end{aligned} \tag{34}$$

The associated first order conditions are

$$\begin{aligned} u^E_{c^E} &= \lambda^E & (35)\\ u^E_{h^E} &= \lambda^E p^F(x^F) & (36)\\ u^N_{c^N} &= \lambda^N & (37)\\ u^N_{h^N} &= \lambda^N p^S(x^S) + \frac{\pi(x^S)}{1-\pi(x^S)}\lambda^E p^S(x^S), & (38) \end{aligned}$$

which can be summarized to yield

$$\begin{aligned} u^E_{h^E} &= u^E_{c^E} p^F(x^F) & (39)\\ u^N_{h^N} &= p^S(x^S)\left(u^N_{c^N} + \frac{\pi(x^S)}{1-\pi(x^S)} u^E_{c^E}\right). & (40) \end{aligned}$$

Secondly, the informal housing market only coexists with the formal housing market, if households enjoy the same level of expected utility $\bar{u}$. Therefore, the slope of the bid rent function of the informal housing market can be found by totally differentiating the uniform utility condition described in equation (16) with respect to x^S.

This yields

$$\begin{aligned}0 =&\pi(x^S)\{u^E_{c^E}[-t - p^{S\prime}(x^S)h^N(x^S) - p^S(x^S)h^{N\prime}(x^S) - p^{F\prime}(x^S)h^E(x^S)\\ &- p^F(x^S)h^{E\prime}(x^S)] + u^E_{h^E}h^{E\prime}(x^S) + u^E_a a'(x^S)\}\\ &+ \big(1-\pi(x^S)\big)\{u^N_{c^N}[-t - p^{S\prime}(x^S)h^N(x^S) - p^S(x^S)h^{N\prime}(x^S)]\\ &+ u^N_{h^N}h^{N\prime}(x^S) + u^N_a a'(x^S)\} + \pi'(x^S)(u^E - u^N)\end{aligned} \tag{41}$$

Equation (41) can then be rearranged and simplified with the help of equations (39) and (40) to yield

$$\begin{aligned}p^{S\prime}(x^S) =& -\frac{t}{h^N(x^S)} + \frac{a'(x^S)}{h^N(x^S)}\left(\frac{\pi(x^S)u^E_a + \big(1-\pi(x^S)\big)u^N_a}{\pi(x^S)u^E_c + \big(1-\pi(x^S)\big)u^N_c}\right)\\ &- p^{F\prime}(x^F)\frac{h^E(x^F)}{h^N(x^S)}\frac{\pi(x^S)u^E_c}{\pi(x^S)u^E_c + \big(1-\pi(x^S)\big)u^N_c}\\ &+ \frac{\pi'(x^S)}{h^N(x^S)}\frac{u^E - u^N}{\pi(x^S)u^E_c + \big(1-\pi(x^S)\big)u^N_c}\end{aligned} \tag{42}$$

C.3 The effect of a change in income on the slope of the bid rent curve in the formal housing market; example: Cobb-Douglas utility

To determine how the slope of the bid rent curve changes with income, I simply have to partially differentiate $p^{F\prime}(x)$ with respect to y

$$\begin{aligned}\frac{\partial p^{F\prime}(x)}{\partial y} =& \frac{\partial p^F(x)}{\partial y}\frac{1}{\beta}\left(-\frac{(\alpha+\beta)^2 t}{y-tx} + (1-\alpha-\beta)\frac{a'(x)}{a(x)}\right)\\ &+ \frac{p^F(x)}{\beta}\frac{(\alpha+\beta)^2 t}{(y-tx)^2},\end{aligned} \tag{43}$$

where

$$\frac{\partial p^F(x)}{\partial y} = \left(\frac{\alpha^\alpha}{\bar{u}}\right)^{\frac{1}{\beta}}(\alpha+\beta)\left(\frac{y-tx}{\alpha+\beta}\right)^{\frac{\alpha}{\beta}}a(x)^{\frac{1-\alpha-\beta}{\beta}} \tag{44}$$

The partial derivative can be simplified further to yield

$$\frac{\partial p^{F\prime}(x)}{\partial y} = \left(\frac{\alpha^\alpha}{\bar{u}}\right)^{\frac{1}{\beta}}\frac{\alpha+\beta}{\beta}a(x)^{1-\alpha-\beta}\left(\frac{y-tx}{\alpha+\beta}\right)^{\frac{\alpha}{\beta}}\left(\frac{-t\big((\alpha+\beta)^2-\beta\big)}{y-tx} + (1-\alpha-\beta)\frac{a'(x)}{a(x)}\right) \tag{45}$$

Equation (45) is unambiguously negative, if $(\alpha+\beta)^2 - \beta > 0$.

C.4 The effect of a slum-upgrading program on the bid rent curve in the formal housing market; example: Cobb-Douglas utility

To determine the effect of a slum-upgrading program on the bid rent curve in the formal housing market, I partially differentiate $p^F(x)$ with respect to $a(x)$

$$\begin{aligned}\frac{\partial p^F(x)}{\partial a(x)} &= \left(\frac{\alpha^\alpha}{\bar{u}}\right)^{\frac{1}{\beta}} \beta \left(\frac{y-tx}{\alpha+\beta}\right)^{\frac{\alpha+\beta}{\beta}} \frac{1-\alpha-\beta}{\beta} a(x)^{\frac{1-\alpha-2\beta}{\beta}} \\ &= \frac{p^F(x)}{a(x)} \frac{1-\alpha-\beta}{\beta} > 0,\end{aligned} \tag{46}$$

where the second equality uses equation (32). Consequently, the bid rent offered in the formal housing market increases after a slum-upgrading program. Note that this effect is increasing in income y, because

$$\frac{\partial(\frac{\partial p^F(x)}{\partial a(x)})}{\partial y} = \frac{\partial p^F(x)}{\partial y} \frac{1-\alpha-\beta}{\beta a(x)} > 0, \tag{47}$$

where

$$\frac{\partial p^F(x)}{\partial y} = \left(\frac{\alpha^\alpha}{\bar{u}}\right)^{\frac{1}{\beta}} (\alpha+\beta) \left(\frac{y-tx}{\alpha}\right)^{\frac{\alpha}{\beta}} a(x)^{\frac{1-\alpha-\beta}{\beta}} > 0. \tag{48}$$

To determine the effect of a slum-upgrading program on the slope of the formal housing market, I partially differentiate $p^{F'}(x)$ with respect to $a(x)$

$$\frac{\partial p^{F'}(x)}{\partial a(x)} = \frac{\partial p^F(x)}{\partial a(x)} \frac{1}{\beta} \left(-\frac{(\alpha+\beta)^2 t}{y-tx} + (1-\alpha-\beta)\frac{a'(x)}{a(x)} \right) - \frac{p^F(x)}{\beta}(1-\alpha-\beta)\frac{a'(x)}{a(x)^2} \tag{49}$$

This expression can be further simplified with equation (46) to yield

$$\frac{\partial p^{F'}(x)}{\partial a(x)} = \frac{p^F(x)}{a(x)} \frac{1-\alpha-\beta}{\beta^2} \left(-\frac{(\alpha+\beta)^2 t}{y-tx} + (1-\alpha-2\beta)\frac{a'(x)}{a(x)} \right) \tag{50}$$

The overall sign of this expression is ambiguous, as the first term in the brackets is negative and the second positive given that $1-\alpha-2\beta < 0$ and $a'(x) < 0$. However, the first term in brackets is very small, since $0 < \alpha < 1$, $0 < \beta < 1$ and $y - tx$ is large. Furthermore, the second term is larger, the smaller the amenity level $a(x)$ and the steeper its slope $a'(x)$. Therefore, it is likely that the slope of the bid rent of the formal housing market increases after a slum-upgrading program on the 'low quality' plot.

This effect is further reinforced as $p^{F'}(x)$ increases with $a'(x)$, because

$$\frac{\partial p^{F'}(x)}{\partial a'(x)} = p^F(x)\frac{(1-\alpha-\beta)}{\beta a(x)} > 0 \tag{51}$$

Note that both of these effects are also increasing in y, since

$$\begin{aligned}\frac{\partial(\frac{\partial p^{F'}(x)}{\partial a(x)})}{\partial y} &= \frac{\partial p^F(x)}{\partial y}\frac{(1-\alpha-\beta)}{\beta^2 a(x)}\left(-\frac{(\alpha+\beta)^2 t}{y-tx}+(1-\alpha-2\beta)\frac{a'(x)}{a(x)}\right)\\ &+p^F(x)\frac{(1-\alpha-\beta)}{\beta^2 a(x)}\frac{(\alpha+\beta)^2 t}{(y-tx)^2}>0\end{aligned} \tag{52}$$

if $\left(-\frac{(\alpha+\beta)^2 t}{y-tx}+(1-\alpha-2\beta)\frac{a'(x)}{a(x)}\right)>0$ and $\frac{\partial p^F(x)}{\partial y}>0$. Moreover,

$$\frac{\partial(\frac{\partial p^{F'}(x)}{\partial a'(x)})}{\partial y}=\frac{\partial p^F(x)}{\partial y}\frac{(1-\alpha-\beta)}{\beta a(x)}>0, \tag{53}$$

if $\frac{\partial p^F(x)}{\partial y}>0$.

C.5 The effect of a slum-upgrading program on the bid rent curve in the informal housing market

To determine the effect of an increase in the amenity level of the squat $a(x^S)$ on the bid rent curve in the informal housing market $p^S(x^S)$, I use the implicit function, which yields

$$\frac{\partial p^S(x^S)}{\partial a(x^S)}=\frac{-\left(\pi(x^S)u_a^E+\left(1-\pi(x^S)\right)u_a^N\right)}{\pi(x^S)\left(u_h^E\frac{\partial h^E}{\partial p^S}-u_c^E h^E(x^F)\right)+\left(1-\pi(x^S)\right)\left(u_h^N\frac{\partial h^N}{\partial p^S}-u_c^N h^N(x^S)\right)} \tag{54}$$

This expression is positive, as $\frac{\partial h^E(x^F)}{\partial p^S(x^S)}<0$ and $\frac{\partial h^N(x^S)}{\partial p^S(x^S)}<0$.

To determine the effect of a slum-upgrading program on the slope of the informal housing market, I partially differentiate $p^{F'}(x)$ with respect to $a'(x^S)$

$$\frac{\partial p^{S'}(x^S)}{\partial a'(x^S)}=\frac{1}{h^N(x^S)}\left(\frac{\pi(x^S)u_a^E+\left(1-\pi(x^S)\right)u_a^N}{\pi(x^S)u_c^E+\left(1-\pi(x^S)\right)u_c^N}\right)>0 \tag{55}$$

The partial derivative of the slope of the $p^{F'}(x)$ with respect to $a(x^S)$ has been omitted, since the expression is unwieldy and difficult to sign without a closed-form solution.

References

Akhmedov, Akhmed and Ekaterina Zhuravskaya, "Opportunistic Political Cycles: Test in A Young Democracy Setting," *Quarterly Journal of Economics*, 2004, *119* (4), 1301–1338.

Albala-Bertrand, Jose-Miguel, *Political Economy of large Natural Disasters: with Special Reference to Developing Countries*, Oxford, United Kingdom: Clarendon Press, 1993.

Alesina, Alberto, "Macroeconomic Policy in a Two-Party System as a Repeated Game," *The Quarterly Journal of Economics*, 1987, *102* (3), 651–678.

Alonoso, William, *Location and Land Use: Toward a General Theory of Land Rent*, Cambridge (MA), United States of America: Harvard University Press, 1964.

Anbarci, Nejat, Monica Escalares, and Charles A. Register, "Earthquake Fatalities: The Interaction of Nature and Political Economy," *Journal of Public Economics*, 2005, *89*, 1907–1933.

Asner, Gregory P., "Cloud Cover in Landsat Observations of the Brazilian Amazon," *International Journal of Remote Sensing*, 2001, *22*, 3855–3862.

Au, Chun-Chung and J. Vernon Henderson, "Are Chinese Cities Too Small?," *Review of Economic Studies*, 2006, *73*, 549–576.

– and Vernon J. Henderson, "How Migration Restrictions Limit Agglomeration and Productivity in China," *Journal of Economic Development*, 2006, *80*, 350–388.

Badcock, Blair, *Unfairly Structured Cities*, Oxford, United Kingdom: Blackwell, 1984.

BAPS, "Disaster Relief 1990-1999 - 1999 Supercyclone of Orissa," 1999.

Bar-Ilan, Avner and William C. Strange, "Urban Development with Lags," *Journal of Urban Economics*, 1996, *39* (1), 87–113.

Barber, Charles V., "The Legal and Regulatory Framework and Forest Production in Indonesia," in Charles Zerner, ed., *Community Land Rights, Customary Law, and the Law of Timber Concessions in Indonesia's Forests: Legal Options and Alternatives in designing the Commons*, number UFT/INS/065. In 'Forestry Studies report.', Jakarta, Indonesia: Ministry of Forestry/FAO, 1990, chapter Appendix. Draft (March 1990).

– and Gregory Churchill, "Land Policy in Irian Jaya: Issues and Strategies," Project Report INS/83/013, Government of Indonesia; United Nations Development Program; International Bank of Reconstruction and Development, Jakarta, Indonesia 1987.

Barr, Christopher, Eva Wollenberg, Godwin Limberg, Njau Anau, Ramses Iwan, I Made Sudana, Moira Moeliono, and Tony Djogo, "The Impacts of Decentralisation on Forests and Forest-Dependent Communities in

Malinau District, East Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 3, Center for International Forestry Research, Bogor, Indonesia 2001.

– , **Ida Aju Pradnja Resosudarmo, Ahmad Dermawan, and Bambang Setiono**, "Decentralization's Effects on Forest Concessions and Timber Production," in Christopher Barr, Ida Aju Pradnja Resosudarmo, Ahmad Dermawan, and John McCarthy, eds., *Decentralization of Forest Administration in Indonesia: Implications for Forest Sustainability, Economic Development and Community Livelihoods*, Bogor, Indonesia: Center for International Forestry Research, 2006, chapter 5, pp. 87–107.

Bartone, Carl, Janis Bernstein, Josef Leitmann, and Jochen Eigen, "Toward Environmental Strategies for Cities: Policy Considerations for Urban Environmental Management in Developing Countries," Urban Management Programme Policy Paper 18, World Bank, Washington DC, United States of America 1994.

Bertrand, Jacques, "Indonesia's Quasi-federalist Approach: Accommodation amid Strong Integrationist Tendencies," *International Journal of Constitutional Law*, 2008, *5*, 576–605.

Besley, Timothy, "Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries," *Journal of Economic Perspectives*, 1995, *9* (3), 115–127.

– , "Property Rights and Investment Incentives: Theory and Evidence from Ghana," *Journal of Political Economy*, 1995, *103* (5), 903–937.

– **and Anne Case**, "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition," *The American Economic Review*, 1995, *85* (1), 25–45.

– **and Robin Burgess**, "The Political Economy of Government Responsiveness: Theory and Evidence from India," *Quarterly Journal of Economics*, November 2002, *117* (4), 1415–1451.

Bluedorn, John C. and Elizabeth U. Cascio, "Education and Intergenerational Mobility: Evidence from a Natural Experiment in Puerto Rico," Economics Papers 21, Economics Group, Nuffield College, University of Oxford April 2005.

Boccard, Nicolas and Yves Zenou, "Racial Discrimination and Redlining in Cities," *Journal of Urban Economics*, September 2000, *48* (2), 260–285.

BPS, "Daftar Nama Provinsi/Kabupaten/Kota Menurut Dasar Hukum Pembentukan Wilayah (Municipality and Regency Codes)," http://dds.bps.go.id/eng/mstkab/mfd2007.pdf? 2007.

Braid, Ralph M., "Residential Spatial Growth with Perfect Foresight and Multiple Income Groups," *Journal of Urban Economics*, 1991, *30* (3), 385–407.

Breiman, Leo, "Bagging Predictors," *Machine Learning*, 1996, *26*, 123–140.

– , **Jerome Friedman, Richard A. Olshen, and Charles J. Stone**, *Classification and Regression Trees*, Monterey (CA), United States of America: Wadsworth, 1984.

Brodjonegoro, Bambang and Jorge Martinez-Vazquez, "An Analysis of Indonesia's Transfer System: Recent Performance and Future Prospects," Working Paper 13, International Studies Program May 2002.

Brown, Gardner M., "Renewable Natural Resource Management and Use without Markets," *Journal of Economic Literature*, 2000, pp. 875–914.

Brueckner, Jan K., "The Structure of Urban Equilibria: a unified Treatment of the Muth-Mills Model," in Edwin S. Mills, ed., *Handbook of Regional and Urban Economics*, Vol. 2, Amsterdam,The Netherlands: Elsevier Science Publications, 1987.

– **and Burkhard von Rabenau**, "Dynamics of Land-Use for a Closed City," *Regional Science and Urban Economics*, 1981, *11* (1), 1–17.

– **and Harris Selod**, "A Theory of Urban Squatting and Land-Tenure Formalization in Developing Countries," CESifo Working Paper 2328, CESifo Group Munich June 2008.

– **, Jacques F. Thisse, and Yves Zenou**, "Why is Central Paris rich and Downtown Detroit poor? An Amenity-based Theory," *European Economic Review*, 1999, *43*, 91–107.

Buckley, Robert M. and Jerry Kalarickal, *Thirty Years of World Shelter Lending. What Have we Learned?*, Washington DC, United States of America: World Bank, 2006.

Carter, Michael and Keith Wiebe, *Land Tenure Reform in Sub-Saharan Africa*, Kendall: Hunt Press, 1994.

– **and Pedro Olinto**, "Getting Institutions "Right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment," *American Journal of Agricultural Economics*, February 2003, *85* (1), 173–186.

Casson, Anne, "Decentralisation of Policies Affecting Forests and Estate Crops in Kotawaringin Timur District, Central Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 5, Center for International Forestry Research, Bogor, Indonesia 2001.

– , "Decentralisation of Policies Affecting Forests and Estate Crops in Kutai Barat District, East Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 4, Center for International Forestry Research, Bogor, Indonesia 2001.

– **, Agus Setyarso, Mario Boccucci, and David W. Brown**, "A Background to Illegal Logging and Law Enforcement in Indonesia," Technical Report, World Bank/WWF Alliance 2006.

– **and Krystof Obidzinski**, "From New Order to Regional Autonomy: Shifting Dynamics of "Illegal" Logging in Kalimantan, Indonesia," *World Development*, 2002, *30* (12), 2133–2151.

Christensen, Jens Hesselbjerg, Bruce Hewitson, Aristita Busuioc, Anthony Chen, Xuejie Gao, Isaac Held, Richard Jones, Rupa Kumar

Kolli, Won-Tae Kwon, René Laprise, Victor Maga na Rueda, Linda Mearns, Claudio Guillermo Menéndez, Jouni Räisänen, Annette Rinke, Abdoulaye Sarr, and Penny Whetton, "Regional Climate Projections," in Susan Solomon, Dahe Qin, Martin Manning, Zhenlin Chen, Melinda Marquis, Kristen Averyt, Melinda M.B. Tignor, and Jr. Miller, Henry LeRoy, eds., *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom and New York, United States of America: Cambridge University Press, 2007, chapter 11, pp. 847–940.

CIFOR, "Generating Economic Growth, Rural Livelihoods, and Environmental Benefits from Indonesia's Forests: A Summary of Issues and Policy Options," 2004. Center for International Forestry Research, Report prepared for the World Bank.

Cohen, Margot, "Tackling a bitter Legacy," *Far Eastern Economic Review*, 2 July 1998.

– , "New Indonesian Districts Radically Reshape Country," *Far Eastern Economic Review*, 2003.

Cohen, Michael, "Learning by Doing: World Bank Lending for Urban Development, 1972-1982," Technical Report, World Bank 1983.

Cole, Shawn, Andrew Healy, and Eric Werker, "Do Voters Appreciate Responsive Governments? Evidence from Indian Disaster Relief," Finance Working Paper 09-050, Harvard Business School July 2008.

– , Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert M. Townsend, and James Vickery, "Barriers to Household Risk Management: Evidence from India," Finance Working Paper 09-116, Harvard Business School October 2009.

Cuaresma, Jesús Crespo, Jaroslava Hlouskova, and Michael Obersteiner, "Natural Disasters as Creative Destruction? Evidence from Developing Countries," *Economic Inquiry*, April 2008, *46* (2), 214–226.

Curran, Lisa M., Simon N. Trigg, Alice K. McDonald, Dwi Astiani, Y.M. Hardiono, P. Siregar, I. Caniago, and Eric S. Kasischke, "Lowland forest loss in protected areas of Indonesian Borneo," *Science*, February 2004, *303* (5660), 1000–1003.

Curran, Paul J., "Multispectral Remote Sensing of Vegetation Amount," *Progress in Physical Geography*, 1980, *4*, 315–341.

Dasgupta, Partha, *The Control of Resources*, Cambridge (MA), United States of America: Harvard University Press, 1982.

– and Geoffrey Heal, "The Optimal Depletion of Exhaustible Resources," *The Review of Economic Studies*, 1974, *41*, 3–28.

Davis, Donald R. and David E. Weinstein, "Bones, Bombs, and Break Points: The Geography of Economic Activity," *The American Economic Review*, December 2002, *92* (5), 1269–1289.

Davis, Mike, *Planet of Slums*, London, United Kingdom: Verso, 2006.

Deaton, Angus, "Household Saving in LDCs: Credit Markets, Insurance and Welfare," *Scandinavian Journal of Economics*, 1992, *94* (2), 253–273.

Denman, Kenneth L., Guy Brasseur, Amnat Chidthaisong, Philippe Ciais, Peter M. Cox, Robert E. Dickinson, Didier Hauglustaine, Christoph Heinze, Elisabeth Holland, Daniel Jacob, Ulrike Lohmann, Ramachandran, Pedro Leite da Silva Dias, Steven C. Wofsy, and Xiaoye Zhang, "Couplings Between Changes in the Climate System and Biogeochemistry," in Susan Solomon, Dahe Qin, Martin Manning, Zhenlin Chen, Melinda Marquis, Kristen Averyt, Melinda M.B. Tignor, and Henry LeRoy Miller, eds., *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom and New York, United States of America: Cambridge University Press, 2007, chapter 7.

Dercon, Stefan, "Income Risk, Coping Strategies and Safety Nets," Working Paper Series 136, Centre for the Study of African Economies 2000.

Dilley, Maxx, Robert S. Chen, Uwe Deichmann, Arthur L. Lerner-Lam, and Margaret Arnold, *Natural Disaster Hotspots: A Global Risk Analysis* number 5. In 'Disaster Risk Management Series.', Washington DC, United States of America: World Bank, 2005.

Dinar, Ariel, Robert Mendelsohn, Robert Evenson, Jyoti Parikh, Apurva Sanghi, Kavi Kumar, James Jr. McKinsey, and Stephen Lonergan, "Measuring the Impact of Climate Change on Indian Agriculture," Technical Paper 402, World Bank, Washington DC, United States of America 1998.

DiTella, Rafael, Sebastian Galiani, and Ernesto Schargrodsky, "The Formation of Beliefs: Evidence from the Allocation of Land Titles to Squatters," *Quarterly Journal of Economics*, 2007, *122*, 209–241.

Dubois, Pierre, "Assurance Compète, Hétérogénéité des Préférences et Métayage au Pakistan," *Annales d'Économie et de Statistique*, 2000, (59), 1–35.

Duflo, Esther and Rohini Pande, "Dams," *Quarterly Journal of Economics*, May 2007, *122* (2), 601–646.

Eisensee, Thomas and David Strömberg, "New Floods, New Droughts, and U.S. Disaster Relief," *Quarterly Journal of Economics*, 2007, *122* (2), 693–728.

Emanuel, Kerry, "Increasing Destructiveness of Tropical Cyclones over the past 30 Years," *Nature*, August 2005, *436*, 686–688. Letters.

Emmerson, Donald K., *Indonesia beyond Suharto: Polity, Economy, Society, Transition*, Armonk (NY), United States of America: M.E. Sharpe, 1999. published in cooperation with the Asia Society.

Erb, Maribeth and Priyambudi Sulistiyanto, "Indonesia and the Quest for

Democracy," in Maribeth Erb and Priyambudi Sulistiyanto, eds., *Deepening Democracy in Indonesia? Direct Elections for Local Leaders (Pilkada)*, Singapore: Institute of South East Asian Studies, 2009.

Fafchamps, Marcel and Susan Lund, "Risk Sharing Networks in Rural Philippines," Technical Report, Mimeograph August 1997.

FAO, "FAO Statistical Yearbook 2009," http://www.fao.org/economic/ess/publications-studies/statistical-yearbook/fao-statistical-yearbook-2009/en/ 2009. Accessed on 10th July 2010.

Field, Erica, "Entitled to Work: Urban Tenure Security and Labor Supply in Peru," *Quarterly Journal of Economics*, 2007, *122* (4), 1561–1602.

Finkelstein, Amy, "E-ZTAX: Tax Salience and Tax Rates," *Quarterly Journal of Economics*, 2009, *124* (3), 969–1010.

Firman, Tommy, "Land Conversion and Urban Development in the Northern Region of West Java, Indonesia," *Urban Studies*, 1997, *34* (7), 1027–1046.

Fisman, Raymond, "Estimating the Value of Political Connections," *The American Economic Review*, September 2001, *91* (4), 1095–1102.

Fitrani, Fitria, Bert Hofman, and Kai Kaiser, "Unity in Diversity? The Creation of New Local Governments in a Decentralising Indonesia," *Bulletin of Indonesian Economic Studies*, 2005, *41* (1), 57–79.

Flood, J., "Secure Tenure Survey Final Report," Technical Report 2006.

Foster, Andrew D., "Prices, Credit Markets and Child Growth in Low-Income Rural Areas," *Economic Journal*, 1995, *105* (430), 551–570.

Friedman, Joseph, Emmanuel Jiminez, and Stephen K. Mayo, "The Demand for Tenure Security in Developing Countries," *Journal of Development Economics*, 1988, *29*, 185–198.

Fujita, Masahisa, "Urban Spatial Dynamics: A Review," *Sistemi Urbani*, 1983, *3*, 411–475.

— , *Urban Economic Theory, Land Use and City Size*, Cambridge, United Kingdom: Cambridge University Press, 1989.

FWI/GFW, *The State of the Forest - Indonesia*, Bogor, Indonesia; Washington DC: World Resources Institute, Forest Watch Indonesia, Global Forest Watch, 2002.

Galal, Ahmed and Omar Razzaz, "Reforming Land and Real Estate Markets," Policy Research Working Paper 2616, World Bank 2001.

Galiani, Sebastian, "Effects of Land Titling on Child Health," *Economics and Human Biology*, 2004, *2*, 353–372.

Gausman, Harold W., "Reflectance of Leaf Components," *Remote Sensing of Environment*, 1977, *6*, 1–9.

Gholz, Henry L., "Environmental Limits on above Ground Net Primary Production, Leaf Area, and Biomass in Vegetation Zones of the Pacific

Northwest," *Ecology*, 1982, *54*, 152–159.

GOI, *Indian Crop Calendar*, 4 ed., New Delhi, India: Manager of Publications, Civil Lines, 1967.

Greenpeace, "Stop Pembabatan Hutan Untuk Dana Kampanye," January 2009.

Grimard, Franque, "Household Consumption Smoothing through Ethnic Ties: Evidence from Cote d'Ivoire," *Journal of Development Economics*, 1997, *53*, 391–422.

Guiteras, Raymond, "The Impact of Climate Change on Indian Agriculture." Mimeo, Massachusetts Institute of Technology 2007.

Hansen, Matthew C. and Ruth S. DeFries, "Detecting Long-term Global Forest Change Using Continuous Fields of Tree-Cover Maps from 8-km Advanced Very High Resolution Radiometer (AVHRR) Data for the Years 1982-99," *Ecosystems*, 2004, *7*, 695–716.

– , Stephen V. Stehman, Peter V. Potapov, Belinda Arunarwati, Fred Stolle, and Kyle Pittman, "Quantifying Changes in the Rates of Forest Clearing in Indonesia from 1990 to 2005 using Remotely Sensed Data Sets," 2009. submitted to IOP Publishing for peer review 25 May 2009.

– , – , – , Thomas R. Loveland, John R. G. Townshend, Ruth S. DeFries, Kyle W. Pittman, Belinda Arunarwati, Fred Stolle, Marc K. Steininger, Mark Carroll, and Charlene DiMiceli, "Humid Tropical Forest Clearing from 2000 to 2005 Quantified using Multitemporal and Multi-resolution Remotely Sensed Data," *Proceedings of the National Academy of Sciences*, 2008, *105*, 9439–9444.

Hausman, Jerry A., Bronwyn H. Hall, and Zvi Griliches, "Econometric Models for Count Data with an Application to the Patents-R and D Relationship," *Econometrica*, 1984, *52* (4), 909–938.

Henderson, J. Vernon and Anthony J. Venables, "The Dynamics of City Formation," *Review of Economic Dynamics*, April 2009, *12* (2), 233–254.

– and Hyoung Gun Wang, "Aspects of the Rural-Urban Transformation of Countries," *Journal of Economic Geography*, 2005, *5* (1), 23–42.

– and Jacques F. Thisse, eds, *Handbook of Urban and Regional Economics: Cities and Geography*, Vol. 4, Amsterdam, The Netherlands; Oxford, United Kingdom: North Holland, 2004.

Henderson-Sellers, Ann, Huqiang Zhang, Gerhard Berz, Kerry Emanuel, William M. Gray, Chris Landsea, Greg J. Holland, James. Lighthill, Shinn-Liang Shieh, Peter J. Webster, and Kendal McGuffie, "Tropical Cyclones and Global Climate Change: a post-IPCC Assessment," *Bulletin of the American Meteorological Society*, 1998, *79*, 19–38.

Hofman, Bert and Kai Kaiser, "The Making of the Big Bang and its Aftermath: A Political Economy Perspective," in James Alm, Martinez Vazquez Jorge, and Sri Mulyani Indrawati, eds., *Reforming Intergovernmental Fiscal*

Relations and the Rebuilding of Indonesia, Cheltenham, United Kingdom: Edward Elgar, 2004.

Holland, Greg, "An Analytic Model of the Wind and Pressure Profiles in Hurricanes," *Monthly Weather Review*, 1980, *108*, 1212–1218.

— , "Windfield Dynamics of Landfalling Tropical Cyclones," Technical Report, Bureau of Meteorological Research Center, Melbourne 28-30 May 1997. Paper presented at the Workshop on Windfield Dynamics of Landfalling Tropical Cyclones.

Holmes, Derek A., "Where have all the Forests Gone?," Environment and Social Development East Asia and Pacific Region Discussion Papers 24462, World Bank June 2002.

Hooijer, Aljosja, Marcel Silvius, Henk Wösten, and Susan Page, "PEAT-CO2 Assessment of CO2 emissions from drained peatlands in SE Asia," Delft Hydraulics Report 3943, Wetland International, Delft, The Netherlands 2006.

Hornbeck, Richard A., "Quantifying Long-term Adjustment to Environmental Change: Evidence from the American Dust Bowl." Mimeo, Massachusetts Institute of Technology 2008.

Hoy, Michael and Emmanuel Jiminez, "Squatter Rights and Urban Development: an Economic Perspective," *Economica*, February 1991, *58* (229), 79–92.

HPC, *Report of the High Power Committee on Disaster Management*, New Delhi, India: Government of India, 2001.

Hull, Terence, "Striking a Most Delicate Balance: The Implications of Otonomi Daerah for the Planning and Implementation of Development Cooperation Projects," Final Report, Australian National University, Research School of Social Sciences 1999.

IMD, *Tracks of Storms and Depressions in the Bay of Bengal and the Arabian Sea 1877-1970*, 2nd ed., New Delhi, India: India Meteorological Department, 1979.

— , "Cyclone eAtlas: Tracks of Cyclones and Depressions in the Bay of Bengal and the Arabian Sea 1891-2007," June 2008. electronic version 1.0/2008.

— , "Hazards of Cyclones, Heavy Rainfall, Strong Winds and Strom Surges," http://www.imdmumbai.gov.in/cycdisasters.htm 2009. Accessed on 12th March 2009.

IPCC, *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom: Cambridge University Press, 2007.

— , *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Geneva: IPCC, 2007.

ISDR, "Global Assessment Report on Disaster Risk Reduction," Technical Report, International Strategy for Disaster Reduction Secretariat, United Nations, Geneva, Switzerland 2009.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), 538–575.

Jiminez, Emmanuel, "Tenure Security and Urban Squatting," *Review of Economics and Statistics*, 1984, *66*, 556–67.

— , "Urban Squatting and Community Organization in Developing Countries," *Journal of Public Economics*, June 1985, *27* (1), 69–92.

Ju, Junchang and David P. Roy, "The Availability of Cloud-free Landsat ETM+ Data over the Conterminous United States and globally," *Remote Sensing of Environment*, 2008, *112*, 1196–1211.

KaltimPost, "Anggota Dewan Juga Dapat Insentif Migas (Local Parliament Members Also Received Oil and Gas Incentive)," *Kaltim Post*, July 23rd 2009.

— , "Muspida Kukar Tak Kembalikan (Conference of District Leaders Funds for Kutai Kartanegara Not Returned)," *Kaltim Post*, July 22nd 2009.

Katz, Lawrence and Kenneth T. Rosen, "Growth Management and Land Use Controls: the San Francisco Bay Area Experience," *American Real Estate and Urban Economics Association Journal*, 1981, *9* (4), 321–344.

— **and** — , "The Interjurisdictional Effects of Growth Controls on Housing Prices," *Journal of Law and Economics*, 1987, *30* (1), 149–160.

Khan, Matthew E., "The Death Toll from Natural Disasters: the Role of Income, Geography, and Institutions," *Review of Economics and Statistics*, May 2005, *87* (2), 271–284.

Kindermann, Georg, Michael Obersteiner, Brent Sohngen, Jayant Sathaye, Kenneth Andrasko, Ewald Rametsteiner, Bernhard Schlamadinger, Sven Wunder, and Robert Beach, "Global Cost Estimates of Reducing Carbon Emissions Through Avoided Deforestation," *Proceedings of the National Academy of Sciences*, 2008, *105* (30), 10302–12307.

Kompas, "KPK Tahan Hamid Rizal (Anti-Corruption Comission Arrests Hamid Rizal)," *Kompas*, October 12th 2009.

KoranTempo, "Polisi Sidik Dugaan Korupsi Rp. 6.5 Milyar (Police Alledge Corruption of Rp. 6.5 billion)," *Koran Tempo*, January 3rd 2006.

Kunreuther, Howard and Mark Pauly, "Neglecting Disaster: Why don't People insure against large Losses?," *Journal of Risk and Uncertainty*, 2004, *28* (1), 5–21.

Kuntson, Thomas R. and Robert A. Tuleya, "Impact of CO2-induced Warming on simulated Hurricane Intensity and Precipitation: Sensitivity to the Choice of Climate Model and convective Parameterization," *Journal of Climate*, 2004, *17*, 3477–3495.

Lanjouw, Jean O. and Philip I. Levy, "Untitled: A Study of Formal and Informal Property Rights in Urban Ecuador," *Economic Journal*, October 2002, *112* (482), 986–1019.

Larson, Bruce A. and Daniel W. Bromley, "Property rights, externalities, and resource degradation: Locating the tragedy," *Journal of Development Economics*, 1990, *33* (2), 235–262.

Levitt, Steven D., "Using Electoral Cycles in Police Hiring to estimate the Effect of Police on Crime," *The American Economic Review*, 1997, *87* (3), 270–290.

Lewis, Tracy R. and Richard Schmalensee, "On Oligopolistic Markets for Nonrenewable Natural Resources," *Quarterly Journal of Economics*, 1980, *95* (3), 475–491.

Ligon, Ethan and Urvashi Narain, "Government Management of Village Commons: Comparing Two Forest Policies," *Journal of Environmental Economics and Management*, 1999, *37* (3), 272–289.

Lovell, K., "Review of Empirical Tropical Cyclone Wind and Pressure Field Models," Master's thesis, Department of Civil Engineering, University of Western Australia 1990.

MacRae, C. Duncan, "A Political Model of the Business Cycle," *The Journal of Political Economy*, 1977, *85* (2), 239–263.

Maria, Mark De, Sim D. Aberson, and Katsuyuki V. Ooyama, "A Nested Spectral Model for Hurricane Track Forecasting," *Monthly Weather Review*, 1992, *120*, 1628–1640.

McCarthy, John F., "Decentralisation and Forest Management in Kapuas District, Central Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 2, Center for International Forestry Research, Bogor, Indonesia 2001.

Merrill, Robert T., "Tropical Cyclone Structure," in Holland Greg J., ed., *Global Guide to Tropical Cyclone Forecasting*, Melbourne, Australia: Bureau of Meteorology Research Centre, 2000, chapter 2.

Mietzner, Marcus, "Party Financing in Post Soeharto Indonesia: Between State Subsidies and Political Corruption," *Contemporary Southeast Asia*, 2007, *29* (2), 238–263.

Miguel, Edward and Gérard Roland, "The Long-Run Impact of Bombing Vietnam," *NBER Working Paper*, 2006, (11954).

Mills, David E., "Growth, Speculation and Sprawl in a Monocentric City," *Journal of Urban Economics*, 1981, *10* (2), 201–266.

Mills, Edwin S., "An Aggregative Model of Resource Allocation in a Metropolitan Area," *The American Economic Review*, 1967, *57*, 197–210.

Morduch, Jonathan, "Income Smoothing and Consumption Smoothing," *Journal of Economic Perspectives*, 1995, *9* (3), 103–114.

Murphy, Denis, "A Decent Place to Live: Urban Poor in Asia," Technical

Report, Asian Coalition for Housing Rights, Habitat International Coalition, Bangkok, Thailand 1990.

Muth, Richard F., *Cities and Housing: the Spatial Pattern of Urban Residential Land Use*, Chicago, United States of America: University of Chicago Press, 1969.

Nabuurs, Gert Jan, Omar Masera, Kenneth Andrasko, Pablo Benitez-Ponce, Rizaldi Boer, Michael Dutschke, Elnour Elsiddig, Justin Ford-Robertson, Peter Frumhoff, Timo Karjalainen, Olga Krankina, Werner A. Kurz, Mitsuo Matsumoto, Walter Oyhantcabal, N.H. Ravindranath, Maria José Sanz Sanchez, and Xiaquan Zhang, "Forestry," in Bert Metz, Ogunlade Davidson, Peter Bosch, Rutu Dave, and Leo A. Meyer, eds., *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom and New York, United States of America: Cambridge University Press, 2007, chapter 9.

Neuwirth, Robert, *Shadow Cities: A Billion Squatters, A New Urban World*, London, United Kingdom: Routledge, 2005.

Niehaus, Paul and Sandip Sukhtankar, "Corruption Dynamics: The Golden Goose Effect," *BREAD Working Paper*, 2009, *223.*

Niessen, Nicole, *Municipal Government in Indonesia: Policy, Law, and Practice of Decentralization and Urban Spatial Planning*, 1 ed., Leiden, The Netherlands: CNWS Publications, Leiden University, 1999.

Nordhaus, William D., "The Political Business Cycle," *The Review of Economic Studies*, 1975, *42* (2), 169–190.

Obidzinski, Krystof and Christopher Barr, "The Effects Decentralisation on Forests and Forest Industries in Berau District, East Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 9, Center for International Forestry Research, Bogor, Indonesia 2003.

Ohls, James C. and David Pines, "Discontinuous Urban Development and Economic Efficiency," *Land Economics*, August 1975, *81* (3), 224–234.

O'Keefe, Phil and Ken N. Westgate, "Some Definitions of Disasters," Occasional Paper 4, Disaster Research Unit, University of Bradford, Bradford 1976.

Olken, Benjamin A., "Monitoring corruption: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 2007, *115* (2), 200–249.

– and Patrick Barron, "The simple economics of extortion: evidence from trucking in Aceh," *Journal of Political Economy*, 2009, *117* (3), 417–452.

Parasuraman, S. and P.V. Unnikrishnan, eds, *India Disasters Report - Towards a Policy Initiative*, New Delhi, India: Oxford University Press, 2000.

Patel, Sheela, Celine d'Cruz, and Sundar Burra, "Beyond Eviction in a Global City: People-managed Resettlement in Mumbai," *Environment and Urbanization*, 2002, *14* (1), 159–172.

Poterba, James M., "State Responses to Fiscal Crises: The Effects of budgetary Institutions and Politics," *Journal of Political Economy*, 1994, *102* (4), 799–821.

Pugatch, Todd and Dean Yang, "The Economic and Health Effects of Hurricanes in Mexico," Technical Report, Working Paper 2008.

Rogoff, Kenneth and Anne Sibert, "Elections and Macroeconomic Policy Cycles," *The Review of Economic Studies*, 1988, *55* (1), 1–16.

ROI, "Undang Undang Nomor 5 Tahun 1967 tentang Pokok-pokok Kehutanan (Basic Forestry Law)," 1967. Republic of Indonesia.

— , "Peraturan Penerintah Nomor 21 Tahun 1970 tentang Hak Pengusahaan Hutan dan Hak Pemungutan Hasil Hutan (Government Regulation on Right of Forest Exploitation and Forest Product Harvesting Permit)," 1970. Republic of Indonesia.

— , "Peraturan Penerintah Nomor 62 Tahun 1998 tentang Penyerahan Sebagian Urusan Pemerintahan di Bidang Kehutanan kepada Daerah (Government Regulation on Delegation of Partial Authority in the Forestry Sector to the Regions)," 1998. Republic of Indonesia.

— , "Undang Undang Nomor 22 Tahun 1999 tentang Pemerintahan Daerah (Law on Regional Governance)," 1999. Republic of Indonesia.

— , "Undang Undang Nomor 25 Tahun 1999 tentang Perimbangan Keuangan antara Pemerintah Pusat dan Daerah (Law on Fiscal Balancing Between the Central Government and the Regions)," 1999. Republic of Indonesia.

— , "Peraturan Penerintah Nomor 34 Tahun 2002 tentang Tata Hutan dan Penyusunan Rencana Pengelolaan Hutan, Pemanfaatan Hutan dan Penggunaan Kawasan Hutan (Government Regulation on Forest Administration and the Formulation of Plans for Forest Management, Forest Utilization, and the Use of the Forest Estate)," 2002. Republic of Indonesia.

— , "Undang Undang Nomor 32 Tahun 2004 tentang Pemerintahan Daerah (Law on Regional Governance)," 2004. Republic of Indonesia.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India," *Journal of Political Economy*, 1993, *101* (2), 223–244.

Samsu, Heru Komarudin, Sian McGrath, Yan Ngau, and Dt. Iman Suramenggala, "Small Scale 100 ha Logging Concessions' Contribution to Regional Finance: Case Study in Bulungan District," Decentralisation Brief 2, Center for International Forestry Research, Bogor, Indonesia September 2004.

Samuelson, Paul A., "Economics of Forestry in an Evolving Society," *Economic Inquiry*, 1976, *14* (4), 466–492.

Sawhney, Kalidas and J.A. Daji, *Handbook of Agriculture: Facts and Figures for Farmers, Students and all engaged or interested in Farming*, New Delhi, India: Indian Council of Agricultural Research, 1961.

Shleifer, A. and R.W. Vishny, "Corruption," *Quarterly Journal of Economics*, 1993, *108 (3)*, 599–617.

Simha, Rakesh K., "New Delhi: The World's Shanty Capital in the Making," http://us.oneworld.net/node/66699 August 26 2006.

Skoufias, Emmanuel, Ambar Narayan, Kai Kaiser, and Basab Dasgupta, "Electoral Accountability, Fiscal Decentralization and Service Delivery in Indonesia," 2010. World Bank.

Smith, J., Krystof Obidzinski, Subarudi, and Suramenggala, "Illegal Logging, Collusive Corruption and Fragmented Illegal Logging, Governments in Kalimantan, Indonesia," *International Forestry Review*, 2003, *2003* (3), 293–302.

Soetarto, Endriatmo, MT Felix Sitorus, and M Yusup Napiri, "Decentralisation of Administration, Policy Making and Forest Management in Ketapang District, West Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 8, Center for International Forestry Research, Bogor, Indonesia 2003.

Stern, Nicholas, "Stern Review on The Economics of Climate Change," http://www.webcitation.org/5nCeyEYJr 2006. Accessed on 10th July 2010.

Strömberg, David, "Natural Disasters, Economic Development, and Humanitarian Aid," *Journal of Economic Perspectives*, Summer 2007, *21* (3), 199–222.

Tadjoeddin, Mohammad Zulfan, Widjajantai I. Suharyo, and Satish Mishra, "Regional Disparity and Vertical Conflict in Indonesia," Mimeo 38, UNDP/UNSFIR (United National Special Facility for Indonesian Recovery, Jakarta, Indonesia 2001.

Tambunan, Mangara, "Indonesia's new Challenges and Opportunities: Blueprint for Reform after the Economic Crisis," *East Asia: an International Quarterly*, 2000, *18* (2), 50–74.

Taylor, Colin J., "TGV Est lifts the record," *Railway Gazette International*, September 2007, pp. 553–558.

Topalova, Petia, "Trade Liberalization and Firm Productivity: The Case of India," Working Paper 28, International Monetary Fund February 2004.

Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), 539–591.

Toya, Hideki and Mark Skidmore, "Do Natural Disasters promote Long-Run Growth?," *Economic Inquiry*, October 2002, *40* (4), 664–687.

— and — , "Economic Development and the Impacts of Natural Disasters," *Economics Letters*, 2007, *94*, 20–25.

Tucker, Compton J., "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation," *Remote Sensing of Environment*, 1979, *8*, 127–150.

Turnbull, Geoffrey K., "Residential Development in an Open City," *Regional Science and Urban Economics*, 1988, *18* (2), 307–320.

— , "Squatting, Eviction and Development," Andrew Young School of Policy Studies Working Paper 0407, Georgia State University 2004.

Turner, Matthew A., "Landscape Preferences and Patterns of Residential Development," *Journal of Urban Economics*, 2005, *57*, 19–54.

Udry, Christopher, "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria," *Review of Economic Studies*, 1994, *61* (3), 495–526.

UN-Habitat, "Defining Slums: Towards and Operational Definition for Measuring Slums," Background Paper 2, Expert Group Meeting on Slum Indicators, Nairobi October 2002.

— , "Revised Draft Report," Technical Report, Expert Group Meeting on Slum Indicators, Nairobi October 2002.

— , *The Challenge of the Slums: Global Report on Human Settlements 2003*, London, United Kingdom; Sterling (VA), United States of America: Earthscan Publications Ltd, 2003.

USNO, *1999 Annual Tropical Cyclone Report*, Pearl Harbor, Hawaii: U.S. Naval Pacific Meteorology and Oceanography Center/ Joint Typhoon Warning Center,

van Vliet, Willem, "Cities in a Globalizing World: From Engines of Growth to Agents of Change," *Environment and Urbanization*, 2002, *14* (1), 31–40.

Venables, Anthony J., "Spatial Disparities in Developing Countries: Cities, Regions, and International Trade," *Journal of Economic Geography*, 2005, *5*, 3–21.

VerHoef, Jay M. and Peter L. Boveng, "Quasi-Poisson vs. Negative Binomial Regression: How should we Model Overdispersed Count Data?," *Ecology*, November 2007, *88* (11), 2766–2772.

Vermote, Eric F., Nazmi Z. El Saleous, and Christopher O. Justice, "Atmospheric Correction of MODIS Data in the Visible to Middle Infrared: First Results," *Remote Sensing of Environment*, 2002, *83*, 97–111.

Wan, Zhengming, Yulin Zhang, Qincheng Zhang, and Zhao liang Li, "Validation of the Land Surface Temperature Products retrieved from Terra Moderate Resolution Imaging Spectroradiometer Data," *Remote Sensing of Environment*, 2002, *83*, 163–180.

WB, "Brazil Progressive Low-Income Housing: Alternatives for the Poor," World Bank Report 22032, World Bank 2002.

— , "Decentralizing Indonesia: A Regional Public Expenditure Review," Overview Report, World Bank, Jakarta, Indonesia 2003.

— , "Dhaka: Improving Living Conditions for the Urban Poor," Bangladesh Development Series Paper 17, World Bank 2007.

Wheaton, William C., "A Comparative Static Analysis of Urban Spatial Structure," *Journal of Economic Theory*, 1974, *9* (2), 223–237.

Wooldridge, Jeffrey M., "Distribution-free Estimation of some Nonlinear Panel Data Models," *Journal of Econometrics*, 1999, *90*, 77–97.

— , *Econometric Analysis of Cross Section and Panel Data*, Cambridge (MA), United States of America: MIT Press, 2002.

Wulan, Yuliana Cahya, Yurdi Yasmi, Christian Purba, and Eva Wollenberg, "An Analysis of Forestry Sector Conflict in Indonesia 1997-2003," Governance Brief 1, Center for International Forestry Research, Bogor, Indonesia September 2004.

Wulder, Mike, "Optical Remote-Sensing Techniques for the Assessment of Forest Inventory and Biophysical Parameters," *Progress in Physical Geography*, 1998, *22* (4), 449–476.

Yang, Dean, "Coping with Disaster: the Impact of Hurricanes on International Financial Flows, 1970-2002," *B.E. Journal of Economic Analysis & Policy*, 2008, *8* (1), 1–43. Article 13.

Yasmi, Yurdi, Gusti Z. Anshari, Syarief Alqadrie, Tri Budiarto, Ngusmanto, Erdi Abidin, Heru Komarudin, Sian McGrath, Zulkifli, and Afifudin, "The Complexities of Managing Forest Resources in Post-decentralization Indonesia: A Case Study from Sintang District, West Kalimantan," Case Studies on Decentralisation and Forests in Indonesia 10, Center for International Forestry Research, Bogor, Indonesia 2005.

Zenou, Yves, "How do Firms Redline Workers?," *Journal of Urban Economics*, November 2002, *52* (3), 391–408.