Essays in Development and Political Economics

Mathias Iwanowsky

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Friday June 1 2018 at 14.00 in Nordenskiöldsalen, Geovetenskapens hus, Svante Arrhenius väg 12.

Abstract

This thesis consists of three self-contained essays in economics.

Property Rights, Resources, and Wealth: Evidence from a land reform in the United States: This paper compares the effectiveness of two alternative property rights regimes to overcome the Tragedy of the Commons. One regime is to distribute access rights under public ownership, as proposed by Samuelson, the other is to sell land to generate private ownership as proposed by Coase. However, as property rights are not randomly allocated, causal evidence on the relative effectiveness of these two regimes is scarce. I exploit a spatial discontinuity generated by the 1934 Taylor Grazing Act, which created 20,000 miles of plausibly exogenous boundaries that separated publicly owned rangeland from open-access rangeland. I combine these boundaries with data on the timing of private-property sales to jointly estimate the effects of public and private ownership on resource exploitation and income in a spatial regression discontinuity design. Using satellite-based vegetation data, I find that both property rights regimes increased vegetation by about 10%, relative to the open-access control. Census-block-level income data reveals that public ownership raised private household income by 13% and decreased poverty rates by 18%. To study mechanisms, I exploit variation in pre-reform police presence and panel data on farm values, and show that legal enforcement through police presence is a necessary condition for the positive and long-lasting effects of both regimes to arise.

State Repression, Exit, and Voice: Living in the Shadow of Cambodia's Killing Fields: This paper asks whether state repression is an effective strategy for silencing dissent and changing political beliefs. We use evidence from history's most severe episode of state-led repression, the genocide in Cambodia under the Khmer Rouge, to estimate the effects of political violence on political behavior four decades later. To establish causality, we rely on the Khmer Rouge's desire to create an agrarian society, moving forced labor to areas experiencing higher agricultural productivity. Using historic rainfall to generate exogenous variation in productivity shows that more people died in productive communes. Higher productivity under the Khmer Rouge leads to more votes in favor of the opposition over the authoritarian incumbent and increased support for democratic principles. At the same time, citizens become more cautious in their interactions with the local community as captured by lower participation in community organizations and less trust. Our results suggest that state repression makes people more convinced about the need for opposing views but more careful in expressing them, making politics less personal and more competitive.

The Effects of Migration and Ethnicity on African Economic Development: Migration between countries has been shown to have positive effects on economic outcomes such as trade by fostering economic and cultural integration. In Africa, where ethnic identification is reasonably strong, omitting ethnic links between countries likely introduces a considerable bias in the estimates. Following the literature, I use past migrant clusters as instruments to show that migration in 1990 led to more bilateral exports for neighboring countries in the period 1989-2014. To account for the ethnic heterogeneity of African countries, I generalize this approach and use pre-colonial ethnic linkages between of home- and foreign-countries as an instrument for migration. The results suggest a downward bias when not accounting for ethnic heterogeneity. I discuss potential concerns of pre-colonial ethnic linkages and find no evidence of omitted variable biases caused by similar languages, preferences, or conflict. Ethnic connections instead facilitate trade, especially for groups that are excluded from government coalitions. The results are consistent with a model of international trade where cross border connections decrease the fixed costs of exporting.

Keywords: *Development Economics, Property Rights, Ethnic Partitioning, Migration, State Repression.*

Stockholm 2018
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-155122

ISBN 978-91-7797-254-9
ISBN 978-91-7797-255-6
ISSN 0346-6892

Stockholm University

Department of Economics

Stockholm University, 106 91 Stockholm

ESSAYS IN DEVELOPMENT AND POLITICAL ECONOMICS

Mathias Iwanowsky

Stockholm University

Essays in Development and Political Economics

Mathias Iwanowsky

ISBN print 978-91-7797-254-9
ISBN PDF 978-91-7797-255-6
ISSN 0346-6892

Cover picture: Monument Valley, Utah. Taken by the author during the trip that inspired the first chapter of this Thesis.

Printed in Sweden by Universitetsservice US-AB, Stockholm 2018
Distributor: Institute for International Economic Studies, Stockholm University

An economists opinion:

From these considerations it follows that direct governmental regulation will not necessarily give better results than leaving the problem to be solved by the market or firm. But equally, there is no reason why, on occasion, such government administrative regulation should not lead to an improvement in economic efficiency.

Ronald H. Coase, “The Problem of Social Cost”, 1960, p.18

Acknowledgments

Whenever I think of obtaining a doctoral degree, as a first in my family, I think of my great-grandmother who called me 'the absent-minded professor'. Essentially no one who knew me back in Germany would've taken me for an academic, and many still reject this notion. I don't hold a grudge against you. I am puzzled myself. Instead, I laud today's world in which anyone, including me, can obtain a PhD and herald it as a feature of a world full of opportunities.

Naturally, I wouldn't have been able to finish this PhD without a countably infinite set of companions, whom I can't possibly pay justice to in these few pages. First, I want to thank my family, who gave me the ambition, directness, naiveté, loudness, and perseverance to embark on this journey. Second, similar to my biological family, I wish to thank my supervisors for their near endless patience with my lack of clarity, confusing thought processes, and my inability to make myself understood. I hope to have learned from Jakob's sense for simplification and focusing on the essential parts of a project, Torsten's sense for interesting 'big picture questions', and Andreas, who instilled a sense for identification and robustness in me that is borderline fanatic.

There were numerous other faculty members who had to endure my underdeveloped sense for nuance. To name a few, Konrad Burchardi and Anna Tompsett gave so much honest feedback in plain English, so even I could understand when to quit a project, although I never did. Jon de Quidt and Kurt Mitman guided me through the stormy, cold winds (quite literally -20 degrees in Philadelphia) of the Job Market, which I am enormously grateful for. And even more so, after having talked to other candidates without such professional preparation.

I also want to shout out to our amazing support staff. Christina always found a solution to any problem we encountered, and Annika was all ears for our moaning and who taught me one lesson: The first two coffees don't count. They only open the eyelids. Thanks goes to Åsa, Hanna, Viktoria, my Marvel buddy Ulrika, and Miriam, who were jointly responsible for lifting the mood significantly.[1]

In my daily live, I've encountered many PhD students who all contributed to this thesis in one or the other way. Richard contributed many lines of code, as mine was apparently 'disgusting', and helped co-found the CC team. Thanks goes to Jonna and Saman who enriched numerous lunch discussions. Daniel, Julia, and Sirus who introduced me to the Swedish way of life and happily invited us or tagged along when we needed a break from studying. The Italian team from SSE, Eleonora, Domenico, Andrea, Marta, and its honorary members Thomas and Nadiia who introduced me to good wine, pizza, and the Italian lifestyle. Naturally, my Italian team here at the Institute, Benni, Serena and Selene, helped me in various discussions about identification and the various hand gestures needed to survive in Milan. Thanks also goes to all 'new' members of the Stockholm PhD life. To John's never ending quest to win the pub quiz, Xueping's endless enthusiasm, and all members of Law Maximum Viskanic - next year is your year!

I also want to pay tribute to those who are no longer at the IIES, but whom I benefited from during my tenure here. Nathan, who introduced me to spatial econometrics which is the basis of all my work in this Thesis. Thorsten, who encouraged me from the very first day to start doing research, and had the patience to listen to my 99 problems although he had none. Alex, whom I failed to turn into a morning person and Audinga whom I enjoyed running, talking, cooking, and hanging out with.

I'll never forget my office mates. George, who gladly introduced me to his family and from whom I've learned that working 12 hours on your PhD is nothing compared to raising a family. Your work ethic encouraged me to keep pushing all these years. Josef, with whom I had my battles that helped me reflect on my personality, and with whom I enjoyed an amazing year at Harvard/MIT. Thanks goes to Andrea Guariso, who is

[1] Results not shown but are available upon request.

as welcoming a person as you can find. And of course my compatriots in the last two years, Divya and Matilda, who made the final sprint so much more entertaining and bearable.

I had the fortune to have four amazing flat mates, all of whom had to endure my 6:30 morning wake up routine. Leda, who never really moved out and whom I saw only on weekends, and the football/salsa genius Monti. Jaakko, who was always up for going out and amazed me with his research vigor and his relentless motivation to play any type of instrument for you. And last but not least Chris, who destroyed so many of my dinner plans by calling me to dinner when I already had eaten. Needless to say, two dinners are better than one.

Outside of Sweden, I had the fortune to have many friends, with even more fantastic nights out, that made this journey so much more tolerable. My friends from childhood, David and Felix, who visited me in the US and many times in Stockholm as well as my friends from Mannheim, Christoph, Philipp, and Marco, with whom I enjoyed more discussions and drinks than I can count.

Finally, my partner Dorothee, who started all this madness. Were it not for the countless hours in the library, trying to implant economic thought into you during our Master, I would have never gotten the grades to get here. She also was the most understanding person with me, when I was so focused on my work, that I would forget calling her. Or whenever we visited each other and I got up in the middle of the night to test an instrument or specification. Sorry for all this, but it won't get better trying to get tenure. But as to say with a line from one of my favorite songs: 'Du bist kein Hindernis für das was ich tue, du bist die Basis'.

But now, at last, it is time for you, the reader, to judge whether this thesis, forged by colleagues, family, and friends alike, is worth more than the sum of its parts.

Stockholm, June 1, 2018

Contents

Introduction

This thesis consists of three self-contained essays on topics central to development and political economics. Each considers a separate problem in developing countries today, and seeks out to provide answers based on past experiences. How should we allocate property rights to preserve resources for future generations without harming the livelihood of its users? How do people react to state repression and voice their discontent? How does migration and ethnic heterogeneity in African countries affect economic activity? In this thesis, I intend to provide insights into the above questions using primary data, many digitized and used for the first time, and rigorous econometrics to inform future policy discussions.

Property Rights, Resources, and Wealth The first chapter considers the ongoing debate on how to allocate property rights to common-pool resources. In modern day Africa, increased privatization has led to a decrease of agricultural and pasture lands used by the rural population for centuries. These common-pool resources were regulated by customary rights in which the duration of usage determined security and profitability of the resource (Goldstein and Udry, 2008).

As an increase in population pressured the remaining common-pool resources, the resulting over-exploitation decreased the income of its users. This situation was coined by Hardin (1968) as the 'Tragedy of the Commons'. Two solutions are discussed since. The first solution based on privatization is often traced back to an influential paper by Coase (1960). In a situation without transaction costs, total privatization of the resource will lead to an efficient outcome since the least profitable farmers sell their newly acquired property to high productive farmers. Importantly, the initial distribution of ownership is irrelevant as the resource can be traded freely. However, if the initial distribution is concentrated to one individual, individuals face capital constraints, or we impose transaction costs, efficiency is no longer guaranteed.[1] Hence, the other solution which is based on the public goods principle of Samuelson (1954) where an institution owns all the resources. By owning all resources and renting the right to use this resource at the optimal price to users, the efficient outcome can be achieved.

Importantly, both solutions, privatization or public ownership, yield the same outcome under the assumption of perfect markets and no transaction costs. The resource is preserved and its owner obtain a steady stream of income. However, as property rights form endogenously, economists have limited empirical evidence on key questions: Compared to a system without any form of property rights, should a policy maker privatize the resource or rent out access to this resource and manage it using an institution? Which solution is better for resource management? Which solution is better for farmers?

In this chapter, I provide evidence on these questions using a unique large-scale land reform in the United States in 1934. Here, more than 140 million acres were placed under public management and farmers with prior use were allocated rights to these grazing lands to graze cattle. Simultaneously, land was privatized up until 1934 based on a strict homestead rule, that allows me to compare the effects of privatization and public management on productivity. As the amount of vegetation is a good predictor of the amount of cattle a range can sustain, I use modern day satellite data and compare vegetation on either side of the fence.

The results suggest that if grazing rights are enforced, the initial predictions from Coase and Samuelson are correct. Privatization and Public

[1] Hence, the quote in the beginning of this Thesis. In two sentences, Coase reminds us that privatization is better than public management, but also that there exists no reason that public management can not be better than privatization.

management increase vegetation, measured 50 years later, by the same amount, suggesting that they are equally efficient. However, comparing farmers with access to these grazing ranges to farmers without, we observe higher income and wealth for farmers with access more than 50 years after the implementation of this reform. Importantly, these wealth results only occur in areas with lower transaction costs, suggesting that farmers traded these rights and reallocated into other professions.

Going full circle, policy makers in developing countries today were right to privatize resources in areas with low enforcement. However, the results from this chapter suggest, that if farmers are credit constraint or there is an information asymmetry between banks and farmers, public management might potentially be more efficient. For a farmer who used a common-pool resource before, a long-term contract for the same quantity and duration that is enforced and guaranteed by the state, is a property right. This increases the book value of this farmer and either alleviates credit constraints, or allows this farmer to sell his farm at a higher price than before. Importantly, invoking the Coase-Theorem once more, after all transactions have been conducted, the allocation is optimal. In this situation, if the efficient outcome is for few to own large swaths of land, this will happen regardless. The only difference is the direction of transfers which now involve farmers and not the government.

State Repression, Exit, and Voice The second chapter discusses an unfortunately ubiquitous post-conflict question. How do people react to observing atrocities? Aside the quantitative parts, there is a more fundamental question to it: How do we as researchers believe people should react to observing atrocities? On the one hand, if one observes a foreign force committing violent acts, one might assume that the optimal action is to increase 'in-group' cohesion to stand a better chance at preventing further acts of violence. On the other hand, if these acts of violence are conducted by members of your own group, distrusting strong leaders that call for such actions might be optimal.

Our starting point is one of the worst genocides in modern history where about 30% of the population was massacred by their own people based on their perceived usefulness to the state. When the Khmer Rouge took power in Cambodia in 1975, people were divided into groups of people useful to create an agricultural empire, and those who were not. People residing in urban areas, were sent to labor camps in the country side to increase productivity and eventually die of starvation, or be killed.

As the Cambodian society was built on strong patron-client relationships, where the trust in the strong leaders was strong, these atrocities found no resistance in rural communities. Our point of departure is then, how did people who observed these atrocities react? Based on seminal work done by (Hirschman, 1970) we develop a simple framework where the survivor has the opportunity to voice their discontent using the anonymity of elections or exit form civil society to reduce the probability of being detected as a dissident. By contrasting unobservable to observable actions, we are able to differentiate between a permanent change in preferences and the cost of being detected which is tied to the current leadership.

Our results suggest that observing violence changes preferences dramatically. Instead of trusting the beliefs and information of their local leaders, survivors inform themselves and form their own individual preferences. This leads to the rejection of strong leaders and to more democratic values, in particular at the ballot box. Moreover, people disengage from their local life, as the repercussions from being detected as a dissident have been ingrained by their experiences during the genocide. Voters hence use exit and voice, as they are willing to voice their discontent, but are careful about the means they do so.

As the electorate becomes more informed about the political process and develops stronger preferences for pluralism, elections become more contested and more competent candidates are elected. These politicians then find their ability to extract rents limited leading to better policies being implemented. As in developing countries the ability to extract rents is closely linked to selling local resources we look at the impact of deforestation in Cambodia. While still illegal in Cambodia, deforestation is wide spread in the country-side where corruption is rampant. We document a significant decrease in deforestation as a result of changed preferences as fewer land concessions are granted by politicians. Moreover, in line with our results on decreased trust, individuals themselves are less likely to participate in illegal activities that further destroys Cambodia's prime forest.

Migration and Ethnic Heterogeneity In the third chapter of this thesis, I consider the old, but relevant, question whether migration is good for society and try to inform a gap in the literature. While we have ample evidence that migration from rich countries to rich countries is enhancing exports and cultural experiences, migration from poor-to-rich, or poor-to-poor countries is far less well understood, despite constituting the majority of migration flows.

By focusing on exports between African countries, I aim at the heart of poor-to-poor migration but face two major obstacles to qualified research. First, both migration and export flows are measured with considerable noise. Second, the identification with ethnicity is much stronger than with nationality such that the standard approaches are likely to produce biased estimates.

In this paper, I use bilateral export and migration data to show that the stock of migrants in 1990 is correlated to exports in the period 1989–2014. To overcome potential endogeneity issues, I use the standard approach in the literature and instrument the stock of migrants in 1990 with their 1960s equivalent. However, as this standard approach is based on the assumption of national identity and migration in 1960 might be driven by initial income differences between countries that shape exports, I argue that this approach does not provide unbiased estimates of the true effect.

To overcome this potential bias, Africa allows me to draw upon a rich history of research on ethnic identification. In particular, I use a map showing the pre-colonial distribution of ethnic groups and intersect this with current country border to obtain a measure of ethnic connectiveness between two bordering countries in Africa. Importantly, since these borders were drawn by imperial powers in the late 19th-century, they do not reflect the preferences of groups in Africa. Moreover, since the distribution of these ethnicities is determined prior to independence, it is not caused by initial income differences and hence provides a reasonable instrument for migration.

Accounting for ethnic heterogeneity using the pre-colonial distribution of ethnicities I show that the estimated effects of migration on exports are about twice as large. To conclude that this estimate is unbiased by an omitted variable, I reject hypotheses based on language similarity, similar preferences that bind countries together, and conflicts.

Using data on government coalitions I find support for the hypothesis that ethnicities substitute favorable institutions for ethnic connections. The effects are larger for ethnicities that are a minority or actively discriminated against and are not part of government coalitions. I confirm this result using historical information about the political centralization of ethnicities. Groups that were historically more centralized were likely to capture more of the governments apparatus and hence rely less on ethnic connections.

The results from this chapter suggest two conclusions. First, the positive effects of split ethnicities on exports suggests that the view of African underdevelopment based on ethnic fractionalization is incomplete as ethnicities use existing ties to overcome barriers to trade and economic development. Second, since using their ethnic networks across countries is likely to be inferior to good institutions, including those ethnicities into government coalitions and improving institutions could further increase African economic development.

Chapter 1

Property Rights, Resources, and Wealth: Evidence from a land reform in the United States[1]

1.1 Introduction

Property rights are a central concept in economics, but how to best implement them is a contentious policy issue in many countries around the world. While it is well established that secure private property rights are important for productivity (Besley, 1995; Goldstein and Udry, 2008), income (Field, 2005, 2007), and wealth (Besley et al., 2012), little is known about the effects of poorly defined property rights to common-pool resources (Ostrom, 1990). Here, limited excludability of entrants and enforcement of rules may condemn farmers to poverty, if they have to compete with other farmers for use of the same resource. Resources are then extracted at a non-sustainable rate and many policymakers conclude that

[1] I am indebted to my advisors Konrad Burchardi, Andreas Madestam, Torsten Persson, Jakob Svensson and Anna Tompsett for guidance, numerous discussions and an open door. I benefited greatly from discussions with Gani Aldashev, Hoyt Bleakley, Jon de Quidt, Mitch Downey, Solomon Hsiang, Gary Libecap, Racheal Meager, Kyle Meng, Arash Nekoei, and Wolfram Schlenker. Thanks to many participants at conferences and seminars at Columbia, UCSB, ULB, and the IIES for valuable comments.

overcoming this so-called Tragedy of the Commons (Hardin, 1968) is an important step in ending poverty.

Many policy makers and economists agree that allocating and enforcing formal property rights solves the Tragedy of the Commons. While Coase (1960) advocated well-defined property rights on privately owned land, Samuelson (1954) argued for property rights on publicly owned land through a system of well-defined access rights. As the type of property right established is usually influenced by the productivity of a resource, empirical evidence on the relative effectiveness of these regimes is virtually absent. Moreover, in theory both property rights regimes can be efficient (Lindahl, 1919; Samuelson, 1954; Foley, 1970), but privately owned land may be preferable in areas with weak law enforcement or imperfect information, and less suitable for areas where transactions between individuals are costly (Coase, 1960). Hence, lack of causal evidence and theoretical ambiguity present a major challenge for policy makers.

This paper sets out to provide the first answers to two key questions: First, how effective are different property rights in reducing resource exploitation and improving living standards? Second, are there necessary conditions on enforcement, financial access and transaction costs for these property rights to be effective? In answering these questions, I focus on historical variation of ownership caused by the 1934 Taylor Grazing Act, which divided the livestock grazing ranges of the American West into different property rights regimes.

The main challenge in estimating the causal effects of the two property rights regimes is that their allocation is correlated with unobservable characteristics. In the context of the United States, high productivity lands are most suitable for farming crops and private rights tend to become established on these valuable lands. In contrast, lands unsuitable for farming are instead used to graze livestock and either formal access rights or informal property rights are established. As a result, estimating the effect of two property rights regimes may be plagued by biases, in particular due to the underlying differences in productivity. To achieve identification, I thus require a separate identification strategy for each property rights regime.

I overcome these endogeneity concerns by exploiting a spatial discontinuity in a large-scale land reform in the United States, namely the 1934 Taylor Grazing Act. The reform created large grazing districts in nine states to which renewable access rights stating a fixed price and quantity of livestock were distributed to nearby farmers.[2] By law, the total area of grazing districts was capped at 142,000,000 acres and exclusively selected from vacant and unappropriated open-access rangeland.[3] Grazing boundaries were drawn using plausibly exogenous grid lines from the Public Land Survey System, which were originally constructed to register land ownership during between 1851 and 1880. These grid lines were set within open-access rangeland 50 years prior to the reform, such that public ownership is quasi-randomly allocated in a very narrow bandwidth around the grazing boundary. Hence, close to the boundary, open-access rangeland serves as a valid control group to estimate the effect of access rights on resource extraction, defined as property rights on public lands.

To identify the effect of property rights on private lands, I explore the timing and location of purchased rangeland as recorded by the General Land Office. Since price and quantity were fixed under the Homestead Act of 1862, quality was the only margin of choice for buyers. As a result, the most productive rangeland was sold first and a comparison between the early privatized land and unsold rangeland would be biased by underlying productivity differences. Some decades later, the remaining public rangeland was of low quality and overgrazed. In these areas of

---

[2] Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, and Wyoming (see Figure 1.2). Prices were heavily subsidized and quantities determined after surveying the carrying capacity of the district.

[3] 'Open-access' and 'common pool' resources concern the same resource. While the resource is completely unmanaged in an open-access setting, common pool resources are managed by a finite number of people in a community. As argued by Bromley and Cernea (1989), many policy debates actually confuse the two, as the definitions are fluid, especially when local institutions are undermined. The implications here are unaffected by this distinction, as they are equivalent at the very local level.

arguably equal productivity, I provide evidence that a farmers' decision to purchase rangeland was uncorrelated with potential yields. As the grazing boundaries divided these areas, they enable me to compare private rights to nearby open-access rangeland holding productivity constant. To ensure comparability with the access rights treatment, I focus on privatized plots inside the grazing districts, since these plots would have been treated with access rights, had the reform been passed earlier. A private plot inside the grazing districts is thus defined as the private rights treatment if the plot was purchased between 1916 and 1934.[4] In this setting, land quality is balanced for all treatments inside the grazing districts and the open-access control outside the grazing districts. Land is thus quasi-randomly selected into 'private rights' and 'access rights', as well as the 'open-access' control.

Given the quasi-random allocation, I compare treatments and control in a regression discontinuity design. To proxy for productivity, I digitized maps of soil erosion in 1934 which were used to draw the original grazing boundaries and collected additional data on vegetation, temperature and rain. Additionally, I digitized maps of 6,830 minor civil divisions which I use to link census data on population statistics and individual characteristics to its sub-county division in 1930. I validate the identifying assumptions of the regression discontinuity by showing that all covariates are balanced and continuous at the boundary.

Using high-resolution satellite imagery as a proxy for productivity, I show that access rights and private rights have no differential effect on vegetation close to the boundary. Both property rights regimes increase

[4]The results do not change if I instead use privatized plots on either side of the boundary. The selection here is taken to illustrate the effects and has the same regression discontinuity design. In 1916, the Stock Raising Homestead Act quadrupled the available acreage farmers could purchase to 640 acres, in a response to the ongoing degradation in land quality. More flexible specifications involving decade-of-purchase fixed effects from 1864 onwards show that plots bought after 1916 are balanced in terms of productivity to the access rights treatment and the open-access control, while earlier ones were more productive.

vegetation by 10%, compared to open-access rangeland within one and five miles. Image recognition and machine-learning techniques confirm that this increase in vegetation directly translates into 25 million acres less of the least productive, but abundant, shrub land. In the Sub-Saharan Africa-context, a 10% increase in vegetation would imply a 30% (150 million acres) reduction in the lowest quality lands.

Using wealth and income data on more than 16,000 census blocks in nine states from 1990, 2000, and 2010, I show that establishing property rights raised family income by 13% ($5,000), increased the likelihood of completing high school by 4%, and decreased the poverty rate by at least 18%.[5] This program is likely to be welfare improving as the $71 million annual costs could be easily covered by a 1.4% tax on the additional income of people living close to the boundary.

Finally, I provide evidence against hypotheses of differential population growth, migration or privatization, and try to pin down the necessary conditions for property rights reforms to raise private wealth. Using pre-reform data on police presence to proxy for law enforcement, I show that vegetation and wealth increased only in areas with enforced access rights. Since enforced access rights validate off-farm income, farmers could either use the additional value as collateral or obtain a higher selling price. I use pre-reform data on bank presence to proxy for financial access and pre-reform data on newspapers to proxy for lower transaction costs. I show that financial access has no differential impact on long term outcomes. Lower transaction costs, however, greatly decrease resource exploitation and increase income and wealth.[6]

[5]Since census blocks are larger than the resolution of the ownership data, I estimate a compound effect of access rights and private rights on wealth. This is to my advantage, as comparing the wealth effect of access rights requires someone to live inside and outside the grazing districts.

[6]Importantly, the effect is not driven by the spread of information as radio penetration, defined as the share of people with radio in 1930, does not have the same impact as newspapers.

My combined results suggest that under ideal conditions, selling resources and renting out access to resources have the same effect on sustainability. An individual with enforced and exclusive rights to a resource is likely to extract resources at the efficient rate. My results show that farmers in areas with stronger enforcement and lower transaction costs may prefer access rights to privatization for two reasons. First, the least productive people get a wealth shock, fostering relocation to more profitable occupations. Second, as formal access rights may mirror informal existing rights, they might be easier to implement in developing countries where property rights are a contentious policy issue. Potentially, the results may be explained by alternative hypotheses, but I find suggestive and indecisive evidence against the hypothesis that police presence proxies for counties with better public service provision.

My results on vegetation contribute to the literature on managing common-pool resources (Ostrom, 1990). I complement evidence from experimental designs on the probability of destruction of resources and the impact of time preferences on exploitation (Walker and Gardner, 1992; Fehr and Leibbrandt, 2011). I add to the literature of economic effects of ambiguous property rights (Goldstein and Udry, 2008), and well-identified historical evidence on issuing private rights in colonial Congo (Vinez, 2017) or Liberia (Christensen et al., 2017). I extend this literature in two substantial ways. First, I exploit a geographical discontinuity to estimate the causal impact of access rights and private rights on resources and wealth at the same boundary. By extending the literatures view on private rights to access rights, this paper presents the first causal evidence on the relative effectiveness of both regimes. Second, by estimating the causal long-term effect of property rights on vegetation, my estimates are likely to encompass equilibrium effects, which would be hard to gauge in experimental and local settings.

My results on wealth add to the literature on property rights and wealth. The evidence on the wealth effect of access rights extends previous work on the effect of secure private rights on investments (Besley, 1995; Field, 2005; Hornbeck, 2010), labor supply (Field, 2007), assets (Besley et al., 2012), the distribution of income and crop choice (Montero, 2017), and human capital investments (Bleakley and Ferrie, 2016). I confirm the results in the literature that farmers benefit directly, and show potentially large spill-overs to the non-farming population in the same census-block more than 60 years after the reform. The wealth effects are larger than the estimates in the literature on private rights enforcement, probably because access rights are more equally distributed among farmers.

My findings on mechanisms shed some light on how higher wealth may come about. Using data on police presence in 1930, I confirm previous results on secure and enforced private rights (Besley, 1995; Svensson, 1998) and extend the implications to access rights where enforcement by the government arguably plays a larger role.[7] Secure access rights appear more important than financial access, even though access rights increase the value of collateralizable assets (De Soto, 2000). The results confirm that access to finance does have a smaller impact than enforcement on increasing investments by the poor (Johnson et al., 2002; Galiani and Schargrodsky, 2010). Consistent with recent evidence on First Nations' treaties in Canada, the introduction of enforced contracts increased incomes (Aragón, 2015) and alleviated the effects of mis-allocation of property rights by realizing the gains from trade (Chernina et al., 2015; Restuccia, 2016; Chen et al., 2017).

In what follows, I briefly describe the historical background and the Taylor Grazing Act before describing the data used in Section 3. I then highlight the identification challenges, the empirical strategy and validity of my approach in Section 4. In Section 5, I present the main results before discussing mechanisms and identifying necessary conditions in Section 6. I focus on the effects of farmers in Section 7, before discussing the implications of my findings. Section 9 concludes the paper.

[7] See Alston et al. (2000) for evidence from Brazil.

1.2 Background on the Land Reform

In many respects, the western United States in 1934 was similar to many developing countries today. Rangeland was mismanaged by local farmers and ranchers who used extra-legal methods to control public rangeland as privatization attempts failed to bear fruit. Combined with poor agricultural practices, this situation contributed to a severe land degradation, with the Dust Bowl (Hansen and Libecap, 2004; Hornbeck, 2012) being the most prominent example. To stop the ongoing land degradation, Congress passed the Taylor Grazing Act.

In this section, I briefly introduce how property rights were distributed prior to the Taylor Grazing Act and how the act changed the way in which property rights were defined in the western United States.

1.2.1 Privatization of the Public Domain

During the westward expansion of the United States, the federal government disposed of vast amounts of land. It considered these lands to be a source of revenue and handed over 72 million acres to eleven western states, 90 million acres to railroad companies, and more than 285 million acres to homesteading citizens.[8]

The first Homesteading Act of 1862 enabled citizens to apply for 160 acres of public land. After living on their homesteads for five years and documenting improvements to it, they were awarded land titles for a small fee of 10$.[9] Since price and quantity were fixed, the margin of differentiation for settlers was quality. As the homesteaded land needed to sustain its owners, the earliest plots usually encompassed the most productive lands. As productivity in the western states was generally low

[8]In their efforts to connect the coastal regions, railroad companies were partially reimbursed with lands close to the tracks. They were supposed to sell this land off to settlers, but many companies kept their lands as assets.

[9]The price to purchase land outright was 1.25$ per acre, a substantial amount in 1862.

and decreased further with ongoing privatization, Congress responded by increasing the acreage to 640 acres in 1916.[10] By the end of 1934, 236 million acres (38.9%) of the land area in the nine states had been sold to private individuals. As shown in Figure 1.1, the density of privatizations was lower in states such as Arizona and Nevada with large amounts of desert land. The continuous decrease in productivity created comparable plots – those that were sold shortly before the Taylor Grazing Act and those that were left unsold due to the passage of the act by Congress in 1934.

As most lands in the western United States were unsuitable for agriculture, farmers turned to grazing cattle and sheep.[11] To feed their livestock, farmers grazed their animals on their rangeland, as well as on nearby public rangeland. Once the first farmer reaped the benefits on public rangeland, this land was never able to fully recover from overgrazing and eventually became depleted. Farmers were painfully aware that they needed to overgraze public ranges without recovery periods, as this ensured their customary right to these ranges.[12] As these customary rights to open-access resources were threatened by an inflow of new farmers and a series of bad rain seasons, overgrazing on the public domain contributed to the

[10]Powell (1878) suggested that in order to make a profitable living, a homesteader required 2,580 acres in total. In 1877, there was another increase for some lands in the Desert Lands Act, but the act referred to here was the Stock-Raising Homestead Act of 1916.

[11]Many lands were also destroyed by agricultural technologies unsuited to local conditions (Foss, 1960).

[12]A New Mexico rancher in 1915:
"I can better afford to take the $2,500 loss of stock which I know I will have when the dry years come than to take my stock off my range and try to save some grass which I know I will need in those dry years. I hold on to my range only by having stock on it. If I take my stock off, someone else will take my range, and I can afford to lose the stock better than to lose the range."
Wooten, E.O. (1915) "Factors Affecting Range Management in New Mexico" U.S. Department of Agriculture Bulletin 211. In Figure 1.5, I show that the periods just leading up to the Taylor Grazing Act were especially severe.

'Tragedy of the Commons' (Hardin, 1968), which motivated Congress to explore potential solutions.[13]

1.2.2 The public grazing solution (Samuelson, 1954)

Following the proposal from stockmen associations in Montana, Congress established the first trial grazing district on public land in 1928. The Mizpah-Pumpkin Creek grazing district was intended to demonstrate the benefits of public management, as *"it was pretty generally conceded by 1920, that some sort of grazing regulation was imperative."* (Pfeffer, 1951). Ranges were subdivided into parcels to allow for recovery periods, and access was regulated to nearby farmers. The benefits were observed earlier than anticipated when a severe drought hit the western states in 1930 and the trial district went into the 1931 season with 20% more vegetation than adjacent rangeland. As the rangeland was also in better condition and the livestock heavier than on the surrounding rangeland, Congress decided to implement a similar solution on all of the remaining public rangeland (Muhn, 1987).

In line with many nature preserving acts of the time[14] and the disastrous effects of the Dust Bowl (Hansen and Libecap, 2004; Hornbeck, 2012), Congress enacted the Taylor Grazing Act on June 28, 1934.[15] It was enacted to:

[13]The Supreme Court of the United States affirmed the right to graze animals on public rangeland (Rundle, 2004).

[14]Antiquities Act 1906 and National Park Service Organic Act 1916, to name the most commonly known.

[15]The act was preceded by many state-specific laws, most of which aimed at discriminating against sheep, such as the "two mile law" in Idaho. However, none of these laws specified conservation of resources as an objective (Coggins and Lindeberg-Johnson, 1982). Prior to enacting the reform, the government tested the range management in the Mizpah-Pumpkin Creek grazing district in Montana in the early 1900s, which enlisted the support of some ranchers (Rundle, 2004).

stop injury to the public grazing lands by preventing overgrazing and soil deterioration, to provide for their orderly use, improvement, and development, to stabilize the livestock industry dependent upon the public range, and for other purposes.

As a first step, the Taylor Grazing Act prohibited future sales of the remaining public lands in the western United States.[16] As the act stipulated an upper bound on the acreage to be covered in grazing districts, the administration set out to identify the areas with the most need.[17] After an extensive soil reconnaissance survey in late 1934 and public hearings in early 1935, 49 grazing districts in nine states were established by 1936 (Figure 1.2). The clear upper bound on acreage left a number of essentially equivalent areas outside the boundaries that would have been treated had the limit been higher. I will use these areas as my control group.

In each grazing district, range surveys determined the optimal number of cattle or sheep a range can sustain. The resulting animal units per months [AUM] were divided between farmers and types of livestock. Farmers applied for access rights to the grazing district by stating the number of AUM they intended to use. Taking into account the use of ranges five years prior to that time and dependent property, the district board allocated farmers a fixed quantity of AUM at a fixed price of $0.05 per AUM.[18] Prices were kept low for a number of years to gain support among farmers, and revenues were only used within the grazing districts to improve water supply, re-vegetate the ranges and build roads

[16]The annual purchase of land decreased to approximately 200,000 acres from a peak of 18.3 million acres in 1910. See Figure 1.6 for the distribution of plots sold from 1900-1934 in my data.

[17]The original upper bound was 80 million acres, which was corrected to 142 million in 1936.

[18]Although this system was rather strict, elite capture by powerful farmers led to many court cases, as they allocated the majority of AUM between themselves and excluded small farmers (Calef, 1960; Libecap, 1981; Klyza, 1994).

and fences.[19] Since prices were low to enlist participation, independent range surveys determined the carrying capacity and farmers agreed on the need to intervene, this grazing solution may not have been far from fulfilling the Samuelson (1954) condition of optimal provision of public goods.[20] Access rights were issued for a period of up to ten years and almost automatically renewable to ensure cooperation by farmers. Furthermore, since revoking access rights when pledged as collateral was only possible in the case of grazing violations, these rights became de facto property rights tied to farms.

Combined with prior privatization of public lands, this institutional reform allows me to compare the effect of two types of property rights, private rights and access rights, in terms of their efficiency to overcome the 'Tragedy of the Commons'. As private individuals have the strongest incentives to behave in an optimal fashion, they serve as a natural benchmark for local resource management.

1.3 Data

I combine several sources of data to estimate the effects of property rights on resource management and wealth. I digitized data on land quality in 1934 and the historical grazing districts from archival sources. This is the first time that these data, covering more than 500 million acres in nine states, have been combined with historical data on the ownership of rangeland and the public land survey system to identify a causal effect of property rights.

---

[19]These range improvements were relatively inexpensive but were expected to have an economic impact since the quality was so low (Calef, 1960).

[20]The condition states that the sum of the marginal benefits is equal to the marginal cost of providing the public good. In the spirit of Lindahl's price system and the allocation mechanism with a cap on quantity, this solution is likely to satisfy this condition. Naturally, no farmer has the incentive to truthfully report her demand for the public good, but as long every farmer asks for $1 + x$ of her demand, the allocation is optimal.

As we have seen, the Taylor Grazing Act regulated the access to, and invested in, public ranges in nine states, with the intent to increase the productivity of rangeland. As the density of vegetation determines the number of livestock the rangeland can support, vegetation is a natural choice to proxy for productivity and was also surveyed at regular intervals by local offices. Unfortunately, local offices may have differed in their subjective judgment of productivity, and only a few original surveys remain.[21] To conduct a large scale, objective, and long-term analysis of the impact of property rights on productivity, I use modern satellite data on vegetation that covers treated and untreated areas of the United States. In this section, I introduce data sources for outcome variables and the main control variables and briefly discuss their construction.

Vegetation Satellite imagery captures different colors across the spectrum of light. Since measuring vegetation was one of the first applications of satellites, second only to espionage, the first Landsat satellites already had cameras that captured red and near-infrared lights. As plants reflect near-infrared light to protect themselves from overheating and soil absorbs near-infrared light, the relation between the red and near-infrared light allows me to identify vegetation from imagery. An example is shown in Figure 1.7a, where I show a test bed of crops together with the satellite vegetation index in Figure 1.7b. More dense vegetation is represented by more near-infrared light being reflected relative to red light. This ratio, called the normalized difference vegetation index [NDVI], is frequently used in the literature for resource management (Scheftic et al., 2014) and economics (Kudamatsu et al., 2016).[22]

[21] One of these surveys has been digitized by Skaggs et al. (2010). Although only available for a small part of New Mexico in 1936, this provides a useful balance test for my analysis. However, according to the authors, these early surveys were hard to classify, and thus, I only use their data as a balance test.

[22] The formula is:

$$NDVI = \frac{NIR - Red}{NIR + Red} \in [-1, 1]$$

In my main analysis, I use data from the Advanced Very High Resolution Radiometer [AVHRR] series. I collapse weekly data from 1989-2016 to reduce measurement error and error term correlations across time periods, as the treatment is cross-sectional (Moulton, 1986). I construct the NDVI from the red and near-infrared channel of the satellite at a pixel resolution of 1×1 km. I show the summary statistics for the estimation sample in Table 1.1, which already show a higher NDVI inside the grazing districts (3rd row).

I use various satellite series, the Moderate Resolution Imaging Spectroradiometer [MODIS] at a pixel resolution of 250 m and the Landsat NDVI index at 30 m, to check the robustness of the findings. Since these three data series come from different satellites and capture different wavelengths, they provide independent observations with limited error correlations.[23]

I use the AVHRR data in my main analysis since their resolution is closest to the size of sections in the Public Land Survey System used to administer ownership and balances the risk for spatial correlation.[24] Satellite imagery is likely to be spatially correlated, and more detailed data increase the severity of this problem. One pixel from the AVHRR data capturing a green forest is equivalent to 16 pixels in the MODIS data, which are perfectly correlated with each other. The downside of reducing the resolution of the data is the loss in power and the increased frequency of partially treated observations. Pixels with centers very close

where NIR stands for near-infrared light and higher values indicating more dense vegetation.

[23]Red wavelength with the AVHRR: 0.58-0.68 μm and near-infrared: 0.725-1.1 μm. Values for the MODIS version 'MOD13Q1' used here are 0.62-0.67 μm and 0.84-0.87 μm, respectively. MODIS data are constructed using Google Earth Engine. Landsat satellites are further sources, but due to the large amount of data, only the access rights treatment is tested (Table A.22).

[24]Since treatment is defined by the Public Land Survey System data, this constitutes the level of variation in the data. Then, since higher resolution data do not change the variation of treatment but risk spatial correlation, the AVHRR data are the preferred choice here.

to the boundary are partially treated because they do not align perfectly with the grazing boundaries. This increases the NDVI values for control pixels at the boundary, thus biasing the estimate downward.[25]

Since the values of the NDVI vary from satellite to satellite, level values are not comparable. Thus, since only the correlation between NDVI sensors has a clear interpretative value, I use image classification techniques to translate the satellite measures into land classifications. I use a random forest to predict land types based on NDVI, elevation, and temperature and then interpret the point estimate on property rights in terms of acres of quality land gained.[26]

Grazing districts Modern grazing districts are likely to have adapted and exchanged areas based on experiences after 1935. To avoid this potential selection bias, I digitize the original grazing maps from archival maps. On these maps, the grazing districts are referenced to the Public Land Survey System [PLSS], a system for administering ownership over the vast western lands.[27] This system dates back to Thomas Jefferson in 1785 and divides the western states into rectangular townships of 6×6 miles and every township into 36 sections of 1×1 mile each, based on reference meridians. Since these meridians were decided between 1855 and 1880 in the nine states, the PLSS is not affected by land quality and grazing districts in 1934. In implementing the Taylor Grazing Act, the grazing agency chose to fix grazing districts to be made up of these 1×1 mile sections so that the boundary of each grazing district corresponds perfectly to the boundary of the predefined areas.

[25] A solution is to drop these partially treated pixels. I show robustness to dropping these pixels in Tables A.6 and A.7.

[26] The land classifications are based on the official land classifications from the US Department of Agriculture.

[27] For a further discussion on the effect that this system has on property rights security, I refer to Libecap and Lueck (2011).

Ownership data Grazing districts were drawn on a large map, but the effective treatment areas varied by ownership status. National parks, national forests, Indian reservations and other reserved areas were not placed under the jurisdiction of the grazing districts and are thus removed from both control and treatment in my data.[28] Sections of townships with private property could also not be administered by the grazing administration. To identify which sections were privately owned in 1935, I web scraped the database on land transaction by the General Land Office [GLO], which provides information on the timing and location of private purchases using the PLSS (Figure 1.6).[29] Combining the two in Figure 1.2, historical grazing districts generally identify treatment and control areas (shaded areas), while the ownership status of every section determines whether a section was treated with access rights (shaded and white), private rights (shaded and gray) or open access (white). In the estimation sample (Table 1.1), 19.3% of the observations are defined as private rights (row 1), and 80.7% are defined as access rights (row 2).

Soil erosion in 1934 The Taylor Grazing Act began with a proclamation *"to stop injury to the public grazing lands"* and initiated a comprehensive soil erosion study covering the western United States in October 1934.[30] Based on this study, maps of soil erosion were drawn for all states, and the most severely damaged public lands were incorporated into the grazing districts. I digitized these erosion maps for nine states and show two levels of erosion in Figure 1.2. While Arizona, Nevada, and Utah are

[28] I drop every 6-mile boundary segment that has a national forest within 6 miles on either side.

[29] The data are available at the sub-section level but considerably more messy. To be conservative, I define a section as private if any part of this section has been sold. Since quarters of sections were sold before 1916 and entire sections after 1916, this does not affect the qualitative findings from the analysis.

[30] Generally, soil erosion is defined as loosened soil caused by cattle or sheep eating the grass that binds the soil. Similar to the Dust Bowl, where soil was blown away as far as Washington from the Midwest, soil without grass was washed away in excessive rains.

the most eroded states and thus have more grazing districts (Figure 1.4), I confirm in the balance section 1.4.4 that erosion is indeed balanced for treatment and control.

Minor Civil Divisions in 1930 As the Taylor Grazing Act covered vacant and unappropriated lands, less populated places were more likely to be included in grazing districts. Moreover, as a larger population is a good indicator of development and the presence of police, banks or newspapers, it is important to verify pretreatment balance to attribute contemporaneous wealth differences to the treatment. I use the grazing boundaries to test the balance of the population in my empirical setup using statistics for all minor civil divisions of each county. Every county in the United States was sub-divided into minor civil divisions in 1930 and 1940, for which data are almost universally available. I digitized the minor civil divisions for Arizona, Montana, Utah, and Wyoming for 1930 and all nine states in 1940 (Figure 1.9). When the 1930s equivalent was not available, I digitized the 1940s and followed given annotations to attribute the 1930s statistics. The population census in 1940 also has numbers for 1930 such that I attribute population statistics to 6,830 minor civil divisions in 1930 (6,537 in 1940) within 312 counties in all nine states. Linking the minor civil divisions to the 5% census sample in 1930, I collect information about families, houses and occupations at the individual level and link them to their geographic position in each county. Summary statistics of these variables are shown in Table 1.1.

Census data in 1990, 2000 and 2010 To estimate the long-term effects on wealth, I use census statistics at the census-block group level in 1990, 2000 and 2010. I obtain 16,248 geocoded observations for 1990, 15,701 for 2000 and 17,527 for 2010 and use information on median family income, median house value and the share of people below the poverty line to capture growth in indicators of long-term economic development (Table

1.1, last four rows).[31] These data can be used in the same regression discontinuity design since they are exceptionally detailed, with the mean size of a census-block being 3,600 acres.

1.4 Empirical Strategy

Property rights are not randomly allocated in space. Farmers choose the most productive sections to purchase, and likewise, access rights are distributed on land that supports livestock. To estimate a causal effect of property rights, I compare property rights within a narrow bandwidth not exceeding 5 miles around historical grazing boundaries based on sections from the Public Land Survey System. By choosing a narrow bandwidth and employing a regression discontinuity design, I capture control and treatment areas that were pretreatment equivalent.

Especially in a setting where I compare the productivity of areas that are in geographic vicinity, it is important that the functional form of latitude and longitude sufficiently captures productivity. Not capturing the underlying productivity of every boundary segment risks misinterpretation of pretreatment productivity differences for the treatment effect.

In this section, I discuss the identification strategy, which is based on the observation that grazing boundaries were set without taking local conditions into account. I introduce three specifications with different assumptions about how to capture the underlying productivity. Having discussed the identification strategy and the specifications, I conclude this section by defining my treatment arms and providing evidence that treatments and control are balanced at the boundary.

[31] The census-block groups change every decade, which is why I construct the data for every year separately before merging them into the final dataset. Census-blocks are significantly smaller than minor civil divisions and only available in recent history. Data are obtained from the National Historical Geographic Information System [NHGIS]. No individual data or earlier data are available at this level.

1.4.1 Identification Strategy

I aim to estimate a causal effect of property rights on resource management and wealth using a regression discontinuity design in a small band around the grazing boundaries. However, as the Taylor Grazing Act stipulated that 'vacant and unappropriated' lands be used, the majority of lands inside the grazing districts are likely of lower average productivity than lands outside the districts. Similarly, since the act was intended to include the most severely eroded public lands, the grazing districts contain worse lands almost by definition.

The act demanded that vacant land *"not exceeding in the aggregate an area of one hundred and forty-two million acres"* be combined in grazing districts.[32] Since lands just outside of grazing districts could well have been treated, had the limit been set higher, they are likely to provide a reasonable control group. To decide which areas were treated, the act stipulated that *"before grazing districts are created in any State as herein provided, a hearing shall be held in the State"*. With input from these hearings and the limitation on maximum acreages, some districts were approved, while others were not.[33] Since these districts were usually adjacent to each other and the boundaries between districts were determined for administrative reasons, the underlying land productivity should not vary significantly at the boundary. Similarly, if preferences influenced the decision to approve districts, it is unlikely that preferences change discontinuously at the boundary.[34]

[32]The original act in 1934 stated 80 million acres, but the situation was so bad that already in 1936, this limit was increased to 142 million acres. The current figure stands at approximately 155 million acres.

[33]An example of this is Nevada, as most of Nevada was suggested to be covered by grazing districts. After the hearings and reaching the upper limit of 142 million acres, they decided to leave out the entire center of Nevada and only focus on the edges close to the other states.

[34]However, preferences are unlikely to be a determining factor, as the overwhelming majority of farmers were pro-regulation at the time. Between 1903 and 1906, the Public Land Commission had surveyed a representative sample of farmers and found that 77% of the farmers who replied favored government control (Foss, 1960, p.42).

To precisely determine the boundaries of grazing districts, I used information about ownership provided by the Public Land Survey System [PLSS]. Since sections of the PLSS are 1×1 mile rectangular and reference lines that were set between 1855 and 1880, the grazing boundary was plausibly set orthogonal to local conditions. I visualize the identification strategy in Figure 1.10a, showing a typical township and its 36 sections. As the grazing boundary is a straight line, treatment and control are defined using the PLSS jointly with ownership data. This pattern is repeated in Figure 1.10b, where a farm is split in half since the PLSS is based on reference points hundredths of miles away without input from local geography. Without local knowledge, many grazing boundaries were set as straight lines for a number of miles, suggesting that treatment and control are quasi-randomly allocated in a wider range around the boundary. As shown in Figure 1.11, the grazing boundary was a visible fence, which separated properly managed rangeland on the right from severely overgrazed rangeland on the left. Importantly, the boundary did not align with other administrative changes, so I can rule out compound treatment effects and isolate the effects of interest (Keele and Titiunik, 2015).

Not all boundaries were placed within vacant land, as some boundaries were predetermined by national parks, national forests and Native American reservations. As the control areas of such boundaries are not comparable, I drop these observations and focus on boundary segments that were placed quasi-randomly in space.[35]

Even if boundaries are quasi-randomly placed, an RDD is invalid if the outcome variables are manipulated at the cutoff. Here, as farmers were caught in the 'Tragedy of the Commons' without incentives to add further cattle, establishing grazing districts on one side of the fence does

[35]The largest fraction by far are national forests, with 25% of the observations within a 6-mile bandwidth. As they are located in areas with higher productivity, I exclude boundary segments if a national park is within 6 miles. This dropping rule drops 11,400 miles of border segments and is robust to excluding any boundary segment with national forests, parks or Indian reservations that drop a further 4,800 miles of border segments. I show robustness to the dropping rule in Tables A.12–A.15.

not affect the optimality condition in the open-access areas. Another potential threat to manipulation arises from the hearings before the grazing districts were created. However, as many boundaries formed long and straight lines that split large farms (Figure 1.10b), local manipulation is unlikely to have been systematic.[36]

In brief summary, the maximum on acreage limited the overall size of grazing districts and created a control group of lands that would have been treated had the limit been set higher. Since boundaries were determined using a system of rectangular townships and sections constructed years prior to the act, land is quasi-randomly placed in treatment and control within a one-mile bandwidth around the boundary. Furthermore, similar to African borders, large parts of the grazing boundaries are straight lines such that areas further from the boundary still constitute valid controls.[37]

1.4.2 Estimation Framework

I follow the literature on geographical discontinuities and use standard versions of spatial regression discontinuity design [RDD] used in the economics literature (Holmes, 1998; Black, 1999; Dell, 2010). The design in its most basic form has two forcing variables in latitude and longitude and relies on two dimensions of choice. First, since the RDD only estimates the local average treatment effect at the boundary, I only compare treatment and control observations in a tight bandwidth of one, two, or three miles. Second, the functional form needs to capture any other variable that varies continuously at the boundary. Hence, I use three different specifications to capture potential continuous differences in productiv-

[36]Even if local manipulation occurred systematically, within one mile, vegetation is randomly placed in the sections of the PLSS and thus randomly assigned to treatment.

[37]In Section 1.B, I alter the identification assumption and use a Difference-in-Differences approach using the Agricultural Census and instrumenting treatment assignment. Despite different identification assumptions, the results confirm the conclusion established in the RDD setting.

ity in an increasingly parsimonious way and three bandwidths to show robustness at a local level.

Baseline The baseline specification controls for pretreatment productivity with a global polynomial in latitude and longitude and estimates the effect of property rights using a simple indicator variable:

$$\log(\mathrm{NDVI}_i) = \beta \times Treatment_i + dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)} \qquad (1.1)$$

Here, I regress the vegetation outcome on pixel i on a binary treatment indicator for whether its center is located inside the grazing district.[38] Controlling for boundary segment fixed effects $\delta_{b(i)}$, the distance to the boundary segment $dist_{b(i)}$ and a global second-order polynomial in latitude and longitude Γ_i, β identifies the local average treatment effect of property rights. Since a valid comparison requires geographic proximity between treatment and control, I only compare observations close the same boundary segment.[39]

Distance to the boundary segment and the global polynomial in latitude and longitude define predetermined productivity globally across the western states. These variables capture more productive areas in the north and less productive areas in the south. Essentially, I assume that productivity across the nine states can be represented by a continuous grid of productivity. The treatment effect is then identified as the differ-

[38] I use log transformation of the index for several reasons. First, the point estimate of this regression directly gives the percentage increase over the control group. Second, the log transformation puts more weight on the most destructed, targeted lands. Third, since negative values are not defined, this transformation ignores water bodies, which would be a bad control. I show robustness to using a standardized measure in Table A.16 for the AVHRR satellites, in Table A.19 for MODIS, and in Table A.22 for Landsat.

[39] I use 60-mile boundary segments and show the robustness to using six-mile boundary segments in Tables A.10 and A.11. However, in the main specification, this effectively compares 24 observations on the treatment side with 24 observations on the control side, severely reducing power.

ence between the expected productivity as defined by this grid and the actual productivity as indicated by the pixel.

Treatment in a spatial RDD is defined by two forcing variables, latitude and longitude. Controlling flexibly for latitude and longitude is sufficient, which is why I do not interact distance to the closest boundary with the treatment indicator (Lee and Lemieux, 2010; Keele and Titiunik, 2015). I compare observations within one mile of the boundary, and since the AVHRR data have a one-kilometer spatial resolution, the coefficient on distance is based on two observations away from the boundary.[40] In such a tight bandwidth, a local linear regression is often the better choice, which is why I use this as a baseline estimation.

Interacted with distance In my second specification, I interact the treatment indicator with distance to the nearest boundary segment to illustrate the discontinuity and the findings:

$$\begin{aligned}\log(\text{NDVI}_i) =&\beta \times Treatment_i + dist_{b(i)} \times Treatment_i + \\ &dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)}\end{aligned} \tag{1.2}$$

By adding the interaction $dist_{b(i)} \times Treatment_i$, this specification is closer to the standard regression discontinuity design based on vote shares in the political economy literature (Lee and Lemieux, 2010). Here, I assume that the productivity is additionally captured by the distance to the border and its functional form may change discontinuously at the border.[41]

However, as every observation within 0.3 miles of the boundary is partially treated, this specification is greatly affected by the choice of band-

[40] Since the bandwidth is 1.6 km, it allows for maximally two 1x1 km squares to be captured within the control and treatment.

[41] Here, I assume that there exists a mapping from productivity $\mathbf{P}^d_{b(i)} \mapsto dist^1_{b(i)}$ with the number of dimensions $d \geq 2$.

width, as it estimates the difference between the control functions of distance at the boundary.[42] Since partial treatment has a positive effect on control observations by increasing their vegetation index and treatment observations have a negative effect by decreasing their vegetation index, these functional forms are pivoted toward each other in small bandwidths.[43] The treatment effect is then identified as the difference between the expected productivity as defined by the global productivity grid and the slopes of productivity as measured by distance to the border. In a two-dimensional RDD, distance to the border is an inferior forcing variable, as treatment is solely determined by longitude and latitude. Hence, I include this specification to visualize the findings but draw inference from the baseline specification.

Boundary-specific productivity In my third specification, I estimate a different functional form of pretreatment productivity for every boundary segment (Dell, 2010):

$$\begin{aligned} \log(\text{NDVI}_i) =&\beta \times Treatment_i + \gamma_{b(i)}^{Lat} Lat_i + \gamma_{b(i)}^{Lon} Lon_i + \\ & dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)} \end{aligned} \tag{1.3}$$

Instead of the global polynomial chosen to capture productivity in the baseline specification (1.1), I allow the underlying productivity grid to vary for every boundary segment. This specification is flexible enough to

[42] The underlying resolution of the AVHRR data is 1 km (0.625 miles). Thus, every pixel that is as close as half this distance is partially treated.

[43] Essentially, partial treatment and productivity are functions of distance. In the smallest bandwidth of 1 mile of the boundary, the treatment effect is affected by the former function of distance since it has a relatively larger number of partially treated observations. Then, the treatment effect is no longer identified as the effect of property rights but as a combined effect of the two counteracting forces. Extending to the maximum bandwidth in the sample, the point estimates are indifferent from the baseline. Excluding partially treated observations, the point estimates are statistically indifferent to the baseline in most cases. The results are shown in Tables A.6 and A.7.

allow productivity to increase with latitude in some areas and decrease with latitude in others.[44]

The flexibility of this specification requires more variation in the data. As every boundary segment has its own latitude and longitude coefficient on top of the fixed effect, this specification requires more variation per segment for the central limit theorem to hold. Hence, especially with few observations in tight bandwidths around boundary segments, inference is affected as the number of observations approaches the number of variables.[45] However, especially with productivity, this specification captures every unobserved variable that varies continuously at the border and identifies the treatment effect exclusively from the discontinuity at the border. Hence, I use the baseline for inference and report this specification as a robustness test that captures productivity most conservatively.

1.4.3 Defining Treatment Status

In my setting, property rights can take the form of either private rights or access rights. Since the two treatments are based on their geographic location and ownership, I define treatment in this section.

First, the access rights treatment is defined as public lands inside the grazing districts, as nearby farmers could use them if they obtained grazing permits. The corresponding open-access controls are vacant and unappropriated lands just outside the grazing districts. Using a three-mile bandwidth around the boundaries, I show in Table 1.1 that 80.7% of the observations inside the grazing district received the access rights treatment.

[44]Here, I assume that there exists a mapping from productivity P to a different function $f()$ for every boundary segment: $\mathbf{P}^d_{b(i)} \mapsto \mathbf{f}(\mathbf{P})^2_{b(i)}$, with the number of dimensions $d \geq 2$. Essentially, the underlying productivity can be represented by a different hyperplane in every boundary segment. I use higher ordered polynomials to capture productivity more flexibly in two dimensions in Tables A.4 and A.5 .

[45]This issue had been noted in previous papers and has thus not been reported in Dell (2010) and only partially reported in Cantoni (2016).

Private lands inside the grazing districts are defined as the private rights treatment and grouped into decades of purchase. Because prices and quantities were fixed, quality was the margin of adjustment farmers used to choose plots. If farmers optimally decided to purchase the most productive plots that were available at the time, the average productivity of remaining plots decreased over time. Thus, a plot sold in 1880 was more productive than a vacant plot in 1935.[46] However, a plot sold in 1934 would have been treated with access rights had the Taylor Grazing Act passed one year earlier. Thus, the selection bias of differential underlying productivities between the two treatments should diminish closer to 1934 since the underlying productivity is continuous at the boundary, which I verify in the balance section.

An alternative representation of the private rights treatment is an indicator for being bought after the Stock-Raising Homestead Act in 1916. As early as 1878, officials discussed that in order to make a sustainable living in most areas in the western states, farmers needed at least 2,560 acres of land (Powell, 1878). However, up until 1916, farmers could only purchase up to 160 acres at subsidized prices from the government. Realizing the sluggish demand for plots, the federal government increased the available acreage to one entire section of 640 acres for the same price.[47] Therefore, 1916 makes for a natural break point in productivity, as a lower average productivity per acre was suddenly sufficient to sustain a living. Due to its simple interpretive value, I use this cutoff to plot the RDD graphs and conduct the heterogeneity analysis.

To fully exploit the local exogeneity of the boundary, I confine the treatment status for the private rights treatment to be within the historical grazing districts for all specifications. While a private plot outside

[46]Graphical evidence of this assumption is provided in Figure 1.14, where I test the balance of pretreatment covariates that might affect productivity. Earlier plots are always more productive than plots sold closer to the Taylor Grazing Act.

[47]To be conservative and due to data quality issues at lower levels of aggregation, I define every section as privately owned in any year if any record shows that any part of it was purchased by a private individual or a company.

the grazing districts is similar to a public plot outside the grazing district in 1934, it was never 'at risk' of being treated and thus violates an assumption of the identification.

To identify the long-term effects on wealth, I use census blocks that are larger than the resolution of the ownership data. Here, I define a treatment indicator for a census block being located inside the grazing district. A pure comparison between private rights and access rights is impossible since no individual lives in sections that were put under the grazing district; this thus estimates a compound effect. I compare farmers with access to open-access rangeland to farmers with access to regulated Taylor Grazing Act land.

1.4.4 Balance of Covariates

A valid spatial regression discontinuity design requires that predetermined covariates vary continuously at the border and are sufficiently captured by the polynomial in latitude and longitude. In this section, I present evidence in support of the local exogeneity of grazing boundaries at two levels. First, I present evidence that control variables for both the access rights and the private rights treatment for vegetation are balanced across treatment and control. Second, moving to population statistics and micro data, I show that the border was set orthogonal to population, income, wealth and other characteristics of the population.

To establish balance, I estimate the treatments separately using only public lands in 1935 (Figure 1.12) and the 1916 cutoff for the private rights treatment (Figure 1.13). All variables are indeed balanced at the boundary. As the Taylor Grazing Act was written to *stop injury to the public grazing lands by (...) soil deterioration*, Figure 1.12a for soil erosion and Figure 1.12b for pretreatment vegetation show the most important balance graphs. Clearly, as we move from 5 miles outside the grazing district to 5 miles within, rangelands are approximately 4% more eroded. At the boundary, however, treatment and control are balanced and continuous. Similarly, Figure 1.12b shows that in New Mexico, there were no pre-existing differences in vegetation at the boundary.

In the remainder of Figure 1.12, I show that other inputs to production, such as elevation, temperature, precipitation and accessibility, are balanced at the boundary. Even though rangeland is further away from cities and less accessible at 5 miles outside of grazing districts, Figure 1.12 shows balance at the boundary. This conclusion is supported by the point estimates in Tables 1.2 and 1.3.

The results carry over to the private rights treatment in Figure 1.13. In this treatment specification, I cannot test the additional identification assumption that the marginal productivity of rangeland decreases closer to the area affected by the Taylor Grazing Act. Hence, I allow for a more flexible specification in Figure 1.14 in which I group privatized rangeland into decades of purchase and report the point estimates of the regression on the covariates.[48] As expected, the earliest rangelands were less eroded due to inherent quality differences (Figure 1.14a). However, already from 1896 onward, there is no difference in erosion relative to the open-access control, i.e., the access rights treatment. Similarly, rangeland bought before the turn of the century is at lower altitudes, is less rugged, and is closer to rivers and cities. The smooth increase in the point estimates in Figure 1.14 provides additional evidence that land quality did indeed decrease over time and, importantly, that all treatments and controls are balanced from 1916 onward. The results suggest that my design provides valid counterfactuals to estimate the effects on resource management and compare the effects of access rights and private rights.

---

[48]Every figure is a separate regression including decades of purchase fixed effects and the access rights treatment. The specification is similar to the baseline in equation (1.1). Specifically, the estimation equation is:

$$\ln NDVI_i = \beta \times \text{Access rights}_i + \sum_d \gamma_d \times \text{Sold in Decade}_{d,i} + dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)}$$

and the Tables and Figures report the coefficients β and γ_d.

Since all covariates jointly determine erosion, I predict the erosion indicator using a probit and linear probability model. The results are shown in the Online Appendix, Figure A.1 and Table A.1. By reducing the dimension of productivity to one variable capturing the severity of erosion, I increase power to detect worse types of land. However, this reduced dimension capturing the probability of erosion is also balanced at the boundary.[49]

Minor civil divisions, a sub-county aggregation available in 1930 and 1940, are larger than a section from the PLSS, which is why the estimation features a simple treatment dummy for being inside the grazing districts.[50] This specification then captures the effect of access rights, as both sides of the boundary have privatized plots. I show that all population numbers and characteristics are balanced and continuous at the boundary (Figure 1.15).[51] As all covariates, including income and earning scores, are balanced prior to the reform, it is likely that any impact on income and wealth stems directly from property rights.

Taken together, the evidence provided here suggests that the Taylor Grazing Act in 1934 provides a valid quasi-experimental setting to evaluate the effects of property rights on resource management, income and wealth.

---

[49] The probability of erosion is balanced in the closest bandwidth of one mile and increases by 0.3 percentage points at the two-mile boundary. However, since the graph shows continuity and no other specification shows any significance, this is balanced at the boundary.

[50] Specifically, the estimation equation is:

$$Y_i = \beta \times \text{Inside Grazing District}_i + dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)}$$

[51] Estimation results are shown in Appendix Tables A.2 and A.3.

1.5 Results

To the owner, property rights guarantee exclusive access to a plot of land. As the Mizpah-Pumpkin Creek experiment showed prior to the Taylor Grazing Act, enforcing exclusive access increases productivity, as farmers are forced to adopt more sustainable grazing strategies. As farmers gain long-term property rights, these become a valuable asset to the farmer. Enforced property rights can then be used as collateral for investments or sold together with the farm to achieve a higher price.

This section documents a causal link between establishing property rights and wealth more than 60 years later. This timespan allows spill-overs and equilibrium effects to manifest themselves and provides an adequate picture of a large-scale property rights reform. The first part documents the impact of a change in property rights on the targeted outcome, resource management, and its impacts on wealth before moving on to potential underlying mechanisms.

In my setting, property rights are defined as both access rights to public lands and private rights on purchased land. I first separate their effects and compare the two solutions to the 'Tragedy of the Commons' by Samuelson (1954) and Coase (1960). Then, using their empirical equivalence, I estimate the joint effect of property rights on wealth using modern day census-block data.

Access rights as property rights The main result for access rights is visualized in Figure 1.16. I plot the residuals, controlling for boundary segment fixed effects, a flexible polynomial in latitude and longitude, and distance to the boundary in a five-mile bandwidth around the boundary.[52] Moving from an open-access regime on the left-hand side of Figure 1.16 to an access rights regime on the right-hand side significantly increases the density of vegetation.

[52]Each bin is 0.125 miles wide, and the confidence intervals are constructed using bootstrapped standard errors with the boundary segment as the sampling cluster.

The graphical finding is corroborated by empirical evidence in Table 1.4 for all empirical specifications. In all specifications and bandwidths, public and vacant land in 1935 that was put under government control shows a 7-12% increase in density of vegetation.[53] Issuing access rights to control the number of cattle on ranges, fencing of ranges and small improvements in water access have a substantial effect on the productivity of the land.[54]

For this to be a causal finding, it is important to compare pixels that were similar prior to treatment. In my setting, this requires the expected productivity function to be continuous across the threshold and sufficiently captured by the specification. In the baseline, I compare pixels along a 60-mile boundary segment and drop entire 6-mile boundary segments if there was a national forest on any side. In Table A.10, I make the comparison within an even tighter corridor of 6 miles along the boundary segment and up to 3 miles in bandwidth. Furthermore, Table A.12 and Table A.14 explore the sensitivity to dropping entire 60-mile segments if they incorporate a forest or keeping all boundary segments.

[53]For more specifications, see Table A.4, and for more bandwidths, see the left panel of Figure A.2. All point estimates are statistically indistinguishable in all bandwidths and specifications, except for the 'Interacted with Distance' specification. However, as argued in the empirical strategy section, the lower point estimate is due to partial treatment. I exclude partially treated observations in Figure A.4a and Table A.6 to show that the point estimates are stable and indistinguishable. More bandwidths are shown in Figure A.3. The results are more stable than the original bandwidths shown in Figure A.2, suggesting that partial treatment implies a downward bias close to the cutoff. Thus, as control pixels close to the boundary are partially treated and show higher NDVI values than their control pixels further outside the grazing districts, the curvature close to the cutoff explains the sensitivity of this specification in small bandwidths.

[54]These results support reports for the 'Mizpah River Pumpkin Creek' experiment prior to the Taylor Grazing Act, where it was noted that *"after three years (...) there is twice as much grass on the Mizpah as before, although the carrying capacity has been increased from 3,000 to 5,000 head"*. Harold Ickes, Secretary of the Interior, "The National Domain and the New Deal" *Saturday Evening Post* December 23, 1933, p.11.

Furthermore, as standard errors are likely spatially correlated, clustering by border segments might lead to wrong inferences. Thus, I explore different cutoffs for spatial clustering in Table A.9 and show that no reasonable assumption on spatial correlation affects the inference. Another way to calculate standard errors is based on randomization inference (Athey and Imbens, 2017). Drawing 200 random grazing districts in the nine states, I report the distribution of the point estimate and T-statistic in Figure A.6. In all six graphs, the baseline point estimates are clear extreme values of the distribution, indicating a significant baseline result.[55]

To translate the estimated treatment effect into changes from shrubland to grassland, I use a machine learning approach to classify lands in the nine western states. Using the U.S. Department of Agriculture cropland data layer from 2016, I identify the most common land usage types in my data.[56] Using these land types, I create a random training sample from the data and train a model to classify land types using the MODIS NDVI data, average temperature and elevation.[57] In Table 1.7 column (2), I show the total area in million acres for the top five land categories. Approximately 300 million acres are shrubland and 100 million acres are grassland, which can be used to graze cattle.

To capture varying degrees of non-linearities, I use two machine learning approaches. Both the support vector machine algorithm (columns 3-5), and the random forest (columns 6-8) results highlight their accuracy in predicting the distribution of land types in the baseline. Both

[55]I repeat the exercise in Figure A.10 for the wealth outcomes.

[56]The top 5 common land types are 51 % shrubland (also known as bush or scrub land), 22% evergreen forest, 11% grassland, 4% deciduous forest and 4% barren lands. Data from CropScape.

[57]Average temperature at the grid cell level and elevations are controls, similar to in a regression. The prediction algorithm is then conditional on temperature and elevation. The Moderate Resolution Imaging Spectroradiometer [MODIS] values for red and near-infrared from version 'MOD13Q1' used here are 0.62-0.67 μm and 0.84-0.87 μm, respectively. Data are processed using Google Earth Engine. The results are shown in Figure A.5a. Results focusing on heterogeneous treatment effects on shrublands suggest that approximately 10% are transformed into grassland, consistent with the results presented here.

algorithms suggest that a ten percent increase in NDVI in column (4) implies that more than 25 million acres of shrubland are transformed into grassland. Even though the random forest results are slightly smaller, they confirm the initial results that for every percentage increase in vegetation, more than 1% of the shrubland is transformed into grasslands.[58] An even larger impact is found in Sub-Saharan Africa, where the average NDVI is similar, but the impact is approximately 3% for each percent increase in vegetation.

In all specifications, access rights significantly increase vegetation compared to an open-access regime. While an important result, the efficiency of this treatment can only be compared using the benchmark of privatization.

Private rights as property rights To compare private rights to open-access regimes, I define comparable treatment and control groups. By grouping private lands into decades of purchase in Figure 1.14, I ensure that private plots sold close to 1934 are comparable to the open-access control and the access rights. From the results on vegetation in Table 1.5 and Figure 1.18, it is clear that plots sold in 1866 are 40% more productive than the open-access control.[59] However, plots sold after the 'Stock-Raising Homestead Act' in 1916 are comparable in effect sizes to the access rights treatment and suggest a 10% increase in productivity. Combined with balanced covariates (Figure 1.14), Figure 1.18 provides additional evidence in favor of a decreasing marginal productivity for rangeland sold closer to the Taylor Grazing Act.

[58] Using the land classification data as output, I find the same point estimates as on NDVI (Table A.23).

[59] The estimation equation is:

$$\ln NDVI_i = \beta \times \text{Access rights}_i + \sum_d \gamma_d \times \text{Sold in Decade}_{d,i} + dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)}$$

and the Tables and Figures report the coefficients β and γ_d.

Thus, I use the Stock-Raising Homestead Act in 1916 as a natural cutoff to have a simple division of treatments. Access rights treatments are defined as unsold plots in 1935 inside the grazing districts, while private rights treatments are defined as those plots sold after 1916. The open-access control groups are unsold plots in 1935 outside the grazing districts.[60]

The RDD graph in Figure 1.17 shows the same discontinuity as the access rights treatment before. In Table 1.6, I empirically test the equality of treatments and cannot reject the null hypothesis of equality in most specifications. Partial treatment heavily affects the 'Interacted with Distance' specifications, and excluding partially treated observations in Table A.7 results in a similarity of effect sizes in eight specifications, with only the largest bandwidth being significantly different at the 5% level.

Once more, to properly control for pretreatment productivity, I compare within 6-mile boundary segments (Table A.11), drop 60-mile boundary segments (Table A.13), keep all segments (Table A.15), and show various bandwidths in the right-hand panel of Figure A.2. To address concerns about the log transformation of the vegetation index, I show the point estimates on the standardized vegetation index and levels in Tables A.16 and A.17, respectively.[61] All results suggest that access rights and private rights have the same effect on vegetation within a tightly defined bandwidth around the grazing district boundaries.[62]

This conclusion is supported by the results in Figure A.5a and Table A.18, where I instead use the MODIS data. Since NDVI values are cal-

[60] Private plots sold prior to 1916 are excluded, as are private plots outside the grazing districts. Including the private plots outside the grazing district does not affect the point estimates (Table A.21).

[61] Although the index is normalized to lie within the interval [-1,1], it is affected by the actual values of red and near-infrared light. Thus, comparisons across sensors are only possible in terms of correlations. I standardize the values by their standard deviation and the mean to have a comparable index across sensors, similar to the log transformation. The results from the MODIS data are comparable (Table A.19).

[62] For spatially corrected standard errors, see Table A.9.

culated from different wavelengths compared to the original data source, the effects are slightly smaller but more stable across all bandwidths.[63]

The results point to an equivalence of private rights and access rights and suggest an answer to the first question. Both forms of property rights are better forms of resource management than open-access management. Furthermore, the equality of both forms of property rights provides support for the Coase (1960) hypothesis, despite a potential selection bias in the private rights point estimate. In a modest interpretation, the presented results show a first stage, where access rights were distributed that were as effective as private rights. While this equivalence speaks to critiques of either type of property rights, it also suggests that access rights could increase the value of assets, similar to property rights.

Property rights effects on wealth The equivalence of government intervention distributing access rights and privatization in terms of vegetation leads to the question of how property rights affect wealth. Enforced property rights for private rights increase the assets of a household. Whether access rights have a similar effect on wealth is an important question when considering welfare effects.

Using census-block data from 1990, 2000 and 2010, I show in Figure 1.19 how income, poverty and house values are affected. In this setup, I only separate between inside and outside the grazing district, as census blocks on both sides have both private rights and open access.[64] I refer

[63] The results are robust to the same robustness checks as the AVHRR data (Table A.20). The same effect is found using Landsat imagery. Here, the average effect is slightly smaller, but the consistency over three bandwidths and specifications highlights the robustness of this finding (Table A.22). The data are processed using Google Earth Engine.

[64] Specifically, the estimation equation is:

$$Y_i = \beta \times \text{Inside Grazing District}_i + log(\text{Size Census Block})_i + dist_{b(i)} + \delta_{b(i)} + \Gamma_i + \varepsilon_{b(i)}$$

where I include the size of the census block to have comparable census blocks. Results are qualitatively the same if dropped.

to the previous results and argue that the effect stems from government intervention distributing access rights.[65]

Families living inside the grazing districts have a slightly larger median income, are less likely to be below the poverty line, and have more assets in terms of housing. Thus, in terms of development indicators, census blocks inside grazing districts show more signs of development. I show in Table 1.8, among other things, that the median family income is 13% larger close to the boundary.[66] Strikingly, poverty rates consistently decrease by 2-6 percentage points, translating into decreases of at least 18% over the average rate of 0.12.[67]

The results suggest that secured tenure and freed capital allowed the farmer to re-invest in education and housing and subsequently grow out of poverty. Since census blocks cover every resident, they contain valuable spill-overs from farmers to non-farming community members. These spillovers are especially welcome, as they magnify the per dollar value of an intervention. Here, distributing the access rights at 7% of the cost

---

[65]It is likely a combined effect, but while the share of private plots is larger inside the grazing districts in 1935, there is no significant difference in 2010. Thus, the effect of private wealth accumulation should cancel out, leaving the access rights vs open-access comparison driving the results.

[66]In 1940, income is balanced at the border using the minor civil divisions. The results are not shown, as they are based on the 1% sample of the census and post-treatment data.

[67]Since these estimates are based on pooled data, I verify in Table A.24 that the results are consistent across all years. Moreover, the results are stable across the robustness checks in Tables A.25–A.29. Various specifications are robust (Table A.25), and dropping partially treated census blocks (Table A.26) increases the robustness to bandwidths as before. Similarly, neither spatially corrected standard errors (Table A.27) nor comparing only within 6 miles (Table A.28) or within the original sample from the satellite data (Table A.29) impacts the inference. Furthermore, the results are robust to narrow bandwidths of up to three miles (Figure A.7), extreme bandwidths of up to 200 miles from the boundary (Figure A.8), and excluding partially treated observations (Figure A.9). Furthermore, additional outcomes in Table A.30 show similar effects in per capita income, the number of bachelor degrees, population, mortgage shares, and social security and public assistance programs. Randomization inference based on drawing 200 random borders highlight the validity of my results (Figure A.10).

of private grazing fees generated income increases of $5,000 per household or $1,300 per capita for one million people living within 3 miles of the boundary. Comparing this to the $71,000,000 annual costs for only these tax payers, a 1.4% tax on the additional income of the people most affected would cover all costs.[68]

Issuing property rights, that is, access rights for public lands, or selling lands to individuals increases the vegetation and economic development. Moreover, there is an equivalence of Coase (1960) and Samuelson (1954) for the average effect on resource management as measured by the denseness of vegetation. However, pretreatment conditions might favor one solution or the other such that drawing conclusions for modern day policies requires an investigation into which conditions shape the effectiveness of property rights.

1.6 Channels

The grazing administration included the worst land types in the grazing districts. The results of the previous section show that these lands now feature substantially more vegetation and richer inhabitants more than 60 years later. In this section, I try to deepen our understanding of the mechanisms behind these results. I show that the worst types of land were the most affected by the grazing districts and that the most likely explanation for the impact on vegetation is an introduction of recovery periods without grazing during the winter months. Credibly enforcing

[68] Comparing this to the 6.4 million residents in these states, the additional tax revenue at the state level would be even smaller. Moreover, as all tax payers obtain some value from non-destroyed lands, this policy is likely to be welfare generating. The numbers are taken from Glaser et al. (2015). Revenues per Animal Unit per Month [AUM] in 2014: $1.35, with 8,594,442 AUM in 2012. Total grazing receipts in 2014: $14,585,000. Costs for Grazing Appropriations in 2013: $85,280,000. In 2016, the grazing fees were at least $1.65, and every state had surcharges per AUM of at least $2.56. Even taking into account foregone income where the price on the private range is approximately $20, the program would be welfare improving due to the immense per year increases in income.

farmers' access rights to ranges makes them willing to move the cattle off the range when needed. To proxy for farmers' beliefs about enforcement, I use the existence of police, as well as the quality of policing.

With increasingly secure property rights, a farm appreciates in value for its owner. Using property rights as collateral, a farmer may be able to obtain a larger loan or a higher selling price for the farm on the market. The data suggest no heterogeneous treatment effects of financial access and instead point toward increased consolidation of farms. Given that farm values increased, evidence suggests that more farmers sold their farms and moved to more profitable occupations. Such relocation is stronger in counties with lower transaction costs, which show larger increases in wealth. The persistence of the resulting wealth effects more than 60 years later suggests that the reallocation of individuals had positive equilibrium effects.

The results suggest that initial constraints on farmers' ability to reap the benefits from investing were lifted by enforcing property rights and that the re-allocation of property rights increased their productivity. Heterogeneous effects of enforcement, financial access and consolidation are likely to depend on the continued presence of these enabling institutions. However, as some institutions might respond endogenously to the Taylor Grazing Act, I restrict myself to predetermined variables. Since the enabling institutions are more likely to be present in cities, I show that population density, distance to the closest large city, or the grazing boundary itself does not predict the presence of these institutions in 1930. To validate the proposed mechanisms, I rule out confounding factors by showing that population growth, migration and privatization cannot explain the results, as each is balanced in 1930 and today.

Implementation of the Reform The Taylor Grazing Act aimed to improve the state of the worst lands in the western United States. Such lands can be characterized by a strong population pressure on resources, barren land or shrubland with the lowest NDVI values, or the most severely eroded lands. Using the minor civil divisions from 1930, I exploit high resolution population data to capture the severity of the 'Tragedy of the Commons'. In areas with low population, resources are unlikely to be over-exploited, as farmers do not overlap in their claims or self-organize to manage the resource (Ostrom, 1990). In Figure 1.20, I divide the population into quantiles and report the average marginal effects in each quantile for access rights.[69] Figure 1.20a on population shows that the effects are concentrated in the upper population quantiles, except for the densest areas. Since the fifth quintile is likely to correspond to more urban areas, the evidence here is consistent with access rights reducing the population pressure on rangeland.

Naturally, the population pressure should also affect the quality of the land. With the data at hand, I can derive two additional direct measures of land quality shown in Figure 1.20b and c. First, I divide the dependent variable into quantiles and show that the effect is driven by the lowest vegetation, shown in the largest quintile. Conditioning on the outcome variable, I estimate the effect in every quintile of the NDVI distribution, instead of the average effect. Together with the land-classification results in Table 1.7, this suggests that an impact is seen only at the lower end, thereby increasing the quality of the worst type of lands.[70]

Another measure of land quality comes from soil erosion. However, as erosion itself is driven by all factors of production, merely using soil erosion as an interaction does not capture the full underlying heterogeneity. Similarly, a heterogeneity analysis using all covariates lacks power and is

[69] The estimation equation interacts access rights (AR) and private rights (PR) with a dummy for each population quintile Q_d:

$$\ln(NDVI) = \beta_{AR}\text{AR} + \beta_{PR}\text{PR} + \sum_{q=2}^{5}[\delta_q + \gamma_{AR,q}\text{AR} + \gamma_{PR,q}\text{PR}] \times \mathbf{I}[\text{Pop} \in (Q_{q-1}, Q_q]]$$

Then, the average marginal effects for each quintile are calculated as $AME_{AR,q} = \beta_{AR} + \gamma_{AR,q} + \delta_q$, evaluated at the mean of the covariates. Figure 1.20 then plots all $AME_{AR,q}$ values for access rights.

[70] This effect is not driven by the weight that the log transformation puts on values close to zero, as the results in Table A.16 show for the average effects.

unlikely to yield significant results. To reduce the dimension of soil quality from the nine variables shown in Table 1.1, I predict soil erosion by the other covariates in a linear probability model.[71] The resulting continuous probability is then split into quintiles and the average marginal effects reported in Figure 1.20c. Lands that were more likely to be severely eroded in 1934 show greater improvements than lands that were not eroded.

Figure 1.21 indicates that the difference between the control and treatment may be explained by an introduction of a recovery period in the winter months. Farmers with secured and enforced access rights were willing to take their livestock off the ranges and let them recover.[72] The resulting productivity increases should allow farmers to have at least the same number of cattle per year, thereby significantly increasing their income.[73]

Enforcement of the Reform More secure tenure on access rights plots is only realizable with strong governance. In the nine states I study, strong governance is represented by the presence and quality of law enforcement. It is plausible to assume that the closer a farmer is to a police officer, the stronger the enforcement of the law is. It is also plausible that if this police officer is more competent, the farmer is more likely to believe that the law will be upheld. Such beliefs should affect both vegetation, as tenure on the plot is secured, and wealth, as the stronger property rights suggest a higher collateral value of the farm.

---

[71] The results are consistent using a probit model or using a random forest to predict the likelihood of classification into erosion based on the covariates.

[72] In contrast: *I hold on to my range only by having stock on it. If I take my stock off, someone else will take my range, and I can afford to lose the stock better than to lose the range."* Wooten, E.O. (1915) "Factors Affecting Range Management in New Mexico" U.S. Department of Agriculture Bulletin 211. The enforcement of recovery periods is consistent with evidence from qualitative surveys in the 1960s (Calef, 1960; Foss, 1960).

[73] The experiment in the Mizpah-Pumpkin Creek shows that even in dry years, farmers have more cattle on the fields for a longer period of time precisely because of secured tenure.

To proxy for governance, I use the existence of police in the 1930 full-count census at the county level and the distance of an observation to the closest city with a civil service reform.[74] As argued by Ornaghi (2016), these measures should gauge the availability and quality of governance in the early 1930s. Splitting the sample into counties with and without law enforcement officers, I find that law enforcement is a driving factor behind both the decrease in resource exploitation and the increases in wealth. These results appear in Table 1.9, columns 2 and 3. The same interpretation arises when interacting the access rights treatment with the distance to the closest city with a civil service reform. The results in Table 1.10 show that a one standard deviation increase in the distance to quality of governance negates the effect on vegetation and decreases the impact on wealth.[75] Naturally, since civil service reforms were enacted in larger cities, other covariates potentially correlate with distance to the closest city with a civil service reform. To alleviate this concern, I show in Table A.33 that distance to the closest city above a population threshold has no heterogeneous effect on my outcome variables.[76]

Economic Channels of the Reform The economic impact of property rights includes two channels, each of which is likely to depend on the enforcement of property rights. First, more secure property rights lead to a higher value of the collateral a farmer can post. Second, higher farm values should result in higher prices for farms, leading to a greater willingness to sell. To identify both channels, I use the presence of banks

[74]The data only have 31 cities with police reforms up until 1940, which I use in this paper. I thank Ornaghi (2016) for sharing. I follow her approach and count individuals with the occupation 'policemen and detectives' in every county and define cities with civil service reforms according to her data. In total, 84 counties had no policemen in 1930. Policemen are defined as individuals who work in the occupation class 'policemen and detectives'. There are 2,539 policemen in the nine states in my sample.

[75]Using an indicator and full results using all specifications shown in Table A.31.

[76]When including both at the same time, only the interaction with civil service reform persists. The results are not shown.

as a proxy for financial access and the presence of local newspapers as a proxy for the ease of placing farms for sale. Since these channels likely benefit from more secure property rights, I show the results for the entire sample, as well as for the sample of counties with law enforcement.

Property rights protection as highlighted by the presence of law enforcement increases the value of potential collateral. As property rights are ensured by the government, banks begin to accept access rights as collateral and issue more credit.[77] Farmers with larger collateral may invest more and grow out of poverty (De Soto, 2000). To obtain a credible measure of financial access in 1934, I use the Federal Deposit Insurance Corporation (2001) and divide the counties according to the existence of a banking institution.[78]

Dividing the sample in columns (4) and (5) of Table 1.9, I find that neither vegetation nor wealth is affected by the presence of banks in 1934. These results suggest that for access rights, increased collateral values had no impact on farmers' economic situation. Considering the nature of the ranching business and the marginal possibilities to increase the growth of young calves, these results are unsurprising.

Another effect of secure property rights stems from the higher sales price of farms. As previously non-verifiable off-farm income from public land is guaranteed with access rights, the price for farms with access rights increased. A subset of farmers retained their farms since prospective buyers did not compensate them for the non-verifiable off-farm income. With access rights, the increased selling price may lead to some farmers selling and switching occupations.[79]

[77] The text of the act also explicitly states that access rights cannot be revoked if they are pledged as part of a bona fide loan.

[78] There are 60 counties without a bank in 1934 in my sample. As financial access prior to establishing banks was mainly through post offices, I verify that the results are robust to using the existence of post offices in 1916 using data from Rogowski (2016). As the importance of post offices declined between 1916 and 1934.

[79] From the buyers side, the expected value is only the income from the farm. Once access rights document off-farm income, they are willing to pay more. See Appendix 1.C.1 for a simple model highlighting this fact.

To proxy this channel, I rely on the presence of local newspapers. Farmers post ads for their farms, including price and grazing rights, in local newspapers where other farmers may search for potential farmland with additional grazing rights. Hence, the availability of local newspapers decreases the transaction cost for buyers and sellers. Using data from Gentzkow et al. (2014), I divide the sample into counties with local newspapers in 1932 and those without. The evidence presented in Table 1.9 shows how important this selling channel was for the effects of the Taylor Grazing Act. Vegetation increases, suggesting that more productive farmers remain in the county, and indicators of wealth go up only in areas with newspapers. Since newspapers also transmit useful information for farmers, I use data from Strömberg (2004) on the share of people having access to radio. In Table A.36, I interact radio share with treatment to separate out the effects of information and farm advertisement. The results on radio share show a significant average impact on wealth indicators. However, once I condition on the existence of a local newspaper, this correlation is insignificant (column 6).[80]

Economic channels depend on the farmers' beliefs about enforcement of the reform. Thus, I explore the interaction of these channels and enforcement in Table 1.11. By conditioning on the presence of police in counties, I isolate the importance of financial access and market consolidation, given that the reform is enforced. Once more, the results suggest no impact of financial access but a positive impact of lower search costs via local newspapers.

To conclude the discussion about channels, I provide evidence against differential migration, population growth or privatization driving the increase in wealth. If issuing access rights increased the value of farms,

[80]Interacting predetermined variables solves potential endogeneity problems, but especially in the case of radio, the interaction captures multiple channels. It is thus hard to directly disentangle the information channel. I argue that jointly with newspapers, the information channel is sufficiently captured by the interaction of treatment with radio share.

this could have been accompanied by an inflow of farmers. Evidence in Figure 1.22, however, suggests that both in 1940 at the individual level and in 1990 and 2000 at the census-block level, migration was balanced for the treatment and control. There is no effect for farmers who have been active in 1935, nor is there a significant difference in the tendency to migrate in any variable from 1985–2000. Thus, as land sales have only been allowed restrictedly since 1976 and had been rare before, it is unlikely that intentional migration into grazing districts could explain an increase in wealth. Corroborating this hypothesis, Figure 1.23 shows no differences in the contemporaneous population (top panel) or modern day privatization (bottom panel). In fact, the privatization rates seem to have increased outside the grazing districts.[81]

Combined, the evidence presented here suggests that a greater police enforcement and a better ability to advertise farms with access rights are important determinants of the effectiveness of property rights. Similar to the previous literature (Johnson et al., 2002), I show that secure and enforceable property rights have a larger effect on wealth than access to finance.[82] Since the previous findings include the entire population of each census block 60 years after, the estimated impact includes important equilibrium effects.

1.7 Policy Discussion

Despite having documented the positive effects, public management of common-pool resources in the United States is a contentious policy issue. While the Bundy family in Nevada (2014) and Oregon (2016) fought to abolish the status quo, ranchers in Montana fight to protect the system

[81]Using the agricultural census and a differences-in-difference estimation, I show in Tables 1.B that, if anything, more farms were sold and farm sizes increased, indicating a consolidation of the local economy.

[82]The results are unlikely to be driven by a greater development potential of treatment census blocks since all covariates, as well as police, banks, and newspaper divisions, are not predicted by treatment.

they require to make a profitable living.[83] Similar incidences are common in Kenya (2017), where cattle herders violate property rights and influence elections, or in Ethiopia (2017), where privatizations to foreign investors, so-called large-scale land acquisitions, threaten the life of the average Somali shepherd.[84]

Especially in Africa, informal property rights have contributed to large-scale land acquisitions in Africa (FAO 2009). Under such customary systems, either the village chief or the most tenured farmer controls who has access to land. These lands are vulnerable to being sold on the private market, as the government does not recognize these customary claims. Here, the largest bidder usually promises to invest in the local communities, but since customary rights and formal rights are formally at odds, many promises are left unfilled (Christensen et al., 2017). The selling of these lands has the largest impact on marginalized farmers who do not own a land title to their plots (Knight, 2010).[85]

In such situations, formalizing customary rights into enforced access rights could prevent this extraction of land from the rural poor at no cost of efficiency in managing the resource. On the contrary, the results in this paper show that many farmers could benefit from the allocation of access rights and would potentially even trade in the land for the opportunity to switch occupations. Moreover, since some customary systems disadvantage women, formalizing such rights and recognizing the status of women could improve the economic security for many families. Moreover, since

---

[83] The Bundy standoffs were nationally televised in both years and cost the life of one supporter in 2016. Newspaper article in Montana: `https://www.theguardian.com/environment/2017/jun/06/montana-land-transfer-american-ranchers`.

[84] Newspaper articles about Kenyan cattle grazers: `https://www.theguardian.com/environment/2017/feb/02/armed-herders-elephant-kenya-wildlife-laikipia`, and Somalian shepherds: `https://www.economist.com/news/middle-east-and-africa/21723155-well-adapted-desert-not-modern-world-hard-life-somali`.

[85] In conversations with Konrad Burchardi about his study in Tanzania, 301 of the 968 farmers who said they own the plot do not possess a formal land title to verify their claim.

many of the benefits arise from a less controversial formalization of customary property rights in terms of access rights that benefit more people, they could lead to substantial decreases in poverty across the developing world.

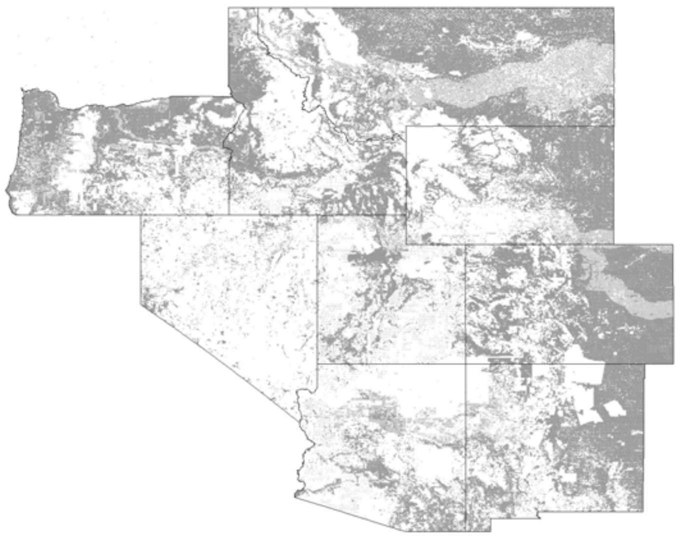
1.8 Conclusion

The results in this paper suggest that government intervention to establish collective access, in the spirit of Samuelson (1954), can have the same effects on resource management as outright privatization, in the spirit of Coase (1960). While property rights are contentious, distributing regulated access rights to all previous users should make this policy more appealing to policy makers. Moreover, as more users benefit from such a policy, wealth effects may be distributed more evenly and decrease poverty rates. However, for such policies to be effective, the issued access rights need to be enforced and easily transferable between farms, as the ability to consolidate and relocate greatly increases the effectiveness of this policy. As access rights document off-farm income, the valuation of sellers and buyers are more aligned, implying more farm sales by the lowest productivity farmers. Then, access rights overcome a market friction, and the relocation into more productive occupations should then imply a net welfare gain for society.

Combined, the results suggest that under ideal conditions, selling resources and renting out access to resources have the same effect on sustainability. As soon as an individual has enforced exclusive rights to a resource, he or she is likely to behave optimally. However, in areas without strong enforcement, privatization may be preferable to access rights. With stronger enforcement and low transaction costs, distributing access rights is preferable to private rights for two reasons. First, more people obtain a wealth shock that leads the least productive farmers to relocate to more profitable occupations. Second, as formal access rights mirror informal existing rights, they might be easier to implement in developing countries.

1.9 Tables and Figures

Figure 1.1: Private rights



Lands sold before 1935 in nine western states shown in gray. Data shows a clear correlation between moving westwards into more desert like regions and privatizations. Data taken from the General Land Office.

Figure 1.2: Access rights

Lands sold by 1935 in nine western states with the extent of the Taylor Grazing Districts overlaid. Treatment is defined as follows. Access rights: Shaded areas with white background. private rights: Shaded areas with grey background. Open-access control: Unshaded white background.

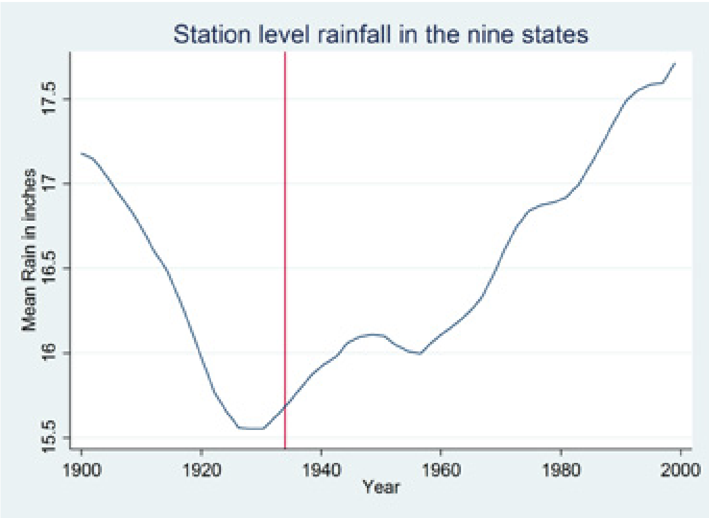
Figure 1.3: Erosion status

Erosion Status in nine western States (October 1934). Severely eroded: Gray areas. Moderately eroded: Gray shaded areas.

Figure 1.4: Erosion Status with Grazing Districts

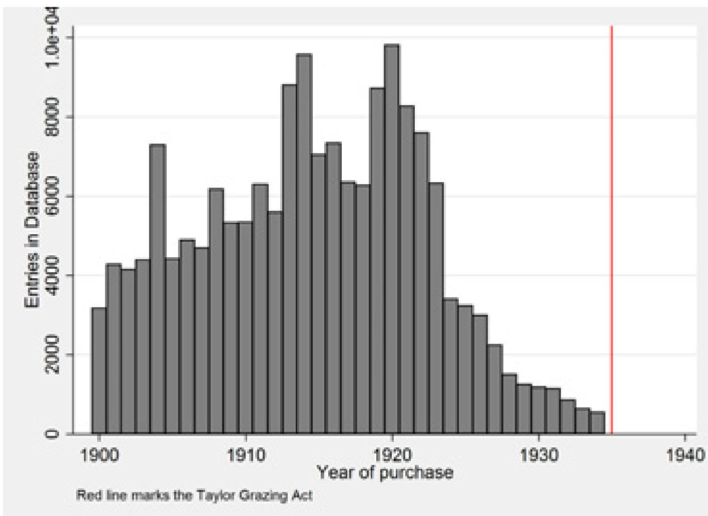
Erosion Status in nine western States (October 1934) with the extent of the Taylor Grazing Districts overlaid. Severely eroded: Gray areas. Moderately eroded: Gray shaded areas. Grazing districts: Solid shaded areas. Data for the east of Colorado is missing.

Figure 1.5: Station level rainfall



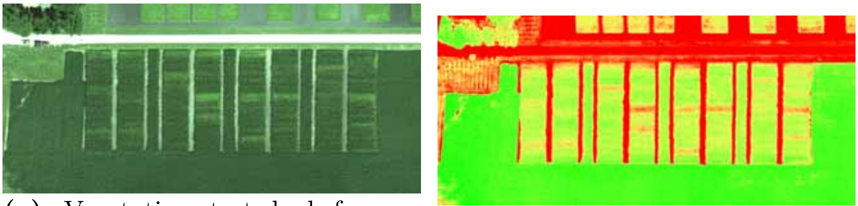
Time series of station level rainfall during in the last century. The years prior to the Taylor Grazing Act (1934) were particularly sever in terms of rainfall, increasing the pressure to pass regulations.

Figure 1.6: Rates of privatization



Histogram of purchases within Sections of the PLSS by year of purchases in the GLO data. The peak years were more than ten years prior to the Taylor Grazing Act (red line), indicating a lower demand due to low quality of the remaining lands.

Figure 1.7: Normalized Difference Vegetation Index (NDVI)

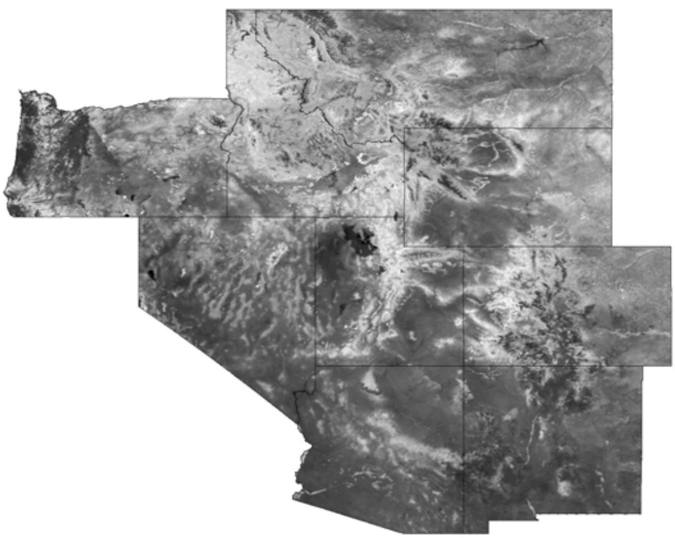


(a) Vegetation test bed from a satellite picture.

(b) As measured with in NDVI.

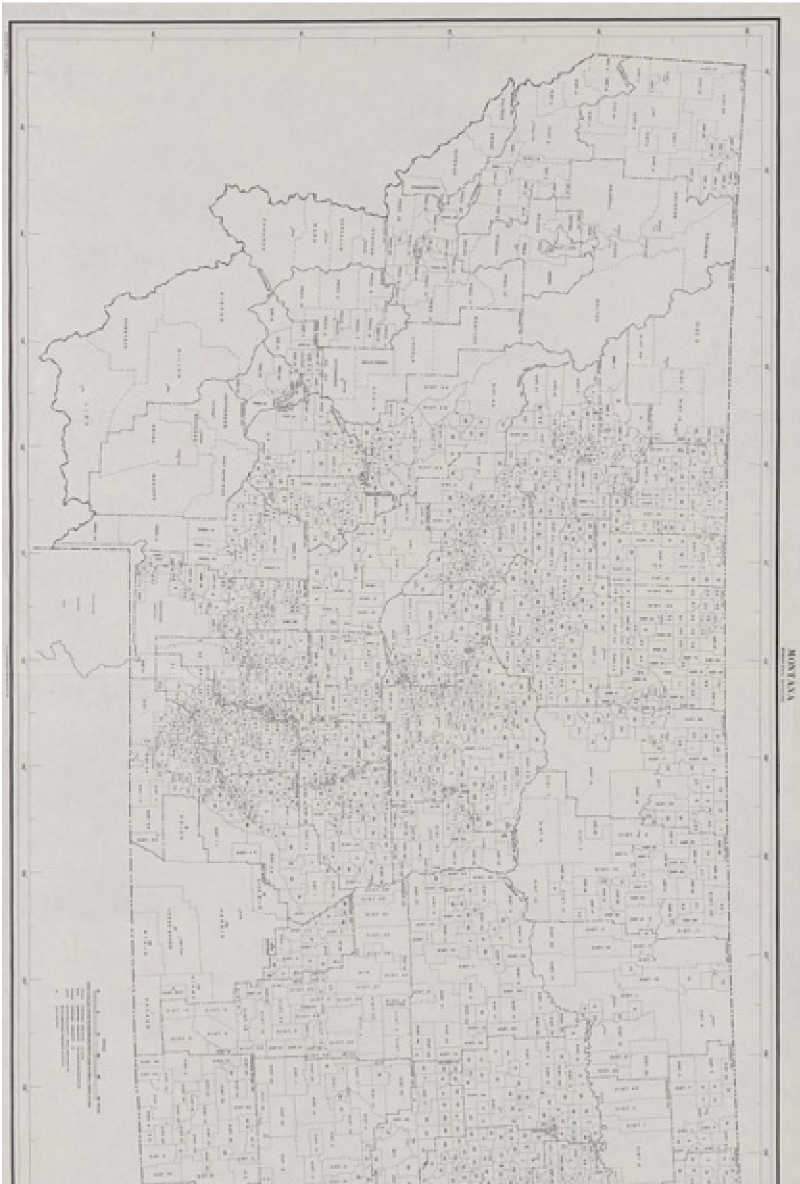
This figure shows how the NDVI values (right) capture the different shades of vegetation in the left picture.

Figure 1.8: Average in-sample NDVI



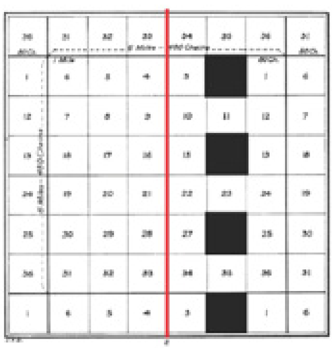
Average NDVI values in nine western states during 1989–2016. Desert like regions in Nevada are shown in dark gray while forest regions in Idaho are shown in lighter colors.

Figure 1.9: Minor Civil Divisions

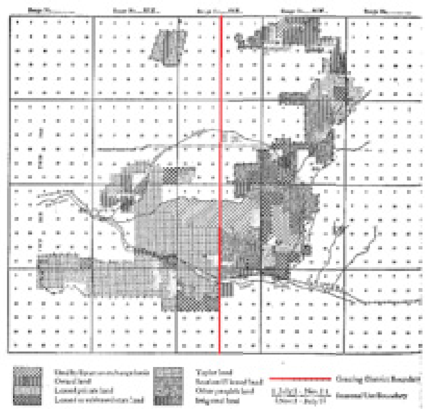


1,177 Minor civil divisions in Montana in 1930 with county information

Figure 1.10: The Public Land Survey System

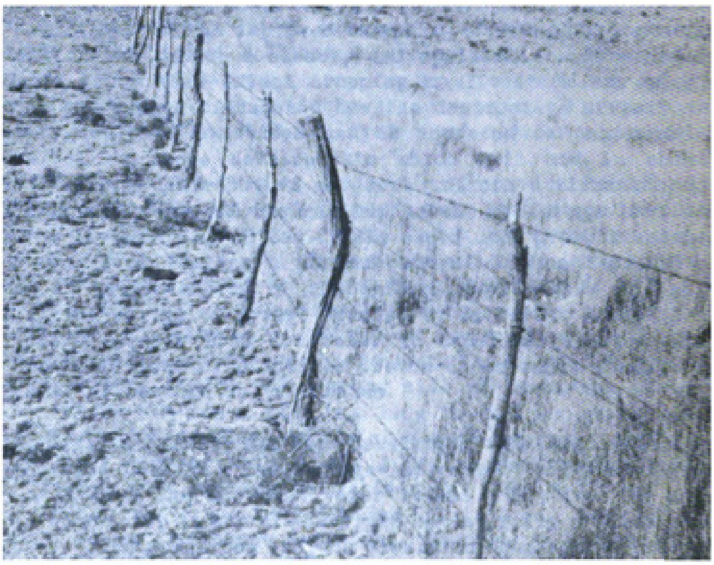


(a) A typical township from the PLSS with 36 sections a 1 x 1 mile. Grazing border marked red, privately owned sections by sold color.



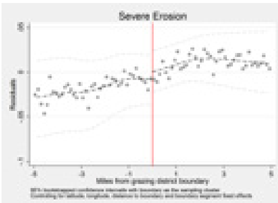
(b) The grazing border splitting a ranch in Wyoming over 24 miles. Figure from (Calef, 1960).

Figure 1.11: Identification example

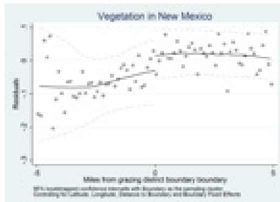


Identification in a picture as explained by a grazing bulletin in 1940: *"The pasture on the right is representative of properly used range. The one on the left has been cropped dangerously close. [...] by annually harvesting only [the optimal] amount of forage [... and] by adjusting the grazing season to permit maximum forage production under use, and by obtaining uniform utilization by proper distribution of livestock and income may be realized."*

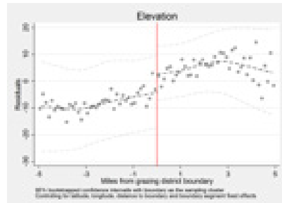
Figure 1.12: Balance graphs: Access rights treatment



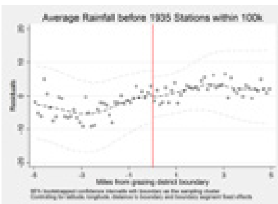
(a) Severe Erosion



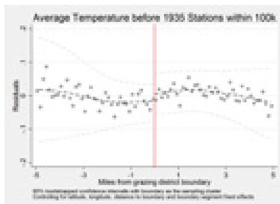
(b) Vegetation in NM



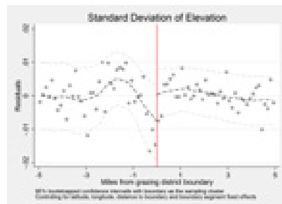
(c) Elevation



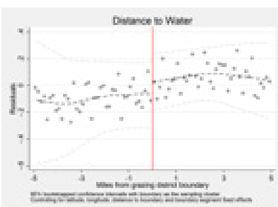
(d) Rainfall 1935



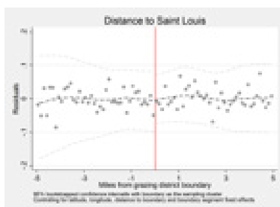
(e) Temperature 1935



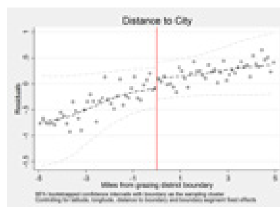
(f) Ruggedness



(g) Distance to nearest river



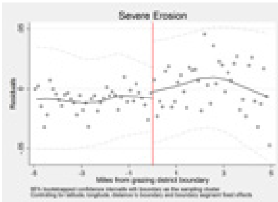
(h) Distance to Saint Louis



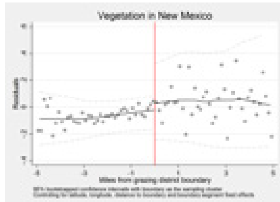
(i) Distance to closest city

Balance regression discontinuity graphs for the access rights treatment in the AVHRR data: Plotting the residuals, controlling flexibly for latitude and longitude, distance to the border as well as boundary fixed effects. Each bin is 0.125 miles wide. Balance table in Table 1.2. Variable description: *Severe Erosion* refers to erosion maps constructed for the nine states in 1934. Those maps were used to determine the extent of the Taylor Grazing act and show the erosion status of the land in 1934. 54% of my sample is classified as severely eroded and 22 % as moderately eroded. *Vegetation in NM* shows the vegetation in a small southern part of New Mexico in 1936 as digitized by Skaggs et al. (2010). Due to the limited geographical extent the numbers of observation is severely reduced and thus this variable is not part of the covariates in any other regression. *Elevation* is constructed from the Global Multi-resolution Terrain Elevation Data (GMTED2010), and shows the mean elevation in a 500m radius around every pixel. *Rainfall prior to 1935* and *Temp. prior to 1935* defines the average yearly rainfall and temperature from 1900–34. I use station level data from all stations within 100km and take the weighted average based on the distance to the pixel. Results are equivalent to using only the closest station. *Ruggedness* calculates the standard deviation of elevation of 8 adjacent cells and denominates it by the average elevation of all 9 cells. The average within a 500m radius around every pixel is reported here. *Distance to nearest river*, *Distance to Saint Louis* and *Distance to closest city* capture varying distances to proxy for water access, remoteness and thus time of settlement, and distance to modern day civilization which might affect the NDVI measure due to green lawns or highways.

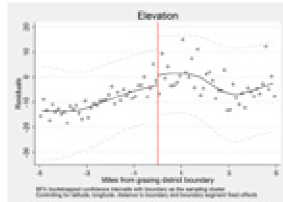
Figure 1.13: Balance graphs: Private rights treatment



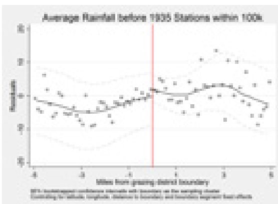
(a) Severe Erosion



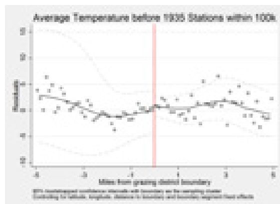
(b) Vegetation in NM



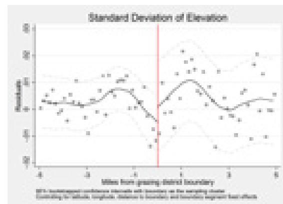
(c) Elevation



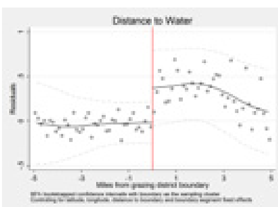
(d) Rainfall 1935



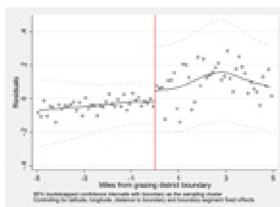
(e) Temperature 1935



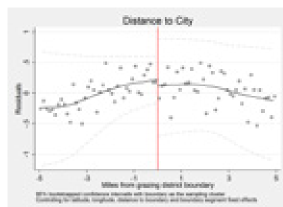
(f) Ruggedness



(g) Distance to nearest river



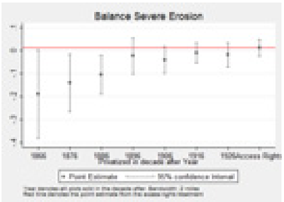
(h) Distance to Saint Louis



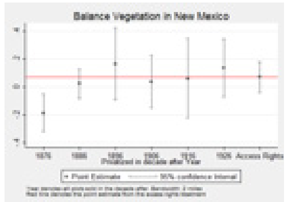
(i) Distance to closest city

Balance regression discontinuity graphs for the private rights treatment in the AVHRR data: Plotting the residuals, controlling flexibly for latitude and longitude, distance to the border as well as boundary fixed effects. Each bin is 0.125 miles wide. Balance table in Table 1.3. Variable description: *Severe Erosion* refers to erosion maps constructed for the nine states in 1934. Those maps were used to determine the extent of the Taylor Grazing act and show the erosion status of the land in 1934. 54% of my sample is classified as severely eroded and 22 % as moderately eroded. *Vegetation in NM* shows the vegetation in a small southern part of New Mexico in 1936 as digitized by Skaggs et al. (2010). Due to the limited geographical extent the numbers of observation is severely reduced and thus this variable is not part of the covariates in any other regression. *Elevation* is constructed from the Global Multi-resolution Terrain Elevation Data (GMTED2010), and shows the mean elevation in a 500m radius around every pixel. *Rainfall prior to 1935* and *Temp. prior to 1935* defines the average yearly rainfall and temperature from 1900–34. I use station level data from all stations within 100km and take the weighted average based on the distance to the pixel. Results are equivalent to using only the closest station. *Ruggedness* calculates the standard deviation of elevation of 8 adjacent cells and denominates it by the average elevation of all 9 cells. The average within a 500m radius around every pixel is reported here. *Distance to nearest river*, *Distance to Saint Louis* and *Distance to closest city* capture varying distances to proxy for water access, remoteness and thus time of settlement, and distance to modern day civilization which might affect the NDVI measure due to green lawns or highways.

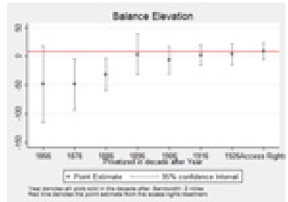
Figure 1.14: Balance graphs: Combined treatments



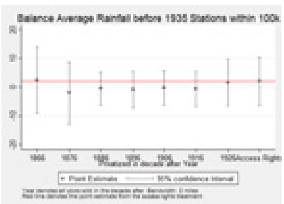
(a) Severe Erosion



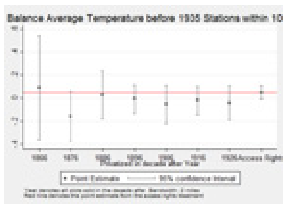
(b) Vegetation in NM



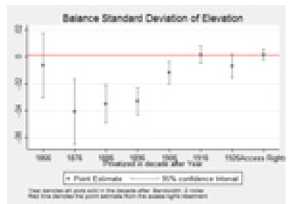
(c) Elevation



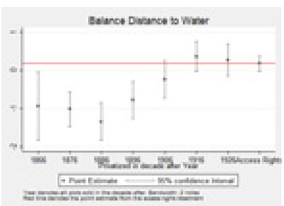
(d) Rainfall 1935



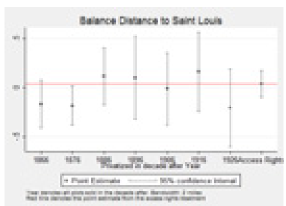
(e) Temperature 1935



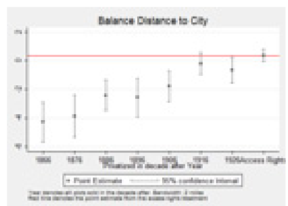
(f) Ruggedness



(g) Distance to nearest river



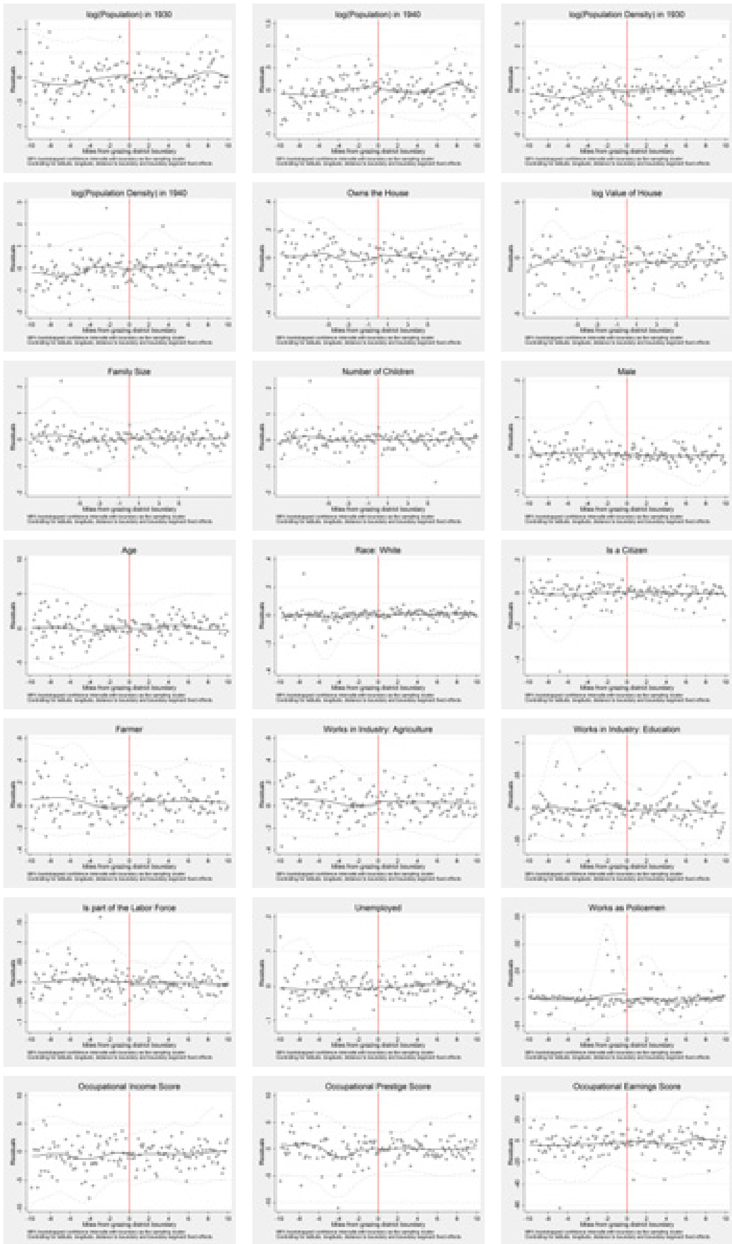
(h) Distance to Saint Louis



(i) Distance to closest city

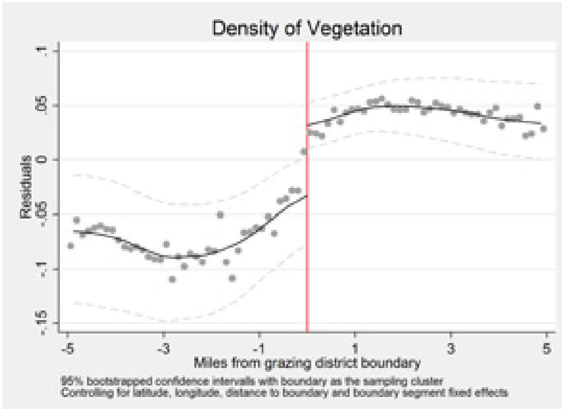
Combined estimates for both treatments in the AVHRR data including decade of purchase. Plotting the point estimate for each decade of purchase and the access rights treatment (red line) from a single regression within two miles. 95% confidence intervals reported. Variable description: *Severe Erosion* refers to erosion maps constructed for the nine states in 1934. Those maps were used to determine the extent of the Taylor Grazing act and show the erosion status of the land in 1934. 54% of my sample is classified as severely eroded and 22 % as moderately eroded. *Vegetation in NM* shows the vegetation in a small southern part of New Mexico in 1936 as digitized by Skaggs et al. (2010). Due to the limited geographical extent the numbers of observation is severely reduced and thus this variable is not part of the covariates in any other regression. *Elevation* is constructed from the Global Multi-resolution Terrain Elevation Data (GMTED2010), and shows the mean elevation in a 500m radius around every pixel. *Rainfall before 1935* and *Temperature before 1935* defines the average yearly rainfall and temperature from 1900–34. I use station level data from all stations within 100km and take the weighted average based on the distance to the pixel. Results are equivalent to using only the closest station. *Ruggedness* calculates the standard deviation of elevation of 8 adjacent cells and denominates it by the average elevation of all 9 cells. The average within a 500m radius around every pixel is reported here. *Distance to nearest river*, *Distance to Saint Louis* and *Distance to closest city* capture varying distances to proxy for water access, remoteness and thus time of settlement, and distance to modern day civilization which might affect the NDVI measure due to green lawns or highways.

Figure 1.15: Balance graphs: Population characteristics



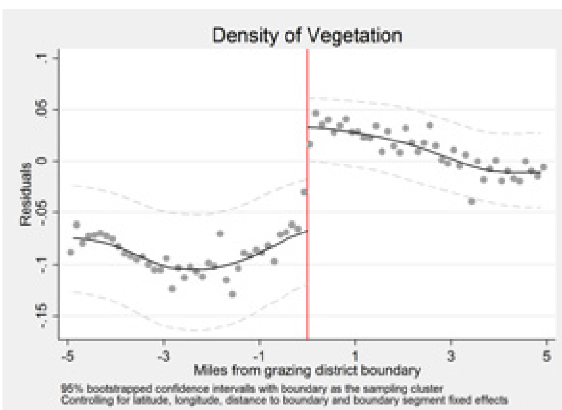
Balance regression discontinuity graphs using the minor civil divisions. Plotting the residuals, each bin is 0.125 miles wide. These RDD graphs complement Table A.2 and A.3.

Figure 1.16: Main result: Public ownership



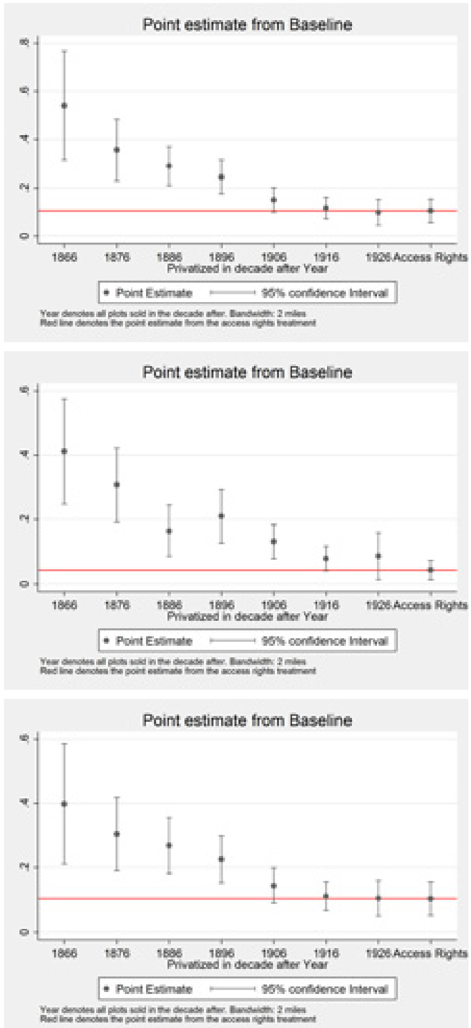
Effects of the ‘access rights’ treatment graph on the AVHRR vegetation index. It shows the residual vegetation after I control flexibly for latitude and longitude, distance to the boundary and boundary fixed effects. Standard errors are constructed by bootstrapping individual boundary segments. Each bin is 0.125 miles wide.

Figure 1.17: Main result: Private ownership



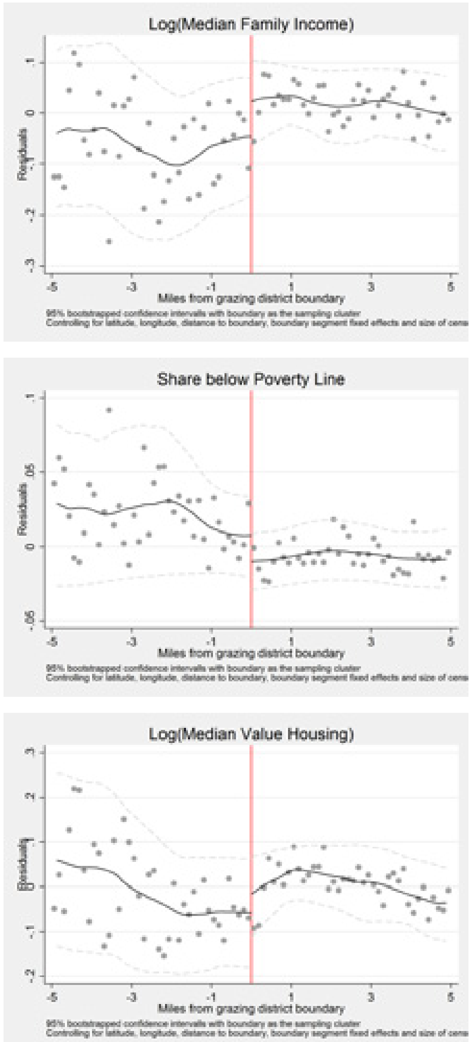
Effects of the private rights treatment graph on the AVHRR vegetation index. It shows the residual vegetation after I control flexibly for latitude and longitude, distance to the boundary and boundary fixed effects. Standard errors are constructed by bootstrapping individual boundary segments. Each bin is 0.125 miles wide.

Figure 1.18: Main result: Combined access rights and private rights



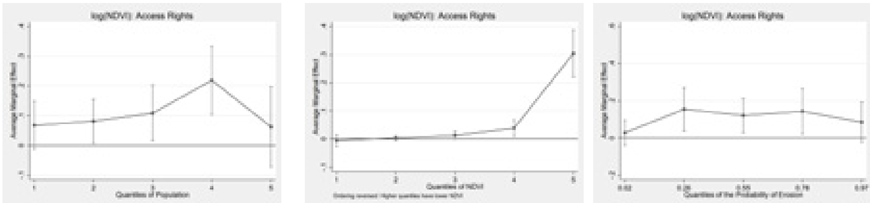
Coefficient plot on decade of purchase on NDVI. Access rights are plots still public in 1935 (red line), and the year indicates that the plot was purchased in the decade thereafter. Combined with Figure 1.14, these indicate that earlier privatized plots were of higher quality and thus not comparable to the access rights plots. Thus, these results provide empirical evidence for my identification strategy of private rights using the 1916 Stock-Grazing Homestead Act. These figures complement Table 1.5.

Figure 1.19: Main result: Income and wealth



Treatment effect on Wealth indicators. Census Blocks inside the grazing districts (right of the red line) show significant increases in income, house values, and reductions in poverty rates. RD-Graph using census-block Groups in 1990, 2000 and 2010. Residuals shown from the baseline shown, including year fixed effects. Each bin is 0.125 miles wide.

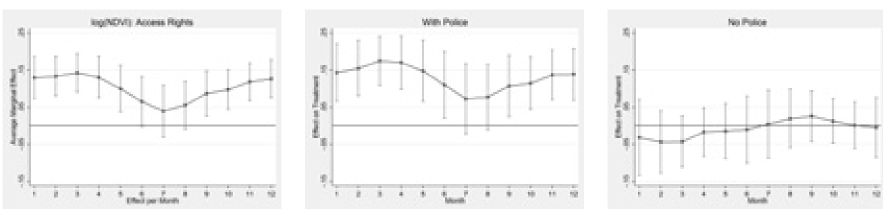
Figure 1.20: Channels: Implementation



(a) Population Quantiles **(b)** NDVI-index Quantiles **(c)** Probability of Erosion

How access rights affect vegetation. Average Marginal Effects for different quantiles. Every point estimate is calculated as the sum of the point estimates on Treatment, the interaction with the population quantile and the quantile itself. Access rights affect areas of high population pressure per minor civil division (a), have the lowest NDVI (b), or the probability of erosion (c) estimated using a linear probability model. Calculated using the `marginsplot` command in `Stata`.

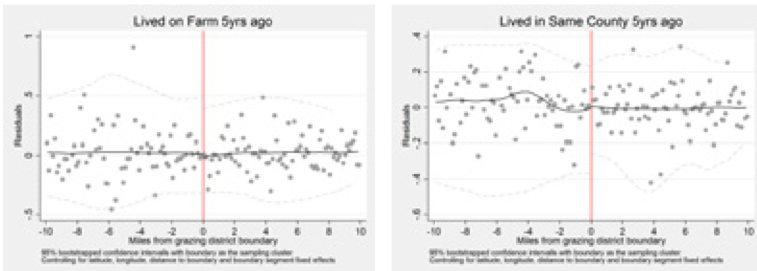
Figure 1.21: Channels: Recovery periods



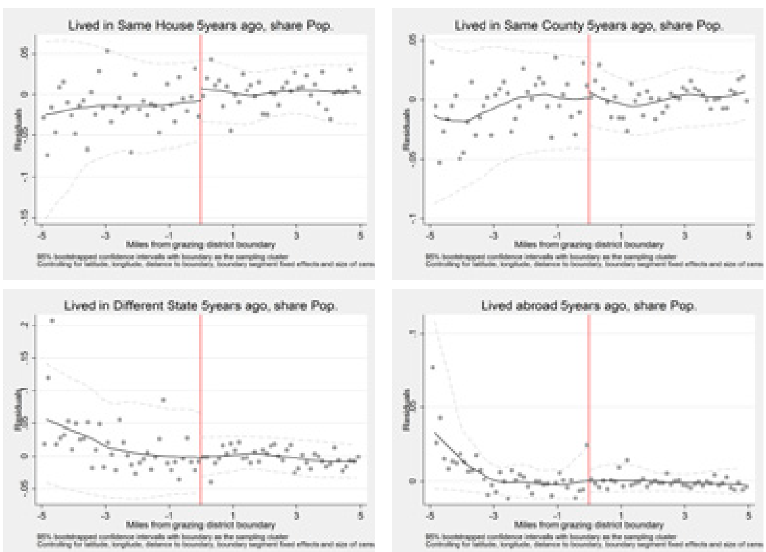
(a) Average Effect for Access rights **(b)** Access rights with Enforcement **(c)** Access rights without Enforcement

How access rights affect vegetation. Increases vegetation by extending the grazing season into the winter. Figure (b) and (c) split the sample for access rights into those counties with (b) and those without (c) police to show that enforcement is crucial for the implementation of recovery periods. Calculated using the `marginsplot` command in `Stata`.

Figure 1.22: Alternative hypothesis: Migration



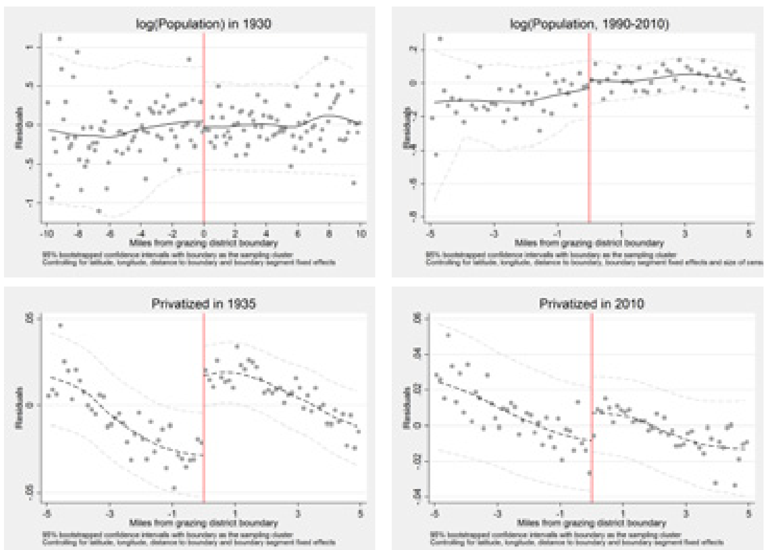
Using Minor Civil Divisions in 1940



Using census-blocks in 1990 & 2000.

Assessing potential confounding channels for the wealth effects: Differential Migration Patterns. Top row using the 1940 1% sample for the minor civil divisions. Next two rows, census blocks in 1990 and 2000, denominated by population in each census block.

Figure 1.23: Alternative hypothesis: Population and privatization



Assessing potential confounding channels for the wealth effects: Differential population growth (top panel)and differential privatization (bottom panel).

Table 1.1: Summary statistics within three miles of the boundary

	Outside Grazing Districts			Inside Grazing Districts		
	Mean	Standard Deviation	Observations	Mean	Standard Deviation	Observations
Treatment Assignment:						
Private rights	0	0	31,968	0.193	0.395	53,342
Access rights	0	0	31,968	0.807	0.395	53,342
Outcome from Satellite Data:						
Normalized Difference Vegetation Index	0.109	0.075	31,968	0.131	0.054	53,342
Covariates for Satellite Data:						
Distance to boundary	2.244	1.381	31,968	2.400	1.382	53,342
Severe erosion	0.481	0.500	31,966	0.540	0.498	53,338
Annual rain prior to 1935	67.135	229.605	31,968	55.584	206.183	53,342
Annual temperature prior to 1935	50.509	6.473	31,968	51.822	7.808	53,342
Elevation	1469.649	454.170	31,968	1417.717	499.409	53,342
Ruggedness	-0.023	0.152	31,968	-0.019	0.152	53,342
Distance to nearest river	5.320	4.856	31,968	5.539	4.954	53,342
Distance to closest city	20.638	13.471	31,968	19.082	12.586	53,342
Distance to Saint Louis	1897.538	333.830	31,968	1876.984	323.538	53,342
Vegetation in New Mexico (1936)	0.093	0.290	1,291	0.251	0.433	3,191
Outcome Data from census blocks in 1990, 2000 and 2010:						
Median family income	44915.090	24616.042	2,077	52804.549	27877.380	5,301
Share poor	0.153	0.135	2,055	0.116	0.117	5,264
Median value of house	147723.080	132826.370	2,071	167709.510	135666.090	5,262
Population statistics from Minor Civil Divisions in 1930:						
Population	619.092	1170.311	261	632.039	1222.981	384
Population Density (per sq mile)	3.154	15.722	144	4.898	26.466	161
Individual controls from Minor Civil Divisions in 1930 (Adults only):						
Male	0.513	0.500	4,143	0.513	0.500	5,609
Age	42.352	14.293	4,143	42.216	13.966	5,609
White	0.956	0.205	4,143	0.949	0.219	5,609
Citizen	0.815	0.388	4,143	0.832	0.374	5,609
Farmer	0.284	0.451	4,143	0.359	0.480	5,609
Works in agriculture	0.305	0.477	2,233	0.415	0.493	3,001
Works in education	0.023	0.149	2,233	0.023	0.151	3,001
Part of the labor force	0.540	0.498	4,143	0.536	0.499	5,608
Is unemployed	0.073	0.260	2,239	0.063	0.242	3,021
Policeman	0.004	0.067	2,244	0.002	0.041	3,035
Household controls from Minor Civil Divisions in 1930 (Household head only):						
Owns her house	0.573	0.495	2,390	0.589	0.492	3,160
Family size	3.848	2.286	2,400	4.013	2.315	3,174
Number of children	1.896	2.029	2,400	2.006	2.052	3,174
House Value	1703.530	300.440	1,341	1496.050	2804.630	1,820
Household income from Minor Civil Divisions in 1930 (Employed adults):						
Occupational income score	22.838	11.835	2,082	21.865	11.465	2,840
Occupational prestige score	36.158	13.042	2,082	36.358	13.048	2,840
Occupational earning score	44.290	56.654	2,082	49.666	106.308	2,840

Summary table for outcomes and covariates. Treatment assignment based on the data by the General Land Office and done for the Satellite data only. private rights defined as privatizations after 1916. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.2: Balance test for issuing access rights

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
				Productivity of Land					
Severe Erosion (1934)	0.004	0.012	0.022	0.002	−0.001	0.000	−0.002	0.002	0.009
	(0.013)	(0.018)	(0.021)	(0.009)	(0.009)	(0.012)	(0.012)	(0.018)	(0.020)
Vegetation in New Mexico (1936)	0.070	0.077	0.083*	0.008	0.042	0.036	0.078*	0.083*	0.077*
	(0.051)	(0.050)	(0.047)	(0.037)	(0.051)	(0.059)	(0.040)	(0.045)	(0.040)
				Inputs into Production of Vegetation					
Elevation	1.543	5.797	10.187	−1.133	0.579	2.029	2.032	5.208	9.771
	(5.325)	(7.562)	(9.047)	(3.704)	(4.224)	(5.267)	(4.651)	(7.824)	(9.841)
Average Rainfall prior to 1935	0.015	2.320	4.600	−5.063***	−3.492**	−2.524	1.323	2.708	4.930
	(2.625)	(4.360)	(6.170)	(1.831)	(1.349)	(2.095)	(2.961)	(4.270)	(4.950)
Average Temperature prior to 1935	0.040*	0.047	0.037	0.046*	0.035	0.046**	−0.001	0.009	0.014
	(0.022)	(0.029)	(0.038)	(0.027)	(0.022)	(0.022)	(0.015)	(0.021)	(0.028)
Standard Deviation of Elevation	0.004	0.000	0.001	0.009*	0.010**	0.005	0.003	−0.001	−0.000
	(0.003)	(0.002)	(0.002)	(0.005)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)
Distance to River	0.058	0.133	0.156	0.016	0.004	0.034	0.122*	0.230*	0.303*
	(0.068)	(0.102)	(0.133)	(0.065)	(0.065)	(0.074)	(0.072)	(0.125)	(0.166)
				Accessibility of Grazing Districts					
Distance to Saint Louis	−0.020	0.009	0.008	0.010	−0.006	0.009	−0.000	−0.002	−0.002
	(0.053)	(0.059)	(0.059)	(0.055)	(0.052)	(0.054)	(0.001)	(0.002)	(0.003)
Distance to Closest City	0.113	0.302	0.420	0.103	−0.023	0.063	0.021	−0.109	−0.157
	(0.164)	(0.219)	(0.281)	(0.164)	(0.164)	(0.183)	(0.125)	(0.206)	(0.274)
Observations	26,506	51,340	75,015	26,506	51,340	75,015	26,506	51,340	75,015

An observation is treated if its center is within the historical grazing districts and is public in 1935. Control observations are open-access pixels, outside the historical grazing districts without prior ownership status. Every cell is a different regression. RD-graphs in Figure 1.12. *Severe Erosion (1934)* refers to erosion maps constructed for the nine states in 1934. Those maps were used to determine the extent of the Taylor Grazing act and show the erosion status of the land in 1934. 54% of my sample is classified as severely eroded and 22 % as moderately eroded. *Vegetation in New Mexico (1936)* shows the vegetation in a small southern part of New Mexico in 1936 as digitized by Skaggs et al. (2010). Due to the limited geographical extent the numbers of observation is severely reduced and thus this variable is not part of the covariates in any other regression. *Elevation* is constructed from the Global Multi-resolution Terrain Elevation Data (GMTED2010), and shows the mean elevation in a 500m radius around every pixel. *Average Rainfall prior to 1935* defines the average yearly rainfall and *Average Temperature prior to 1935* the average temperature in Fahrenheit from 1900-34. I use station level rainfall data from all stations within 100km and takes the weighted average based on the distance to the pixel. Results are equivalent to using only the closest station. *Standard Deviation of Elevation* calculates the standard deviation of elevation of 8 adjacent cells and denominates it by the average elevation of all 9 cells. The average within a 500m radius around every pixel is reported here. *Distance to River*, *Distance to Saint Louis* and *Distance to Closest City* capture varying distances to proxy for water access, remoteness and thus time of settlement, and distance to modern day civilization which might affect the NDVI measure due to green lawns or highways. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.3: Balance test for issuing private rights

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
				Productivity of Land					
Severe Erosion (1934)	−0.009	0.000	0.014	−0.007	−0.008	−0.006	−0.019	−0.016	−0.008
	(0.018)	(0.018)	(0.020)	(0.021)	(0.022)	(0.020)	(0.014)	(0.019)	(0.022)
Vegetation in New Mexico (1936)	0.048	0.108	0.122	−0.012	0.007	−0.013	0.103	0.215^{***}	0.236^{***}
	(0.110)	(0.117)	(0.097)	(0.133)	(0.099)	(0.124)	(0.070)	(0.073)	(0.067)
				Inputs into Production of Vegetation					
Elevation	−10.203	−7.005	−8.197	−11.772	−10.530	−10.033	−4.378	0.264	1.738
	(8.353)	(8.801)	(10.396)	(8.662)	(9.345)	(8.768)	(4.994)	(6.840)	(8.258)
Average Rainfall prior to 1935	−1.878	−2.482	1.390	−2.860	−3.967	−4.236	0.881	0.540	4.372
	(2.739)	(2.733)	(2.675)	(3.329)	(2.823)	(2.693)	(0.899)	(1.375)	(2.659)
Average Temperature prior to 1935	−0.657	−0.718	0.921	−1.376	−1.555	−0.747	−0.351	−0.556	0.975
	(1.472)	(1.572)	(1.601)	(1.696)	(1.720)	(2.395)	(0.720)	(1.139)	(1.693)
Standard Deviation of Elevation	−0.006	−0.002	−0.001	0.001	0.000	0.004	−0.005	−0.005	−0.005
	(0.005)	(0.004)	(0.004)	(0.008)	(0.006)	(0.005)	(0.006)	(0.004)	(0.005)
Distance to River	0.267	0.381^{*}	0.484^{*}	0.096	0.274	0.372^{*}	0.399^{**}	0.630^{***}	0.723^{***}
	(0.204)	(0.219)	(0.252)	(0.161)	(0.199)	(0.221)	(0.199)	(0.215)	(0.232)
				Accessibility of Grazing Districts					
Distance to Saint Louis	0.149	0.226	0.275	0.147	0.195	0.169	−0.001	0.002	−0.000
	(0.249)	(0.258)	(0.278)	(0.186)	(0.222)	(0.225)	(0.002)	(0.004)	(0.006)
Distance to Closest City	−0.547	−0.249	−0.244	−0.637	−0.525	−0.367	0.233	0.004	−0.040
	(0.379)	(0.387)	(0.395)	(0.399)	(0.402)	(0.405)	(0.315)	(0.393)	(0.466)
Observations	15,482	29,512	42,257	15,482	29,512	42,257	15,482	29,512	42,257

An observation is treated if its center is within the historical grazing districts and is privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts without prior ownership status. Every cell is a different regression. RD-graphs in Figure 1.13*Severe Erosion (1934)* refers to erosion maps constructed for the nine states in 1934. Those maps were used to determine the extent of the Taylor Grazing act and show the erosion status of the land in 1934. 54% of my sample is classified as severely eroded and 22 % as moderately eroded. *Vegetation in New Mexico (1936)* shows the vegetation in a small southern part of New Mexico in 1936 as digitized by Skaggs et al. (2010). Due to the limited geographical extent the numbers of observation is severely reduced and thus this variable is not part of the covariates in any other regression. *Elevation* is constructed from the Global Multi-resolution Terrain Elevation Data (GMTED2010), and shows the mean elevation in a 500m radius around every pixel. *Average Rainfall prior to 1935* defines the average yearly rainfall and *Average Temperature prior to 1935* the average temperature in Fahrenheit from 1900-34. I use station level rainfall data from all stations within 100km and takes the weighted average based on the distance to the pixel. Results are equivalent to using only the closest station. *Standard Deviation of Elevation* calculates the standard deviation of elevation of 8 adjacent cells and denominates it by the average elevation of all 9 cells. The average within a 500m radius around every pixel is reported here. *Distance to River*, *Distance to Saint Louis* and *Distance to Closest City* capture varying distances to proxy for water access, remoteness and thus time of settlement, and distance to modern day civilization which might affect the NDVI measure due to green lawns or highways. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.4: Access rights treatment: Effect on density of vegetation

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Without Covariates	0.075^{***}	0.110^{***}	0.126^{***}	0.017	0.043^{***}	0.064^{***}	0.071^{***}	0.108^{***}	0.125^{***}
	(0.020)	(0.028)	(0.031)	(0.011)	(0.016)	(0.021)	(0.021)	(0.030)	(0.034)
With Covariates	0.076^{***}	0.109^{***}	0.122^{***}	0.019^{*}	0.043^{***}	0.063^{***}	0.071^{***}	0.107^{***}	0.121^{***}
	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.028)	(0.030)
Adj. R2 Without Covariates	0.695	0.671	0.652	0.696	0.671	0.652	0.795	0.764	0.745
Adj. R2 With Covariates	0.734	0.718	0.707	0.734	0.719	0.707	0.811	0.786	0.773
Observations	25,870	49,770	72,491	25,870	49,770	72,491	25,870	49,770	72,491
Control Mean	12.262	12.151	12.131	12.262	12.151	12.131	12.262	12.151	12.131

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is treated if its center is within the historical grazing districts and was public land in 1935. Control observations are open-access pixels, outside the historical grazing districts. Covariates are defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5: Access rights and private rights treatment: Effect on density of vegetation

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.074***	0.105***	0.118***	0.021*	0.044***	0.064***	0.072***	0.105***	0.120***
	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.026)	(0.029)
Purchased after 1926	0.101***	0.099***	0.115***	0.040	0.088**	0.098***	0.097***	0.107***	0.125***
	(0.038)	(0.027)	(0.026)	(0.027)	(0.037)	(0.033)	(0.037)	(0.028)	(0.029)
Purchased after 1916	0.088***	0.117***	0.134***	0.048***	0.081***	0.106***	0.082***	0.113***	0.135***
	(0.019)	(0.023)	(0.025)	(0.017)	(0.020)	(0.022)	(0.019)	(0.023)	(0.026)
Purchased after 1906	0.135***	0.151***	0.160***	0.099***	0.133***	0.157***	0.118***	0.146***	0.165***
	(0.026)	(0.026)	(0.026)	(0.024)	(0.027)	(0.029)	(0.027)	(0.027)	(0.028)
Purchased after 1896	0.203***	0.244***	0.268***	0.188***	0.210***	0.245***	0.165***	0.226***	0.257***
	(0.036)	(0.035)	(0.035)	(0.039)	(0.042)	(0.042)	(0.035)	(0.037)	(0.039)
Purchased after 1886	0.220***	0.290***	0.318***	0.150***	0.167***	0.229***	0.189***	0.269***	0.300***
	(0.040)	(0.040)	(0.044)	(0.035)	(0.040)	(0.039)	(0.043)	(0.044)	(0.049)
Purchased after 1876	0.339***	0.358***	0.385***	0.277***	0.308***	0.314***	0.263***	0.305***	0.346***
	(0.063)	(0.065)	(0.072)	(0.052)	(0.059)	(0.059)	(0.056)	(0.058)	(0.061)
Purchased after 1866	0.502***	0.541***	0.539***	0.385***	0.411***	0.470***	0.317***	0.400***	0.430***
	(0.105)	(0.116)	(0.119)	(0.081)	(0.084)	(0.111)	(0.087)	(0.096)	(0.090)
Adj. R2	0.728	0.715	0.701	0.729	0.716	0.702	0.804	0.781	0.767
Observations	33,096	64,317	93,934	33,096	64,317	93,934	33,096	64,317	93,934
Control Mean	12.262	12.151	12.131	12.262	12.151	12.131	12.262	12.151	12.131

Every column constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as 'Purchased after 1866' if the the land title was issued between 1866 and 1875. Control observations are open-access pixels, outside the historical grazing districts. Covariates are included in all regressions and defined in Table 1.2 and 1.3. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.6: Access rights and private rights treatment: Using purchased plots after 1916.

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.074***	0.107***	0.120***	0.018*	0.042***	0.062***	0.070***	0.105***	0.119***
	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.027)	(0.029)
Property rights	0.090***	0.112***	0.131***	0.044**	0.079***	0.102***	0.083***	0.110***	0.129***
	(0.022)	(0.022)	(0.025)	(0.018)	(0.022)	(0.024)	(0.022)	(0.023)	(0.026)
Adj. R2 With Covariates	0.736	0.722	0.710	0.736	0.723	0.710	0.812	0.789	0.776
Observations	29,217	56,661	82,775	29,217	56,661	82,775	29,217	56,661	82,775
F-Test of equality	0.249	0.652	0.398	0.094	0.016	0.008	0.377	0.673	0.336

Every column constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.7: Land classification results after a 10% increase in vegetation using machine learning algorithms

	Average NDVI	Area in 2016	Machine Learning Classifications: Support Vector Machine			Machine Learning Classifications: Random Forest		
	(1)	(2)	(3) Baseline	(4) +10%	(5) Difference	(6) Baseline	(7) +10%	(8) Difference
Western United States								
Barren	0.042	13.440	13.197	10.602	−19.67%	10.645	9.078	−14.72%
Shrub land	0.141	305.981	204.506	174.746	−14.55%	191.306	166.432	−13.00%
Grass land	0.202	101.847	131.082	131.261	0.14%	138.339	138.289	−0.04%
Deciduous forest	0.379	8.885	46.948	48.744	3.82%	33.445	33.718	0.82%
Evergreen forest	0.552	122.176	133.423	163.804	22.77%	155.422	181.640	16.87%
Sub-Saharan Africa								
Bare ground	0.126	72.551	99.800	59.011	−40.87%	50.051	31.239	−37.59%
Open shrub land	0.137	385.059	382.161	283.223	−25.89%	386.598	276.886	−28.38%
Closed shrub land	0.168	328.250	337.995	297.350	−12.03%	364.500	317.957	−12.77%
Grass land	0.207	281.483	563.762	459.573	−18.48%	277.103	169.416	−38.86%
Wooded grass land	0.223	849.709	570.039	671.341	17.77%	514.152	627.265	22.00%
Crop land	0.234	250.286	662.290	476.293	−28.08%	308.251	202.084	−34.44%
Wood land	0.259	1281.306	734.626	1007.978	37.21%	1581.189	1755.501	11.02%
Evergreen forest	0.277	543.762	542.243	554.423	2.25%	682.073	696.104	2.06%
Deciduous forest	0.285	172.149	295.95	379.676	28.29%	24.950	112.414	350.56%

Classification results from a support vector machine (Columns 1–3) and Random Forest (Columns 4–6) controlling for elevation and temperature. Cross validation rate for the support vector machine 0.942, and for the random forest 0.995. Numbers given in million Acres. NDVI is calculated within the training sample using the MODIS NDVI data for the US and the Global AVHRR data for Sub-Saharan Africa. Column (2) gives the area covered by this type of land in the respective sample. Columns (3) and (6) indicate a 10% increase in NDVI over the baseline.

Table 1.8: Wealth effect of property rights

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
log(Median Family Income)	0.131***	0.161***	0.174***	0.047	0.095***	0.108**	0.139***	0.197***	0.233***
	(0.039)	(0.044)	(0.043)	(0.037)	(0.033)	(0.046)	(0.034)	(0.032)	(0.035)
Share below Poverty Line	−0.027***	−0.035***	−0.046***	−0.027**	−0.022***	−0.024***	−0.032***	−0.047***	−0.060***
	(0.008)	(0.012)	(0.011)	(0.014)	(0.008)	(0.009)	(0.006)	(0.011)	(0.012)
log(Median Value Housing)	0.072*	0.110**	0.105**	0.007	0.059*	0.090*	0.093***	0.155***	0.187***
	(0.042)	(0.043)	(0.044)	(0.046)	(0.031)	(0.047)	(0.032)	(0.039)	(0.041)
Observations	1,928	4,325	6,658	1,928	4,325	6,658	1,928	4,325	6,658

Wealth effects using census-block groups in 1990, 2000, and 2010. Every cell is a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the grazing districts with control observations being blocks outside the grazing districts. All columns control for the size of the census-block and year fixed effects. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.9: Heterogeneous effects of property rights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Police in 1930		Bank in 1934		Newspaper in 1932	
		Without	With	Without	With	Without	With
				Vegetation Outcomes			
Access rights	0.108***	−0.007	0.136***	0.136***	0.111***	0.058**	0.252***
	(0.027)	(0.028)	(0.036)	(0.039)	(0.036)	(0.024)	(0.067)
Private rights	0.110***	0.047	0.132***	0.156***	0.096***	0.081***	0.198***
	(0.024)	(0.035)	(0.033)	(0.046)	(0.030)	(0.024)	(0.061)
				Wealth Outcomes			
log(Median Family Income)	0.161***	−0.059	0.185***	0.129	0.166***	−0.038	0.167***
	(0.044)	(0.058)	(0.042)	(0.085)	(0.045)	(0.044)	(0.039)
Share below Poverty Line	−0.035***	0.021*	−0.041***	−0.050*	−0.035***	0.004	−0.038***
	(0.012)	(0.012)	(0.012)	(0.027)	(0.012)	(0.010)	(0.012)
log(Median Value Housing)	0.110**	0.007	0.118***	0.166*	0.108**	−0.008	0.096*
	(0.043)	(0.090)	(0.045)	(0.094)	(0.045)	(0.055)	(0.057)

In the first panel 'Vegetation Outcomes' I run seven different regressions using the baseline model with satellite data and splitting the sample by the variable in the header. In the second panel 'Wealth Outcomes' every cell is a different regression using the census-blocks in 1990, 2000 and 2010. A census-block is treated if its center is within the historical grazing districts. Police is defined as zero if not person in the 1930 census is a police men in that county. Out of 321 counties, 84 counties have no policemen. Bank is defined as zero if the county had no bank in 1934. Downloaded from the Federal Deposit Insurance Corporation (2001). Out of 321 counties, 60 counties had no bank. Newspaper is defined as zero if the county had no newspaper in 1932 (Gentzkow et al., 2014). Out of 321 counties, 230 counties have no Newspaper. None of the split variables is predicted by treatment. Full results for all specification in Table A.31–A.35. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.10: Civil service reform proxy for quality of governance

	(1) log(NDVI)	(2) Log Median Family Income	(3) Share below Poverty Line	(4) Log Median Value Housing
Access rights	0.147***			
	(0.029)			
× Distance to Police Reform	−0.135***			
	(0.027)			
Private rights	0.145***			
	(0.030)			
× Distance to Police Reform	−0.100***			
	(0.025)			
Inside Grazing District		0.145***	−0.032***	0.092***
		(0.030)	(0.010)	(0.032)
× Distance to Police Reform		−0.068**	0.017*	−0.080**
		(0.032)	(0.010)	(0.031)
Observations	56,667	4,325	4,325	4,325

Civil Service Reforms (Ornaghi, 2016) as a proxy for quality of governance. In the first column the distance of every pixel to the closest city with civil service reform is calculated. Distance is standardized to give the interaction an "one standard deviation increase" interpretation. One standard deviation is 80 miles. A census-block is treated if its center is within the grazing districts with control observations being blocks outside the grazing districts. All columns control for the size of the census-block and year fixed effects. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

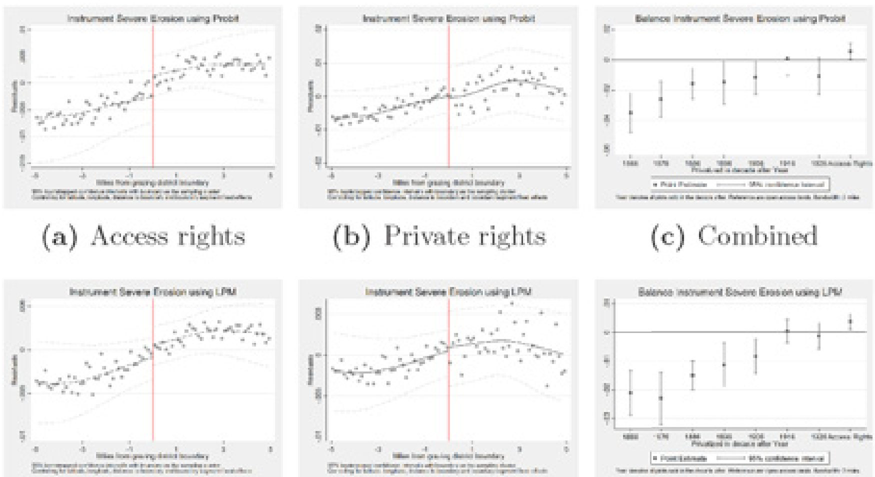
Table 1.11: Heterogeneous effect of property rights: Conditioning on police presence in 1930

	(1)	(2)	(3)	(4)	(5)
	Police in 1930	Bank in 1934		Newspaper in 1932	
	With	Without	With	Without	With
		Vegetation Outcomes			
Access rights	0.136***	0.137***	0.168***	0.073**	0.254***
	(0.036)	(0.047)	(0.054)	(0.034)	(0.068)
Private rights	0.132***	0.169**	0.142***	0.092**	0.201***
	(0.033)	(0.079)	(0.044)	(0.036)	(0.061)
		Wealth Outcomes			
Log(Median Family Income)	0.185***	0.126	0.186***	0.003	0.167***
	(0.042)	(0.094)	(0.044)	(0.053)	(0.039)
Share below Poverty Line	−0.041***	−0.051*	−0.040***	−0.011	−0.038***
	(0.012)	(0.029)	(0.013)	(0.015)	(0.012)
Log(Median Value Housing)	0.118***	0.184	0.119**	−0.000	0.096*
	(0.045)	(0.123)	(0.047)	(0.059)	(0.057)

In the first panel 'Vegetation Outcomes' I run five different regressions using the baseline model with satellite data and splitting the sample by the variable in the header. In the second panel 'Wealth Outcomes' every cell is a different regression using the census-blocks in 1990, 2000 and 2010. A census-block is treated if its center is within the historical grazing districts. Police is defined as zero if not person in the 1930 census is a police men in that county. Out of 321 counties, 84 counties have no policemen. I restrict the sample to counties with police in this table. Bank is defined as zero if the county had no bank in 1934. Downloaded from the Federal Deposit Insurance Corporation (2001). Out of 321 counties, 60 counties had no bank. Newspaper is defined as zero if the county had no newspaper in 1932 (Gentzkow et al., 2014). Out of 321 counties, 230 counties have no Newspaper. None of the split variables is predicted by treatment. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.A Appendix: Robustness

Figure A.1: Probability of erosion

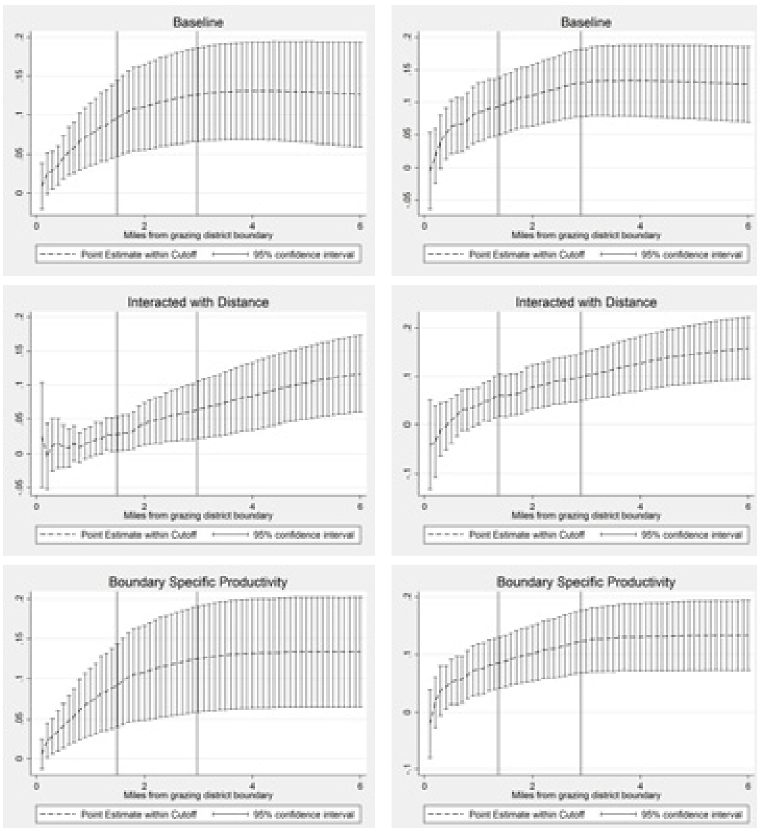


(a) Access rights (b) Private rights (c) Combined

(d) Access rights (e) Private rights (f) Combined

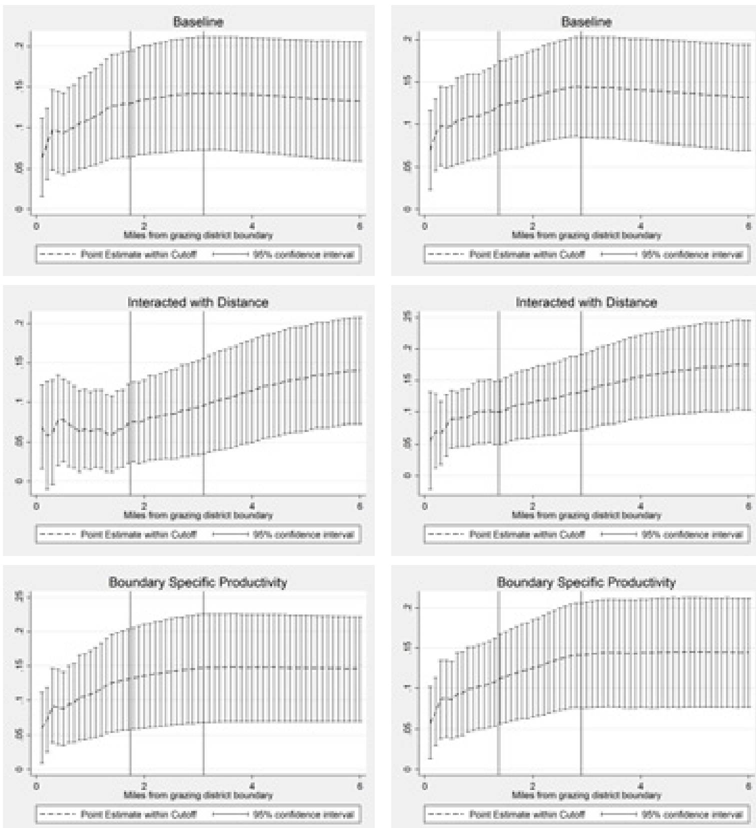
The underlying variable in this figure is the predicted probability of an observation being eroded based on all other covariates, except vegetation in New Mexico. The first row uses a Probit and the second a linear probability model.

Figure A.2: Additional bandwidths



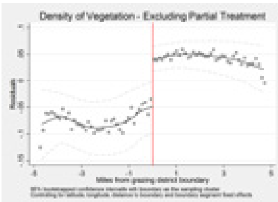
Access rights treatment (left) and private rights treatment (right): Bandwidth choices of the baseline specification (top), interacted with distance to the boundary segment (middle) and boundary specific productivity (bottom). First red line denotes the optimal bandwidth (Calonico et al., 2015), the second the bias corrected optimal bandwidth.

Figure A.3: Additional bandwidths: Excluding partially treated observations

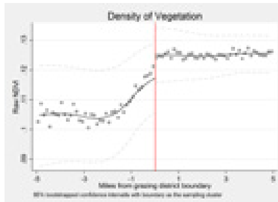


Excluding partially treated: Access rights treatment (left) and private rights treatment (right). Bandwidth choices of the baseline specification (top), interacted with distance to the boundary segment (middle) and boundary specific productivity (bottom). First red line denotes the optimal bandwidth (Calonico et al., 2015), the second the bias corrected optimal bandwidth.

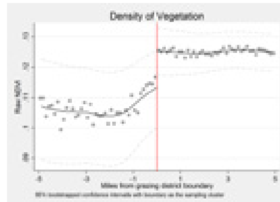
Figure A.4: Excluding partially treated



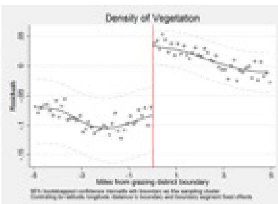
(a) Excl. Partial



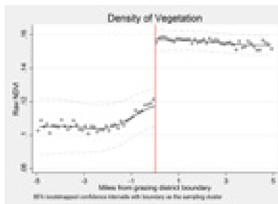
(b) NDVI



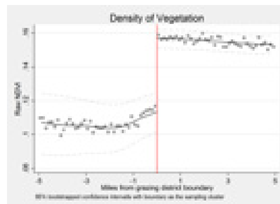
(c) NDVI, excl. Partial



(d) Excl. Partial



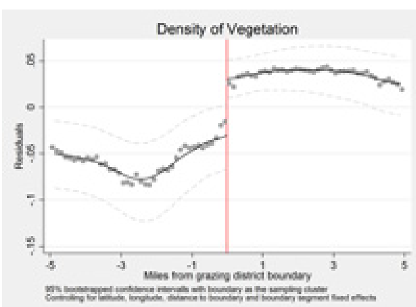
(e) NDVI



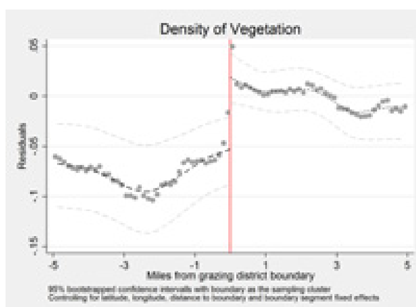
(f) NDVI, excl. Partial

Excluding partially treated and raw data for the access rights treatment (upper panel a-c) and private rights treatment (lower panel d-f). Each bin is 0.125 miles wide.

Figure A.5: MODIS: Main outcome



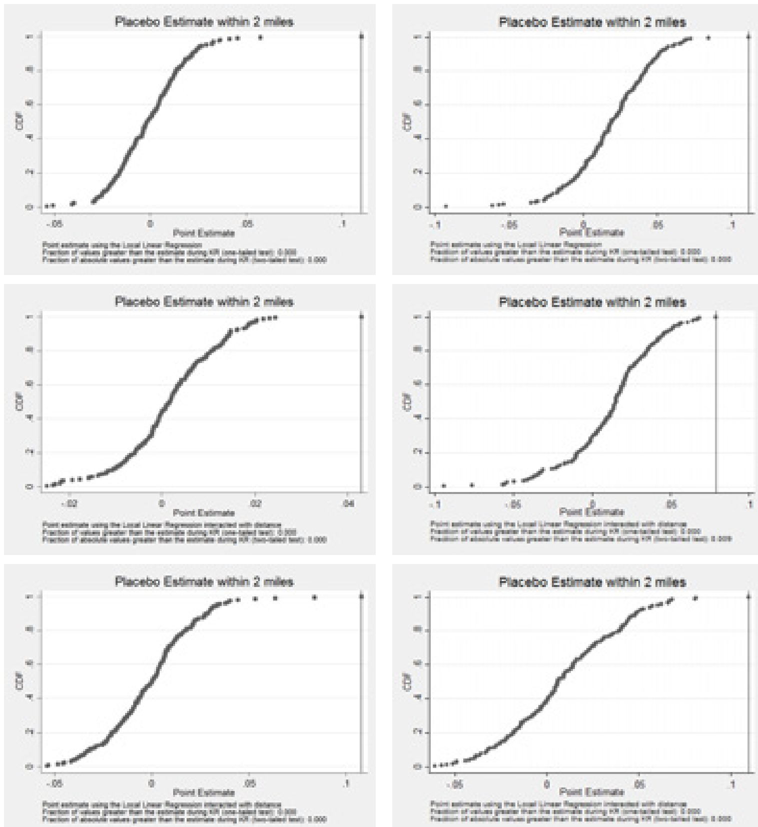
(a) Access rights



(b) Private rights

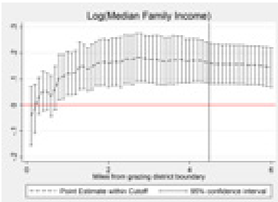
Using the MODIS satellites to derive the vegetation index. It shows the residual vegetation after I control flexibly for latitude and longitude, distance to the boundary and boundary fixed effects. Standard errors are constructed by bootstrapping individual boundary segments. Each bin is 0.125 miles wide.

Figure A.6: Placebo estimates

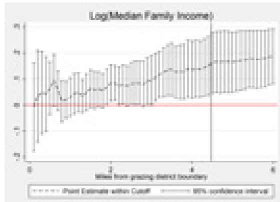


Distribution of the access rights point estimates (left) and private rights point estimates (right) after drawing 200 random grazing boundaries. The reference line marks the baseline estimate. I use all three specification. Baseline (first row), interacted with distances (second row), and interacted with lat&lon (third row).

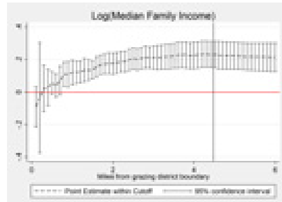
Figure A.7: Wealth: Alternative bandwidths



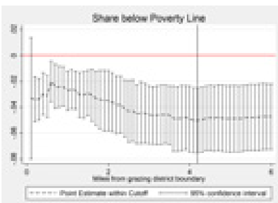
(a) Baseline specification



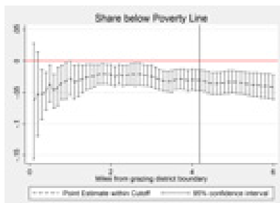
(b) Interacted with Distance



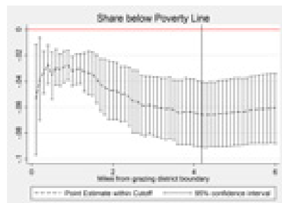
(c) Boundary specific productivity



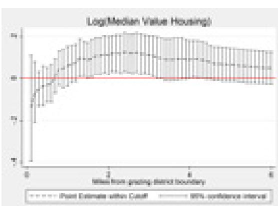
(d) Baseline specification



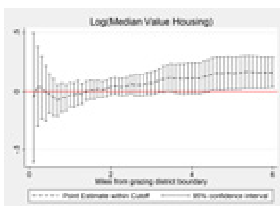
(e) Interacted with Distance



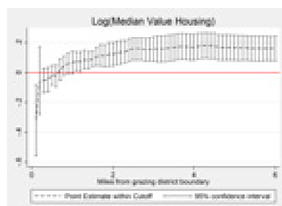
(f) Boundary specific productivity



(g) Baseline specification



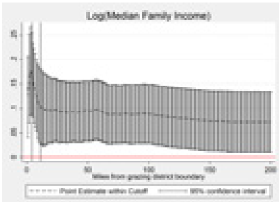
(h) Interacted with Distance



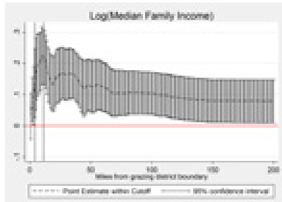
(i) Boundary specific productivity

Bandwidth choices for three variables from the census-blocks, using three specifications. The horizontal red line marks zero, the black vertical line the optimal bandwidth calculated using Calonico et al. (2015).

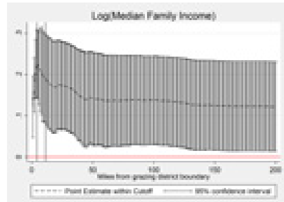
Figure A.8: Wealth: Wide bandwidths



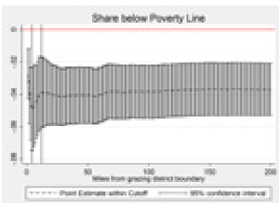
(a) Baseline specification



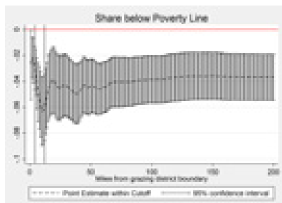
(b) Interacted with Distance



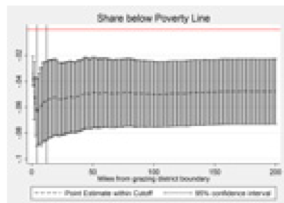
(c) Boundary specific productivity



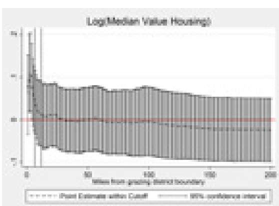
(d) Baseline specification



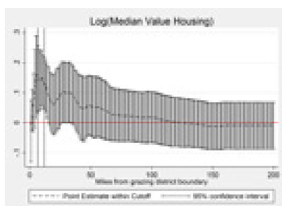
(e) Interacted with Distance



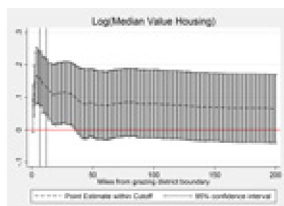
(f) Boundary specific productivity



(g) Baseline specification



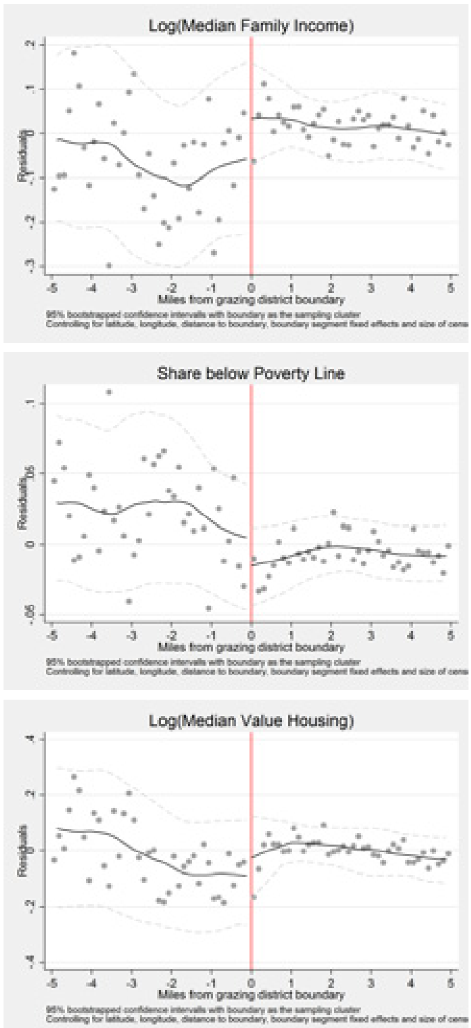
(h) Interacted with Distance



(i) Boundary specific productivity

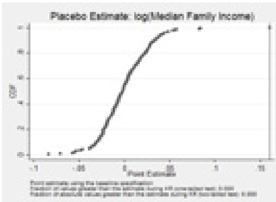
Bandwidth choices in extreme bandwidths up to 200 miles for four variables from the census-blocks, using three specifications. The horizontal red line marks zero, the black vertical line the optimal bandwidth calculated using Calonico et al. (2015).

Figure A.9: Wealth: Excluding partially treated

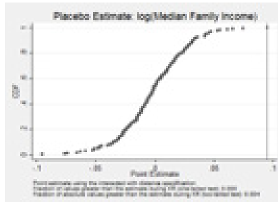


Excluding partial treatment: RD-Graph using census-block Groups in 1990, 2000 and 2010. Residuals shown, including year fixed effects.

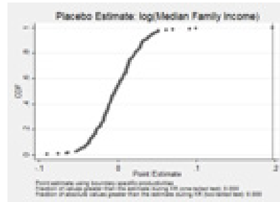
Figure A.10: Placebo Estimates



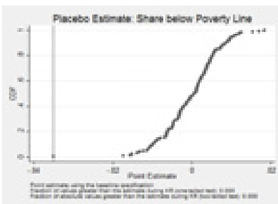
(a) Baseline specification



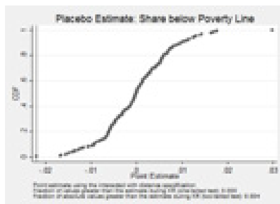
(b) Interacted with Distance



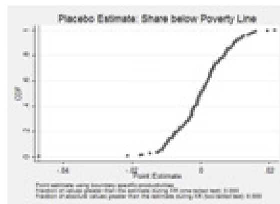
(c) Boundary specific productivity



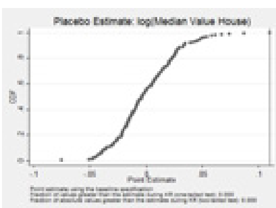
(d) Baseline specification



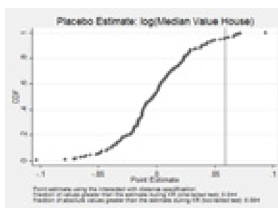
(e) Interacted with Distance



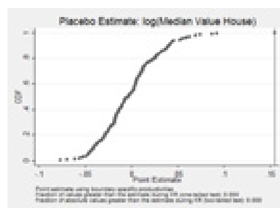
(f) Boundary specific productivity



(g) Baseline specification



(h) Interacted with Distance



(i) Boundary specific productivity

Distribution of the access rights point estimate after drawing 200 random grazing boundaries. Bandwidth: 2 miles. First panel: log(Median Family Income), second panel: Share below Poverty Line, and third panel: log(Median Value Housing).

Table A.1: Balance test for productivity of the land: Instrumenting Severe Erosion with covariates to reduce the dimensionality.

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Instrument Severe erosion using a Linear Probability Model									
Access rights	0.002	0.003**	0.004**	0.001	0.000	0.001	0.001	0.001	0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Private rights	−0.002	−0.000	−0.000	−0.003	−0.002	−0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Instrument Severe erosion using a Probit Model									
Access rights	0.002	0.006**	0.007**	0.002	0.001	0.003	0.001	0.002	0.003
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Private rights	−0.005	−0.002	0.001	−0.006	−0.006	−0.006	−0.000	−0.000	0.001
	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)
Observations	29,857	58,235	85,304	29,857	58,235	85,304	29,857	58,235	85,304

An observation is defined access rights if its center is within the historical grazing districts and is public in 1935 and private rights if its center is within the historical grazing districts and is privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts without prior ownership status. Here I instrument Severe Erosion (1934), by all covariates except vegetation in New Mexico, as defined in Table 1.2. *Severe Erosion (1934)* refers to erosion maps constructed for the nine states in 1934. I combine access rights and private rights in one regression. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Balance test on census data using Minor Civil Divisons - Part I

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 5 miles	(4) 1 mile	(5) 2 miles	(6) 5 miles	(7) 1 mile	(8) 2 miles	(9) 5 miles
				Census of Population					
log(Population) (1930)	-0.283	-0.147	-0.017	-0.682^{*}	-0.396^{*}	-0.086		-0.002	-0.112
	(0.189)	(0.130)	(0.087)	(0.389)	(0.232)	(0.132)		(0.283)	(0.149)
	234	427	1,033	234	427	1,033		427	1,033
log(Population) (1940)	-0.127	-0.173	-0.028	-0.297	-0.263	-0.060		0.123	-0.038
	(0.205)	(0.155)	(0.089)	(0.389)	(0.244)	(0.155)		(0.323)	(0.158)
	260	496	1,171	260	496	1,171		496	1,171
log(Population Density) (1930)	-0.163	0.231	0.383^{**}	-0.731	-0.209	0.285		0.407	-0.031
	(0.266)	(0.236)	(0.151)	(0.617)	(0.428)	(0.267)		(0.608)	(0.246)
	108	195	521	108	195	521		195	521
log(Population Density) (1940)	0.036	0.481^{*}	0.314^{*}	0.109	0.004	0.368		0.259	0.093
	(0.307)	(0.289)	(0.171)	(0.637)	(0.442)	(0.302)		(0.781)	(0.316)
	106	196	512	106	196	512		196	512
				Household Covariates					
Owns the House	0.026	0.058	0.058	-0.027	-0.036	0.022		0.055	0.093
	(0.049)	(0.063)	(0.037)	(0.068)	(0.089)	(0.055)		(0.064)	(0.072)
	2,057	3,906	10,826	2,057	3,906	10,826		3,906	10,826
log(House Value)	-0.615	-0.410	-1.065^{**}	-0.406	-0.893	-0.185		1.375^{*}	-0.432
	(0.589)	(0.395)	(0.495)	(1.153)	(0.929)	(0.496)		(0.706)	(0.787)
	1,134	2,230	5,901	1,134	2,230	5,901		2,230	5,901
Family Size	0.339^{*}	0.054	0.113	0.422	-0.089	-0.104		0.018	0.299^{**}
	(0.197)	(0.163)	(0.107)	(0.378)	(0.305)	(0.179)		(0.280)	(0.143)
	2,068	3,921	10,873	2,068	3,921	10,873		3,921	10,873
Number of Children	0.236	0.050	0.083	0.421	-0.037	-0.092		0.029	0.271^{*}
	(0.179)	(0.140)	(0.094)	(0.319)	(0.259)	(0.156)		(0.227)	(0.141)
	2,068	3,921	10,873	2,068	3,921	10,873		3,921	10,873

An observation is treated if the center of the minor civil division lies within the historical grazing districts in 1935. The third row of each variable indicates the number of observations for this variable in this year. Variables from the 'Census of Population' are derived from official statistic and are available for the universe of counties close to the boundary. Individual and household covariates are taken from the 5% sample of the 1930 census. *Policemen* is derived from the occupation of the individual. All individual statistics are taken from all adults, and all household statistics from the household head. *log(House Value)* is conditioning on owning a house and occupation, industry and unemployment status is conditional on being in the labor force. RD Graphs in Figure 1.15. Column (7) cannot be estimated as the degrees of freedom are zero within one mile of the boundary in that specification. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Balance test on census data using Minor Civil Divisons - Part II

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 5 miles	(4) 1 mile	(5) 2 miles	(6) 5 miles	(7) 1 mile	(8) 2 miles	(9) 5 miles
				Individual Covariates					
Male	0.002	0.004	0.002	0.002	−0.002	0.003		0.004	0.012
	(0.009)	(0.007)	(0.006)	(0.021)	(0.012)	(0.009)		(0.012)	(0.013)
	3,626	6,842	19,046	3,626	6,842	19,046		6,842	19,046
Age	-2.035^{***}	−0.676	0.714	−0.475	-2.147^{*}	−0.549		−1.078	-1.735^{*}
	(0.754)	(0.546)	(0.519)	(1.272)	(1.157)	(0.737)		(0.666)	(0.885)
	3,626	6,842	19,046	3,626	6,842	19,046		6,842	19,046
Race: White	-0.064^{*}	−0.037	0.020	0.066	-0.069^{**}	−0.038		0.001	−0.000
	(0.037)	(0.026)	(0.019)	(0.103)	(0.030)	(0.030)		(0.034)	(0.026)
	3,626	6,842	19,046	3,626	6,842	19,046		6,842	19,046
Citizen	0.019	0.035	0.002	0.001	0.042	0.030		0.087	0.082^{**}
	(0.016)	(0.026)	(0.017)	(0.040)	(0.035)	(0.026)		(0.070)	(0.033)
	3,626	6,842	19,046	3,626	6,842	19,046		6,842	19,046
Farmer	0.089^{*}	0.069	0.099^{**}	0.089	0.094	0.011		−0.078	0.029
	(0.047)	(0.046)	(0.049)	(0.085)	(0.093)	(0.054)		(0.096)	(0.083)
	3,626	6,842	19,046	3,626	6,842	19,046		6,842	19,046
Works in Industry: Agriculture	0.093^{*}	0.052	0.076	0.127	0.061	−0.038		0.070	0.076
	(0.048)	(0.043)	(0.049)	(0.114)	(0.085)	(0.056)		(0.072)	(0.077)
	1,964	3,689	10,194	1,964	3,689	10,194		3,689	10,194
Works in Industry: Education	−0.005	0.005	−0.001	−0.029	−0.015	0.001		0.014	0.012
	(0.011)	(0.011)	(0.006)	(0.019)	(0.014)	(0.010)		(0.017)	(0.012)
	1,964	3,689	10,194	1,964	3,689	10,194		3,689	10,194
Is part of the Labor Force	0.008	−0.001	−0.001	−0.033	−0.014	0.001		-0.051^{***}	−0.010
	(0.019)	(0.014)	(0.009)	(0.027)	(0.024)	(0.013)		(0.016)	(0.021)
	3,626	6,842	19,043	3,626	6,842	19,043		6,842	19,043
Is unemployed	0.033^{*}	0.010	0.003	0.049	0.028	0.021		0.037	0.024
	(0.019)	(0.017)	(0.014)	(0.046)	(0.026)	(0.017)		(0.024)	(0.025)
	1,964	3,689	10,194	1,964	3,689	10,194		3,689	10,194
Is a police men	-0.006^{**}	−0.001	−0.002	−0.006	−0.003	0.000		0.000	-0.005^{**}
	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)	(0.003)		(0.000)	(0.002)
	1,964	3,689	10,194	1,964	3,689	10,194		3,689	10,194
Occupational Income Score	-3.693^{**}	−1.295	−1.074	-6.667^{**}	−2.694	−0.001		0.031	−2.223
	(1.671)	(1.333)	(0.924)	(2.793)	(2.650)	(1.660)		(1.329)	(2.051)
	1,854	3,475	9,544	1,854	3,475	9,544		3,475	9,544
Occupational Prestige Score	-2.639^{*}	−0.494	0.515	-7.135^{***}	−2.347	−0.569		−0.422	−1.064
	(1.427)	(1.195)	(0.669)	(1.780)	(2.018)	(1.327)		(1.817)	(1.246)
	1,854	3,475	9,544	1,854	3,475	9,544		3,475	9,544
Occupational Earnings Score	-13.832^{***}	−2.090	−1.983	-16.637^{**}	−7.732	3.338		6.815	−2.832
	(5.009)	(3.984)	(2.978)	(7.509)	(7.916)	(5.139)		(4.961)	(6.452)
	1,854	3,475	9,544	1,854	3,475	9,544		3,475	9,544

An observation is treated if the center of the minor civil division lies within the historical grazing districts in 1935. The third row of each variable indicates the number of observations for this variable in this year. Variables from the 'Census of Population' are derived from official statistic and are available for the universe of counties close to the boundary. Individual and household covariates are taken from the 5% sample of the 1930 census. *Policemen* is derived from the occupation of the individual. All individual statistics are taken from all adults, and all household statistics from the household head. *log(House Value)* is conditioning on owning a house and occupation, industry and unemployment status is conditional on being in the labor force. RD Graphs in Figure 1.15. Column (7) cannot be estimated as the degrees of freedom are zero within one mile of the boundary in that specification. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Access rights treatment: Specification tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bandwidth:	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	6 miles
					Baseline					
Without Covariates	0.108***	0.108***	0.110***	0.110***	0.110***	0.110***	0.110***	0.108***	0.110***	0.125***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.034)
With Covariates	0.110***	0.109***	0.109***	0.109***	0.109***	0.109***	0.109***	0.108***	0.108***	0.120***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.029)
					Interacted with Distance					
Without Covariates	0.042***	0.042***	0.043***	0.043***	0.043***	0.043***	0.014	0.018	0.004	0.028**
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.011)	(0.016)	(0.018)	(0.013)
With Covariates	0.043***	0.043***	0.043***	0.043***	0.043***	0.043***	0.016	0.020	0.005	0.022*
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.010)	(0.015)	(0.017)	(0.012)
					Boundary specific productivity					
Without Covariates	0.108***	0.108***	0.108***	0.106***	0.103***	0.105***	0.103***	0.101***	0.099***	0.118***
	(0.028)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)	(0.032)	(0.033)
With Covariates	0.110***	0.106***	0.106***	0.105***	0.105***	0.104***	0.104***	0.101***	0.098***	0.113***
	(0.025)	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.029)	(0.030)	(0.030)
Linear Lat/Lon		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lat×Lon			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Distance				Yes	Yes	Yes	Yes	Yes	Yes	Yes
Squared Lat/Lon					Yes	Yes	Yes	Yes	Yes	Yes
Lat×Lon, squared						Yes	Yes	Yes	Yes	Yes
Squared Distance							Yes	Yes	Yes	Yes
Cubic Lat/Lon								Yes	Yes	Yes
Lat×Lon, Cubic								Yes	Yes	Yes
Cubic Distance								Yes	Yes	Yes
Quadratic Distance									Yes	Yes
Quadratic Lat/Lon									Yes	Yes
Lat×Lon, Quadratic									Yes	Yes
Observations	49,770	49,770	49,770	49,770	49,770	49,770	49,770	49,770	49,770	132,200

An observation is treated if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. Column (1) features only boundary segment fixed effects and a binary treatment indicator for access rights. Column (6) constitutes the baseline specification. In the third panel with boundary specific productivity, I additionally interact all boundary segments with all latitude and longitude polynomials and distance polynomials. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Private rights treatment: Specification tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bandwidth:	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	6 miles
					Baseline					
Access rights	0.107***	0.106***	0.107***	0.107***	0.106***	0.106***	0.106***	0.105***	0.106***	0.117***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.029)
Private rights	0.113***	0.113***	0.114***	0.114***	0.112***	0.111***	0.111***	0.110***	0.110***	0.124***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.026)
					Interacted with Distance					
Access rights	0.045***	0.044***	0.044***	0.044***	0.044***	0.044***	0.015	0.019	0.006	0.022*
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.011)	(0.015)	(0.017)	(0.012)
Private rights	0.091***	0.090***	0.091***	0.091***	0.089***	0.088***	0.054***	0.041**	0.016	0.093***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)	(0.020)	(0.024)	(0.025)
					Boundary specific productivity					
Access rights	0.107***	0.104***	0.105***	0.103***	0.097***	0.101***	0.100***	0.097***	0.095***	0.109***
	(0.025)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.028)	(0.029)
Private rights	0.113***	0.108***	0.109***	0.109***	0.103***	0.105***	0.105***	0.104***	0.102***	0.121***
	(0.022)	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.027)
Linear Lat/Lon		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lat×Lon			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Distance				Yes	Yes	Yes	Yes	Yes	Yes	Yes
Squared Lat/Lon					Yes	Yes	Yes	Yes	Yes	Yes
Lat×Lon, squared						Yes	Yes	Yes	Yes	Yes
Squared Distance							Yes	Yes	Yes	Yes
Cubic Lat/Lon								Yes	Yes	Yes
Lat×Lon, Cubic								Yes	Yes	Yes
Cubic Distance								Yes	Yes	Yes
Quadratic Distance									Yes	Yes
Quadratic Lat/Lon									Yes	Yes
Lat×Lon, Quadratic									Yes	Yes
Observations	56,661	56,661	56,661	56,661	56,661	56,661	56,661	56,661	56,661	151,442

An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. Column (1) features only boundary segment fixed effects and a binary treatment indicator for access rights. Column (6) constitutes the baseline specification. In the third panel with boundary specific productivity I additionally interact all boundary segments with all latitude and longitude polynomials and distance polynomials. Covariates included in all regressions. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Access rights treatment: Excluding partially treated

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
Without Covariates	0.098***	0.128***	0.140***	0.076***	0.087***	0.101***	0.095***	0.128***	0.143***
	(0.026)	(0.033)	(0.035)	(0.024)	(0.026)	(0.030)	(0.028)	(0.036)	(0.039)
With Covariates	0.097***	0.125***	0.134***	0.076***	0.084***	0.097***	0.096***	0.128***	0.138***
	(0.024)	(0.029)	(0.031)	(0.023)	(0.024)	(0.028)	(0.026)	(0.033)	(0.035)
Adj. R2 Without Covariates	0.695	0.667	0.648	0.695	0.667	0.648	0.787	0.759	0.743
Adj. R2 With Covariates	0.737	0.718	0.706	0.737	0.719	0.706	0.804	0.784	0.773
Observations	16,224	40,003	62,587	16,224	40,003	62,587	16,224	40,003	62,587
Control Mean	12.260	12.120	12.104	12.260	12.120	12.104	12.260	12.120	12.104

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. Distance is defined as the distance from the edge of the pixel to the grazing district boundary. If the pixel intersects the boundary the pixel is dropped. An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935. Control observations are open-access pixels, outside the historical grazing districts. Covariates are defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Private rights treatment: Excluding partially treated

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
Access rights	0.094^{***}	0.122^{***}	0.131^{***}	0.073^{***}	0.082^{***}	0.095^{***}	0.093^{***}	0.125^{***}	0.136^{***}
	(0.023)	(0.029)	(0.030)	(0.022)	(0.024)	(0.027)	(0.025)	(0.032)	(0.034)
Private rights	0.110^{***}	0.126^{***}	0.142^{***}	0.094^{***}	0.112^{***}	0.127^{***}	0.103^{***}	0.125^{***}	0.144^{***}
	(0.024)	(0.025)	(0.027)	(0.022)	(0.025)	(0.027)	(0.024)	(0.027)	(0.030)
Adj. R2	0.739	0.722	0.709	0.739	0.722	0.710	0.807	0.787	0.776
Observations	18,481	45,777	71,720	18,481	45,777	71,720	18,481	45,777	71,720
F-Test of equality	0.245	0.807	0.453	0.183	0.051	0.031	0.466	0.990	0.472

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. Distance is defined as the distance from the edge of the pixel to the grazing district boundary. If the pixel intersects the boundary the pixel is dropped. An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. Covariates are defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Access rights treatment with Conley standard errors

	Local Linear			Local Linear, Interacted			Local Linear, Boundary×Lat,Lon		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
Without Covariates	0.075	0.110	0.126	0.017	0.043	0.064	0.071	0.108	0.125
Clustered by boundary segment	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.027)	(0.029)
Spatial correlation within 0.5 degrees	(0.021)	(0.029)	(0.032)	(0.012)	(0.016)	(0.021)	(0.022)	(0.031)	(0.033)
Spatial correlation within 1 degree	(0.031)	(0.044)	(0.049)	(0.015)	(0.021)	(0.030)	(0.031)	(0.045)	(0.050)
With Covariates	0.076	0.109	0.122	0.019	0.043	0.063	0.071	0.107	0.121
Clustered by boundary segment	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.027)	(0.029)
Spatial correlation within 0.5 degrees	(0.020)	(0.027)	(0.029)	(0.012)	(0.016)	(0.021)	(0.020)	(0.028)	(0.030)
Spatial correlation within 1 degree	(0.029)	(0.040)	(0.044)	(0.015)	(0.020)	(0.029)	(0.030)	(0.042)	(0.047)
Adj. R2 Without Covariates	0.695	0.671	0.652	0.696	0.671	0.652	0.795	0.764	0.745
Adj. R2 With Covariates	0.734	0.718	0.707	0.734	0.719	0.707	0.811	0.786	0.773
Observations	25,870	49,770	72,491	25,870	49,770	72,491	25,870	49,770	72,491
Control Mean	12.262	12.151	12.131	12.262	12.151	12.131	12.262	12.151	12.131

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is treated if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. Covariates are defined in Table 1.2. Standard errors shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Private rights treatment with Conley standard errors

	Local Linear			Local Linear, Interacted			Local Linear, Boundary×Lat,Lon		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
Access rights	0.074	0.107	0.120	0.018	0.042	0.062	0.070	0.105	0.119
Clustered by boundary segment	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.027)	(0.029)
Spatial correlation within 0.5 degrees	(0.020)	(0.026)	(0.028)	(0.015)	(0.021)	(0.028)	(0.020)	(0.027)	(0.029)
Spatial correlation within 1 degree	(0.029)	(0.040)	(0.043)	(0.020)	(0.028)	(0.035)	(0.031)	(0.035)	(0.038)
Private rights	0.090	0.112	0.131	0.044	0.079	0.102	0.083	0.110	0.129
Clustered by boundary segment	(0.022)	(0.022)	(0.025)	(0.018)	(0.022)	(0.024)	(0.022)	(0.023)	(0.026)
Spatial correlation within 0.5 degree	(0.021)	(0.023)	(0.025)	(0.019)	(0.026)	(0.032)	(0.022)	(0.024)	(0.026)
Spatial correlation within 1 degree	(0.030)	(0.033)	(0.037)	(0.025)	(0.030)	(0.035)	(0.031)	(0.035)	(0.038)
Observations	29,217	56,661	82,775	29,217	56,661	82,775	29,217	56,661	82,775

Every column constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as public management if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private management if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. Covariates are included in all regressions and defined in Table 1.2. Standard errors shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Access rights treatment: using 6 miles boundary segments

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Without Covariates	0.072***	0.101***	0.112***	0.012	0.033**	0.053**	0.021*	0.027**	0.033**
	(0.022)	(0.028)	(0.030)	(0.010)	(0.015)	(0.020)	(0.010)	(0.012)	(0.014)
With Covariates	0.075***	0.105***	0.114***	0.014	0.036**	0.057***	0.021**	0.029**	0.035**
	(0.021)	(0.026)	(0.027)	(0.010)	(0.015)	(0.020)	(0.011)	(0.012)	(0.014)
Adj. R2 Without Covariates	0.845	0.804	0.779	0.846	0.805	0.780	0.934	0.920	0.907
Adj. R2 With Covariates	0.851	0.817	0.798	0.852	0.818	0.799	0.935	0.920	0.908
Observations	25,870	49,770	72,491	25,870	49,770	72,491	25,870	49,770	72,491
Control Mean	12.262	12.151	12.131	12.262	12.151	12.131	12.262	12.151	12.131

An observation is treated if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. Estimations have fixed effects for each 6 miles boundary segments. Standard errors clustered by 60 miles boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Private rights treatment: using 6 miles boundary segments

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.073***	0.103***	0.112***	0.015	0.038**	0.058***	0.021**	0.026**	0.035**
	(0.020)	(0.026)	(0.027)	(0.010)	(0.015)	(0.019)	(0.010)	(0.011)	(0.014)
Private rights	0.077***	0.098***	0.114***	0.048***	0.077***	0.092***	0.024***	0.033***	0.040***
	(0.022)	(0.021)	(0.023)	(0.016)	(0.018)	(0.019)	(0.009)	(0.011)	(0.012)
Adj. R2	0.852	0.820	0.801	0.852	0.821	0.802	0.939	0.925	0.912
Observations	29,217	56,661	82,775	29,217	56,661	82,775	29,217	56,661	82,775
F-Test of equality	0.769	0.653	0.910	0.027	0.002	0.007	0.584	0.167	0.274

An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. Estimations have fixed effects for each 6 miles boundary segments. Standard errors clustered by 60 miles boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Access rights treatment: Dropping 60 miles boundary segments

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Without Covariates	0.100***	0.143***	0.166***	0.032**	0.056**	0.077***	0.093***	0.138***	0.161***
	(0.028)	(0.037)	(0.040)	(0.015)	(0.021)	(0.029)	(0.029)	(0.039)	(0.043)
With Covariates	0.097***	0.134***	0.153***	0.037**	0.057***	0.077***	0.091***	0.132***	0.150***
	(0.026)	(0.033)	(0.036)	(0.015)	(0.021)	(0.028)	(0.027)	(0.036)	(0.039)
Adj. R2 Without Covariates	0.733	0.717	0.706	0.733	0.718	0.707	0.813	0.783	0.768
Adj. R2 With Covariates	0.736	0.719	0.709	0.736	0.720	0.710	0.813	0.783	0.768
Observations	17,042	32,768	47,660	17,042	32,768	47,660	17,042	32,768	47,660
Control Mean	11.356	11.286	11.270	11.356	11.286	11.270	11.356	11.286	11.270

An observation is treated if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. In the baseline I drop 6 miles border segments if any observation associated with that segment is a national forest. Here I drop an entire 60 miles border segment if any observation is a national forest. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: Private rights treatment: Dropping 60 miles boundary segments

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.097***	0.133***	0.152***	0.038**	0.058***	0.076***	0.091***	0.131***	0.149***
	(0.026)	(0.033)	(0.035)	(0.015)	(0.021)	(0.026)	(0.027)	(0.035)	(0.037)
Private rights	0.122***	0.142***	0.161***	0.077***	0.115***	0.134***	0.109***	0.136***	0.161***
	(0.032)	(0.030)	(0.033)	(0.024)	(0.028)	(0.028)	(0.032)	(0.030)	(0.034)
Adj. R2	0.736	0.722	0.711	0.736	0.723	0.712	0.814	0.787	0.772
Observations	19,028	36,889	53,823	19,028	36,889	53,823	19,028	36,889	53,823
F-Test of equality	0.227	0.602	0.559	0.058	0.003	0.003	0.403	0.764	0.414

An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. In the baseline I drop 6 miles border segments if any observation associated with that segment is a national forest. Here I drop an entire 60 miles border segment if any observation is a national forest. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.14: Access rights treatment: Not dropping any boundaries

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Without Covariates	0.047***	0.063***	0.070***	0.034***	0.057***	0.073***	0.038**	0.063***	0.075***
	(0.016)	(0.022)	(0.025)	(0.009)	(0.014)	(0.018)	(0.016)	(0.023)	(0.026)
With Covariates	0.051***	0.074***	0.085***	0.019**	0.041***	0.057***	0.044***	0.073***	0.086***
	(0.015)	(0.020)	(0.022)	(0.008)	(0.012)	(0.016)	(0.015)	(0.021)	(0.024)
Adj. R2 Without Covariates	0.710	0.683	0.660	0.710	0.683	0.660	0.814	0.784	0.763
Adj. R2 With Covariates	0.754	0.736	0.720	0.755	0.736	0.720	0.829	0.806	0.790
Observations	46,573	88,728	128,370	46,573	88,728	128,370	46,573	88,728	128,370
Control Mean	13.982	13.764	13.705	13.982	13.764	13.705	13.982	13.764	13.705

An observation is treated if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. In the baseline I drop 6 miles border segments if any observation associated with that segment is a national forest. In this table I keep all observations regardless. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.15: Private rights treatment: Not dropping any boundaries

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.050***	0.072***	0.082***	0.030***	0.058***	0.074***	0.043***	0.070***	0.082***
	(0.015)	(0.020)	(0.022)	(0.009)	(0.012)	(0.015)	(0.015)	(0.021)	(0.023)
Private rights	0.064***	0.084***	0.098***	0.066***	0.103***	0.118***	0.054***	0.076***	0.089***
	(0.016)	(0.019)	(0.021)	(0.013)	(0.015)	(0.017)	(0.015)	(0.019)	(0.021)
Adj. R2	0.753	0.736	0.719	0.753	0.735	0.719	0.831	0.809	0.793
Observations	57,308	109,613	158,150	57,308	109,613	158,150	57,308	109,613	158,150
F-Test of equality	0.056	0.090	0.037	0.000	0.000	0.000	0.106	0.270	0.214

An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. In the baseline I drop 6 miles border segments if any observation associated with that segment is a national forest. In this table I keep all observations regardless. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.16: Access rights and private rights treatment: Standardize the index instead of using logs.

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.069***	0.112***	0.136***	0.006	0.022	0.051**	0.061**	0.099**	0.114**
	(0.026)	(0.039)	(0.046)	(0.012)	(0.018)	(0.023)	(0.027)	(0.040)	(0.045)
Private rights	0.090***	0.132***	0.174***	0.055*	0.115***	0.148***	0.083**	0.123***	0.149***
	(0.034)	(0.039)	(0.045)	(0.030)	(0.032)	(0.035)	(0.034)	(0.039)	(0.043)
Adj. R2 With Covariates	0.736	0.722	0.710	0.736	0.723	0.710	0.812	0.789	0.776
Observations	29,217	56,661	82,775	29,217	56,661	82,775	29,217	56,661	82,775
F-Test of equality	0.459	0.458	0.179	0.094	0.001	0.001	0.423	0.296	0.114

Every column constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.17: Access rights and private rights treatment: Dependent variable: NDVI×100

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.437***	0.715***	0.866***	0.038	0.141	0.321**	0.388**	0.634**	0.725**
	(0.163)	(0.250)	(0.293)	(0.078)	(0.114)	(0.148)	(0.174)	(0.257)	(0.286)
Private rights	0.574***	0.841***	1.108***	0.348*	0.735***	0.942***	0.525**	0.786***	0.945***
	(0.216)	(0.247)	(0.289)	(0.191)	(0.205)	(0.221)	(0.213)	(0.248)	(0.276)
Adj. R2 With Covariates	0.736	0.722	0.710	0.736	0.723	0.710	0.812	0.789	0.776
Observations	29,217	56,661	82,775	29,217	56,661	82,775	29,217	56,661	82,775
F-Test of equality	0.459	0.458	0.179	0.094	0.001	0.001	0.423	0.296	0.114

Every column constitute a separate regression of the treatment indicator on the standardized index of NDVI using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.18: Access rights and private rights treatment: Using MODIS 250m data

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Access rights	0.065***	0.081***	0.094***	0.036***	0.049***	0.055***	0.065***	0.086***	0.101***
	(0.020)	(0.025)	(0.028)	(0.013)	(0.017)	(0.020)	(0.021)	(0.027)	(0.030)
Private rights	0.054***	0.069***	0.083***	0.035**	0.051**	0.058**	0.048**	0.070***	0.090***
	(0.020)	(0.021)	(0.023)	(0.017)	(0.020)	(0.022)	(0.020)	(0.023)	(0.026)
Adj. R2	0.622	0.604	0.588	0.622	0.604	0.588	0.709	0.690	0.673
Observations	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353
F-Test of equality	0.414	0.319	0.379	0.943	0.879	0.828	0.188	0.164	0.332

Every column constitute a separate regression of the treatment indicator on NDVI×100 using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.19: Access rights and private rights treatment: Using MODIS 250m data and standardize the index and levels.

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
					Standardized NDVI				
Access rights	0.083***	0.097**	0.108***	0.046***	0.058**	0.063**	0.080**	0.096**	0.104**
	(0.031)	(0.038)	(0.040)	(0.017)	(0.024)	(0.030)	(0.032)	(0.040)	(0.042)
Private rights	0.065*	0.083**	0.107**	0.034	0.050	0.049	0.057	0.085**	0.111**
	(0.038)	(0.040)	(0.042)	(0.031)	(0.034)	(0.037)	(0.038)	(0.041)	(0.043)
Observations	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353
F-Test of equality	0.589	0.635	0.960	0.714	0.792	0.660	0.434	0.695	0.773
					NDVI×100				
Access rights	0.779***	0.911**	1.021***	0.435***	0.551**	0.595**	0.750**	0.903**	0.977**
	(0.289)	(0.355)	(0.379)	(0.163)	(0.226)	(0.285)	(0.302)	(0.375)	(0.398)
Property rights	0.617*	0.781**	1.008**	0.324	0.472	0.465	0.534	0.805**	1.044**
	(0.362)	(0.379)	(0.397)	(0.291)	(0.320)	(0.346)	(0.355)	(0.384)	(0.405)
Observations	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353
F-Test of equality	0.589	0.635	0.960	0.714	0.792	0.660	0.434	0.695	0.773

Every column constitute a separate regression of the treatment indicator on standardized index of NDVI or NDVI×100 using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.20: Access rights and private rights treatment: Using MODIS 250m data, robustness checks

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Using 6 miles border segments									
Access rights	0.061***	0.077***	0.087***	0.031**	0.040**	0.046**	0.041***	0.040**	0.037**
	(0.019)	(0.022)	(0.024)	(0.012)	(0.016)	(0.019)	(0.016)	(0.016)	(0.016)
Private rights	0.050***	0.063***	0.075***	0.030**	0.044***	0.048**	0.039***	0.042***	0.040***
	(0.017)	(0.017)	(0.019)	(0.013)	(0.017)	(0.018)	(0.014)	(0.015)	(0.014)
Observations	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353	623,437	1,210,956	1,771,353
F-Test of equality	0.270	0.168	0.214	0.907	0.714	0.886	0.791	0.728	0.623
Dropping 60 Miles Boundary Segments									
Access rights	0.057***	0.073***	0.085***	0.026*	0.040**	0.044**	0.059***	0.077***	0.091***
	(0.021)	(0.024)	(0.026)	(0.013)	(0.019)	(0.022)	(0.022)	(0.026)	(0.029)
Private rights	0.064***	0.072***	0.087***	0.043**	0.063***	0.065**	0.054**	0.073***	0.093***
	(0.024)	(0.023)	(0.025)	(0.021)	(0.024)	(0.025)	(0.023)	(0.023)	(0.026)
Observations	422,277	821,250	1,119,720	422,277	821,250	1,119,720	422,277	821,250	1,119,720
F-Test of equality	0.663	0.925	0.936	0.336	0.182	0.207	0.776	0.723	0.850
Keep entire sample									
Access rights	0.043***	0.054***	0.062***	0.042***	0.049***	0.052***	0.040***	0.054***	0.065***
	(0.015)	(0.018)	(0.020)	(0.009)	(0.013)	(0.015)	(0.015)	(0.020)	(0.022)
Private rights	0.040***	0.054***	0.066***	0.047***	0.056***	0.059***	0.034**	0.051***	0.064***
	(0.014)	(0.017)	(0.019)	(0.012)	(0.014)	(0.016)	(0.014)	(0.017)	(0.020)
Observations	1,200,270	2,294,028	3,309,410	1,200,270	2,294,028	3,309,410	1,200,270	2,294,028	3,309,410
F-Test of equality	0.687	0.978	0.601	0.537	0.387	0.411	0.362	0.535	0.868

Every column constitute a separate regression of the treatment indicator on log(NDVI) using the model in the corresponding header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. An observation is defined as private property rights if its center is within the historical grazing districts and has been privatized after 1916. Control observations are collectively managed pixels, outside the historical grazing districts. The last row tests for equality of coefficients using an F-Test. Covariates are included in all regressions and defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.21: Private rights treatment: Using private rights on either side of the boundary

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
AVHRR data:									
Access rights	0.074***	0.106***	0.118***	0.018	0.041***	0.061***	0.068***	0.101***	0.111***
	(0.019)	(0.025)	(0.027)	(0.011)	(0.015)	(0.020)	(0.020)	(0.026)	(0.028)
Private rights	0.078***	0.094***	0.105***	0.040**	0.070***	0.093***	0.072***	0.089***	0.095***
	(0.019)	(0.020)	(0.021)	(0.016)	(0.020)	(0.022)	(0.018)	(0.020)	(0.021)
Observations	31,221	60,514	88,432	31,221	60,514	88,432	31,221	60,514	88,432
F-Test of equality	0.749	0.363	0.338	0.100	0.036	0.030	0.783	0.329	0.187
MODIS data:									
Access rights	0.063***	0.078***	0.088***	0.036***	0.047***	0.052***	0.064***	0.082***	0.092***
	(0.019)	(0.022)	(0.024)	(0.012)	(0.017)	(0.019)	(0.020)	(0.025)	(0.027)
Private rights	0.062***	0.072***	0.084***	0.043***	0.058***	0.061***	0.057***	0.074***	0.091***
	(0.019)	(0.020)	(0.021)	(0.016)	(0.019)	(0.021)	(0.019)	(0.021)	(0.023)
Observations	623,369	1,210,853	1,771,230	623,369	1,210,853	1,771,230	623,369	1,210,853	1,771,230
F-Test of equality	0.903	0.634	0.738	0.606	0.445	0.512	0.523	0.493	0.893

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model and data in the header. Distance is defined as the distance from the edge of the pixel to the grazing district boundary. If the pixel intersects the boundary the pixel is dropped. An observation is treated if its center is within the historical grazing districts. It is defined as access rights if it was public lands in 1935, and private rights if it has been privatized after 1916. Control observations are open-access pixels, outside the historical grazing districts. Covariates are defined in Table 1.2. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.22: Access rights treatment: Using Landsat 30m data

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
Within 60 miles segments	0.038*	0.043*	0.043	0.026	0.031*	0.038**	0.046**	0.050**	0.049*
	(0.020)	(0.024)	(0.027)	(0.016)	(0.017)	(0.019)	(0.019)	(0.024)	(0.027)
Adj. R2	0.433	0.426	0.420	0.433	0.426	0.420	0.488	0.478	0.469
Within 6 miles segments	0.040**	0.043*	0.044	0.026*	0.029**	0.033*	0.030*	0.029**	0.028*
	(0.019)	(0.024)	(0.027)	(0.013)	(0.014)	(0.017)	(0.016)	(0.014)	(0.014)
Adj. R2	0.613	0.594	0.578	0.613	0.594	0.578	0.676	0.663	0.650
Standardized NDVI value	0.102*	0.133*	0.147*	0.033	0.053	0.075*	0.109**	0.128**	0.136**
	(0.053)	(0.069)	(0.078)	(0.025)	(0.033)	(0.043)	(0.044)	(0.059)	(0.066)
Adj. R2	0.385	0.393	0.396	0.385	0.394	0.397	0.480	0.486	0.488
Observations	37,693,841	73,713,788	107,250,573	37,693,841	73,713,788	107,250,573	37,693,841	73,713,788	107,250,573

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts and has not been privatized in 1935. Control observations are collectively managed pixels, outside the historical grazing districts. In the first column I compare within 60 miles boundary segments and in the second column within 6 mile segments. The last column compares standardized NDVI values within 60 miles boundary segments to compare across sensors. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.23: Access rights treatment: US Department of Agriculture cropland type

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Pr[Type=rangeland]	0.080***	0.120***	0.143***	0.029***	0.046***	0.067***	0.080***	0.116***	0.138***
	(0.015)	(0.019)	(0.020)	(0.011)	(0.013)	(0.016)	(0.016)	(0.020)	(0.023)
Observations	22,703,293	44,526,483	66,004,911	22,703,293	44,526,483	66,004,911	22,703,293	44,526,483	66,004,911
Control Mean	0.643	0.635	0.634	0.643	0.635	0.634	0.643	0.635	0.634
Pr[Type=rangeland|Type $\in$ {rangeland;barren}]	0.056***	0.082***	0.096***	0.022***	0.035***	0.051***	0.054***	0.079***	0.096***
	(0.013)	(0.016)	(0.019)	(0.009)	(0.010)	(0.013)	(0.013)	(0.017)	(0.020)
Observations	18,762,732	37,086,222	55,288,555	18,762,732	37,086,222	55,288,555	18,762,732	37,086,222	55,288,555
Control Mean	0.805	0.798	0.800	0.805	0.798	0.800	0.805	0.798	0.800
Pr[Type=grass|Type $\in$ {grass;shrub}]	0.006	0.009*	0.007	0.000	0.004	0.008*	0.012***	0.017***	0.017***
	(0.004)	(0.005)	(0.006)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
Observations	15,882,187	31,728,891	47,934,780	15,882,187	31,728,891	47,934,780	15,882,187	31,728,891	47,934,780
Control Mean	0.229	0.234	0.238	0.229	0.234	0.238	0.229	0.234	0.238

Every cell constitute a separate regression of the treatment indicator on log(NDVI) using the model in the header. An observation is defined as access rights if its center is within the historical grazing districts. Control observations are collectively managed pixels, outside the historical grazing districts. Every row is a different dependent variable from the NASS Cropland classification. The first row estimates the overall likelihood of observing rangeland, defined as shrub or grass land. The second row compares the probability of rangeland compared to low quality barren land. The third row compares the likelihood of observing higher quality grass land, compared to shrub land. Unconditional probability of observing grass land: 0.713, shrub land: 0.601, grass land: 0.112 and barren lands: 0.068 Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.24: Wealth effect of property rights: Year-by-year effects

	Baseline			Interacted with Distance			Boundary specific productivity		
Year	(1) 1990	(2) 2000	(3) 2010	(4) 1990	(5) 2000	(6) 2010	(7) 1990	(8) 2000	(9) 2010
Log(Median Family Income)	0.175***	0.169***	0.171***	0.114**	0.127***	0.073	0.224***	0.229***	0.193***
	(0.059)	(0.041)	(0.043)	(0.049)	(0.038)	(0.046)	(0.047)	(0.034)	(0.032)
Share below Poverty Line	−0.038**	−0.038***	−0.033***	−0.024*	−0.019*	−0.024***	−0.048**	−0.057***	−0.048***
	(0.019)	(0.011)	(0.010)	(0.013)	(0.011)	(0.009)	(0.020)	(0.010)	(0.010)
High school graduate	0.052**	0.046**	0.042*	0.019	0.028*	0.027**	0.063***	0.070***	0.061***
	(0.021)	(0.021)	(0.022)	(0.015)	(0.014)	(0.013)	(0.017)	(0.018)	(0.019)
Log(Median Value Housing)	0.170***	0.096**	0.100*	0.064	0.094**	0.046	0.193***	0.164***	0.155***
	(0.045)	(0.040)	(0.052)	(0.050)	(0.041)	(0.042)	(0.056)	(0.036)	(0.058)
Observations	1,125	1,510	1,757	1,125	1,510	1,757	1,125	1,510	1,757

Wealth effects using census-block Groups in individual years. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.25: Wealth effect: Specification tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bandwidth:	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles	2 miles
					Baseline					
log(Median Family Income)	0.121**	0.149***	0.156***	0.153***	0.158***	0.161***	0.159***	0.158***	0.156***	0.154***
	(0.051)	(0.048)	(0.047)	(0.047)	(0.048)	(0.046)	(0.046)	(0.046)	(0.046)	(0.047)
Share below Poverty Line	-0.027**	-0.032***	-0.034***	-0.033***	-0.035***	-0.035***	-0.035***	-0.035***	-0.035***	-0.035***
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
log(Median Value Housing)	0.075	0.101**	0.109**	0.106**	0.107**	0.110**	0.106**	0.104**	0.101**	0.099**
	(0.050)	(0.049)	(0.050)	(0.050)	(0.050)	(0.046)	(0.046)	(0.046)	(0.045)	(0.045)
					Interacted with Distance					
log(Median Family Income)	0.061	0.084**	0.093***	0.084**	0.084**	0.085**	0.080**	-0.008	0.035	-0.039
	(0.038)	(0.038)	(0.036)	(0.035)	(0.035)	(0.036)	(0.033)	(0.042)	(0.055)	(0.074)
Share below Poverty Line	-0.016*	-0.021**	-0.023***	-0.021**	-0.021**	-0.022***	-0.022**	-0.028*	-0.037	-0.023
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.015)	(0.023)	(0.017)
log(Median Value Housing)	0.016	0.039	0.050	0.038	0.038	0.041	0.030	-0.065	-0.082	-0.196*
	(0.037)	(0.039)	(0.038)	(0.035)	(0.035)	(0.035)	(0.038)	(0.064)	(0.092)	(0.117)
					Boundary specific productivity					
log(Median Family Income)	0.121**	0.149***	0.184***	0.171***	0.182***	0.180***	0.162***	0.185***	0.178***	0.172***
	(0.051)	(0.048)	(0.034)	(0.031)	(0.036)	(0.041)	(0.036)	(0.045)	(0.052)	(0.063)
Share below Poverty Line	-0.027**	-0.032***	-0.047***	-0.046***	-0.049***	-0.049***	-0.046***	-0.053***	-0.053***	-0.051***
	(0.012)	(0.012)	(0.010)	(0.011)	(0.012)	(0.013)	(0.012)	(0.014)	(0.015)	(0.018)
log(Median Value Housing)	0.075	0.101**	0.136***	0.122***	0.121***	0.117***	0.108**	0.125**	0.107*	0.108
	(0.050)	(0.049)	(0.039)	(0.042)	(0.041)	(0.044)	(0.046)	(0.051)	(0.061)	(0.074)
log(Area)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Lat/Lon			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lat×Lon				Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Distance					Yes	Yes	Yes	Yes	Yes	Yes
Squared Lat/Lon						Yes	Yes	Yes	Yes	Yes
Lat×Lon, squared							Yes	Yes	Yes	Yes
Squared Distance								Yes	Yes	Yes
Cubic Lat/Lon									Yes	Yes
Lat×Lon, Cubic									Yes	Yes
Cubic Distance									Yes	Yes
Quadratic Distance										Yes
Quadratic Lat/Lon										Yes
Lat×Lon, Quadratic										Yes
Observations	4,325	4,325	4,325	4,325	4,325	4,325	4,325	4,325	4,325	4,325

An observation is treated if its center is within the historical grazing districts. Column (1) features only boundary segment fixed effects and a binary treatment indicator. Column (6) constitutes the baseline specification. In the third panel with 'Boundary specific productivity' I additionally interact all boundary segments with all latitude and longitude polynomials and distance polynomials. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.26: Wealth effect of property rights: Excluding partially treated

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Log(Median Family Income)	0.215***	0.231***	0.251***	0.019	0.163***	0.180***	0.216***	0.260***	0.290***
	(0.028)	(0.041)	(0.044)	(0.071)	(0.056)	(0.063)	(0.034)	(0.035)	(0.044)
Share below Poverty Line	−0.039***	−0.049***	−0.063***	−0.011	−0.027*	−0.035**	−0.045***	−0.066***	−0.076***
	(0.010)	(0.016)	(0.016)	(0.014)	(0.014)	(0.014)	(0.008)	(0.017)	(0.018)
Log(Median Value Housing)	0.139***	0.152***	0.153***	−0.013	0.106**	0.145**	0.127**	0.183***	0.213***
	(0.045)	(0.041)	(0.051)	(0.066)	(0.043)	(0.063)	(0.055)	(0.039)	(0.045)
Observations	1,125	3,234	5,351	1,125	3,234	5,351	1,125	3,234	5,351

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.27: Wealth effect of property rights with Conley standard errors

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Log(Median Family Income)	0.131	0.161	0.174	0.047	0.095	0.108	0.139	0.197	0.233
Clustered by boundary segment	(0.039)	(0.044)	(0.043)	(0.037)	(0.033)	(0.046)	(0.034)	(0.032)	(0.035)
Spatial correlation within 0.5 degrees	(0.042)	(0.047)	(0.049)	(0.041)	(0.037)	(0.050)	(0.036)	(0.032)	(0.040)
Spatial correlation within 1 degree	(0.044)	(0.051)	(0.054)	(0.037)	(0.037)	(0.051)	(0.036)	(0.031)	(0.042)
Share below Poverty Line	-0.027	-0.035	-0.046	-0.027	-0.022	-0.024	-0.032	-0.047	-0.060
Clustered by boundary segment	(0.008)	(0.012)	(0.011)	(0.014)	(0.008)	(0.009)	(0.006)	(0.011)	(0.012)
Spatial correlation within 0.5 degrees	(0.009)	(0.011)	(0.011)	(0.013)	(0.009)	(0.010)	(0.007)	(0.009)	(0.011)
Spatial correlation within 1 degree	(0.008)	(0.011)	(0.012)	(0.014)	(0.009)	(0.010)	(0.006)	(0.008)	(0.011)
Log(Median Value Housing)	0.072	0.110	0.105	0.007	0.059	0.090	0.093	0.155	0.187
Clustered by boundary segment	(0.042)	(0.043)	(0.044)	(0.046)	(0.031)	(0.047)	(0.032)	(0.039)	(0.041)
Spatial correlation within 0.5 degrees	(0.042)	(0.045)	(0.045)	(0.047)	(0.032)	(0.046)	(0.027)	(0.035)	(0.041)
Spatial correlation within 1 degree	(0.043)	(0.048)	(0.051)	(0.047)	(0.031)	(0.046)	(0.026)	(0.036)	(0.049)
Observations	1,928	4,325	6,658	1,928	4,325	6,658	1,928	4,325	6,658

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.28: Wealth effect of property rights: using 6 miles boundary segments

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Log(Median Family Income)	0.070***	0.119***	0.152***	−0.023	0.046	0.073**	0.098**	0.102***	0.128***
	(0.024)	(0.019)	(0.035)	(0.055)	(0.031)	(0.031)	(0.047)	(0.031)	(0.027)
Share below Poverty Line	−0.019***	−0.025***	−0.041***	−0.006	−0.014	−0.020***	−0.030***	−0.033***	−0.037***
	(0.006)	(0.005)	(0.009)	(0.010)	(0.009)	(0.007)	(0.009)	(0.011)	(0.011)
Log(Median Value Housing)	0.036*	0.092**	0.101*	−0.092**	0.015	0.066**	0.023	0.048	0.105**
	(0.020)	(0.038)	(0.053)	(0.037)	(0.022)	(0.028)	(0.038)	(0.044)	(0.044)
Observations	1,928	4,325	6,658	1,928	4,325	6,658	1,928	4,325	6,658

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.29: Wealth effect of property rights: Using the boundary segment sample from the AVHRR data

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Log(Median Family Income)	0.084***	0.109***	0.221***	0.063	0.029	−0.002	0.050*	0.202***	0.302***
	(0.024)	(0.038)	(0.065)	(0.052)	(0.043)	(0.047)	(0.030)	(0.040)	(0.083)
Share below Poverty Line	−0.023**	−0.027**	−0.063***	−0.062***	−0.023*	−0.010	−0.029***	−0.057***	−0.084***
	(0.010)	(0.012)	(0.022)	(0.016)	(0.014)	(0.012)	(0.006)	(0.014)	(0.027)
Log(Median Value Housing)	0.108*	0.110	0.211**	0.101*	0.073	0.078	0.084	0.209**	0.315***
	(0.057)	(0.075)	(0.092)	(0.058)	(0.053)	(0.080)	(0.052)	(0.095)	(0.119)
Observations	741	1,631	2,587	741	1,631	2,587	741	1,631	2,587

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.30: Wealth effect of property rights: Additional outcomes

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
Mortgage Share	0.019*	0.029***	0.033***	−0.015	0.024	0.019	0.005	0.014	0.021
	(0.011)	(0.011)	(0.010)	(0.027)	(0.017)	(0.014)	(0.012)	(0.013)	(0.015)
log(Per Capita Income)	0.070*	0.094**	0.094**	0.054	0.065*	0.077	0.103***	0.169***	0.204***
	(0.040)	(0.039)	(0.037)	(0.051)	(0.039)	(0.047)	(0.038)	(0.032)	(0.040)
Bachelor degree	0.032**	0.038***	0.025*	0.011	0.024*	0.034	0.042***	0.062***	0.059***
	(0.014)	(0.014)	(0.014)	(0.016)	(0.014)	(0.021)	(0.011)	(0.010)	(0.013)
Households with Social Security Income	−0.014*	−0.017**	−0.019***	0.007	−0.002	−0.010	−0.006	−0.001	0.003
	(0.008)	(0.008)	(0.007)	(0.022)	(0.013)	(0.012)	(0.008)	(0.006)	(0.005)
Households with Public Assistance Income	−0.005	−0.012***	−0.013***	−0.001	−0.003	−0.006	−0.009***	−0.017***	−0.019***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
Observations	1,928	4,325	6,658	1,928	4,325	6,658	1,928	4,325	6,658

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.31: Heterogeneous effect on vegetation

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles	1 mile	2 miles	3 miles
County Without Police									
Access rights	0.003	−0.007	−0.010	0.013	0.027	0.017	0.028	0.024	0.028
	(0.021)	(0.028)	(0.030)	(0.024)	(0.033)	(0.035)	(0.026)	(0.029)	(0.032)
Private rights	0.023	0.047	0.065	0.027	0.061*	0.073*	0.039	0.066**	0.085**
	(0.027)	(0.035)	(0.043)	(0.024)	(0.035)	(0.039)	(0.028)	(0.032)	(0.039)
County With Police									
Access rights	0.085***	0.136***	0.161***	0.103***	0.115***	0.138***	0.082***	0.131***	0.155***
	(0.026)	(0.036)	(0.041)	(0.034)	(0.035)	(0.040)	(0.029)	(0.041)	(0.046)
Private rights	0.098***	0.132***	0.156***	0.120***	0.140***	0.159***	0.089***	0.122***	0.147***
	(0.032)	(0.033)	(0.037)	(0.034)	(0.039)	(0.042)	(0.033)	(0.036)	(0.041)
City with civil service reform $>$ 100miles									
Access rights	0.002	0.006	0.014	0.004	−0.001	−0.007	−0.005	−0.009	−0.005
	(0.010)	(0.016)	(0.020)	(0.016)	(0.013)	(0.015)	(0.010)	(0.016)	(0.019)
Private rights	0.017	0.028*	0.038*	0.021	0.023	0.026	−0.000	0.010	0.018
	(0.012)	(0.016)	(0.021)	(0.017)	(0.015)	(0.016)	(0.012)	(0.016)	(0.019)
City with civil service reform $<$ 100miles									
Access rights	0.222***	0.331***	0.377***	0.234***	0.285***	0.343***	0.232***	0.347***	0.394***
	(0.049)	(0.064)	(0.070)	(0.057)	(0.063)	(0.072)	(0.053)	(0.070)	(0.077)
Private rights	0.249***	0.322***	0.376***	0.253***	0.336***	0.385***	0.287***	0.354***	0.407***
	(0.071)	(0.068)	(0.074)	(0.061)	(0.074)	(0.080)	(0.074)	(0.071)	(0.076)
County Without Bank									
Access rights	0.095***	0.136***	0.164***	0.098***	0.109**	0.121**	0.053*	0.060*	0.067*
	(0.032)	(0.039)	(0.045)	(0.037)	(0.045)	(0.048)	(0.029)	(0.033)	(0.037)
Private rights	0.151**	0.156***	0.186***	0.141**	0.175***	0.193***	0.125*	0.098**	0.110**
	(0.067)	(0.046)	(0.046)	(0.054)	(0.062)	(0.057)	(0.072)	(0.044)	(0.043)
County With Bank									
Access rights	0.074***	0.111***	0.123***	0.091***	0.097***	0.110***	0.075***	0.116***	0.130***
	(0.026)	(0.036)	(0.040)	(0.032)	(0.033)	(0.039)	(0.028)	(0.040)	(0.045)
Private rights	0.058**	0.096***	0.111***	0.083***	0.098***	0.112***	0.057**	0.099***	0.116***
	(0.023)	(0.030)	(0.032)	(0.029)	(0.033)	(0.037)	(0.022)	(0.032)	(0.038)
County Without Newspaper									
Access rights	0.041**	0.058**	0.069**	0.036*	0.046*	0.044*	0.033**	0.041**	0.050**
	(0.018)	(0.024)	(0.028)	(0.021)	(0.023)	(0.025)	(0.016)	(0.020)	(0.023)
Private rights	0.070***	0.081***	0.099***	0.076***	0.090***	0.096***	0.053*	0.062***	0.079***
	(0.027)	(0.024)	(0.027)	(0.024)	(0.028)	(0.029)	(0.027)	(0.022)	(0.024)
County With Newspaper									
Access rights	0.155***	0.252***	0.296***	0.200***	0.231***	0.282***	0.137**	0.211***	0.240***
	(0.050)	(0.067)	(0.076)	(0.059)	(0.064)	(0.073)	(0.056)	(0.079)	(0.092)
Private rights	0.116**	0.198***	0.238***	0.145**	0.201***	0.249***	0.122**	0.184***	0.213**
	(0.048)	(0.061)	(0.068)	(0.055)	(0.065)	(0.073)	(0.049)	(0.070)	(0.082)

Every column in every panel separates a separate regression using the model in the header, and splitting by the variable in the panel header. Outcome variable is log(NDVI) and an observation is defined as access rights if its center is within the historical grazing district and has not been privatized in 1935. An observation is defined as private rights if its center is within the historical grazing districts and has been privatized after 1916. Police is defined as zero if not person in the 1930 census is a police men in that county. Out of 206 counties, 61 counties have no policemen. Bank is defined as zero if the county had no bank in 1934. Downloaded from the Federal Deposit Insurance Corporation (2001). Out of 206 counties, 54 counties had no bank. Newspaper is defined as zero if the county had no newspaper in 1932 (Gentzkow et al., 2014). Out of 206 counties, 154 counties have no Newspaper. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.32: Heterogeneous wealth effect of property rights: Enforcement

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
				log(Median Family Income)					
No Police	-0.142^{**}	-0.059	-0.002	-0.077	0.003	-0.030	-0.090	-0.090	-0.063
	(0.066)	(0.058)	(0.058)	(0.103)	(0.088)	(0.070)	(0.137)	(0.091)	(0.060)
With Police	0.153^{***}	0.185^{***}	0.199^{***}	0.046	0.106^{***}	0.112^{**}	0.143^{***}	0.218^{***}	0.262^{***}
	(0.036)	(0.042)	(0.041)	(0.038)	(0.033)	(0.046)	(0.036)	(0.033)	(0.032)
				Share below Poverty Line					
No Police	0.034	0.021^{*}	0.015	0.015	0.025	0.028	0.032	0.029	0.017
	(0.020)	(0.012)	(0.013)	(0.023)	(0.019)	(0.018)	(0.034)	(0.022)	(0.019)
With Police	-0.034^{***}	-0.041^{***}	-0.052^{***}	-0.030^{**}	-0.027^{***}	-0.028^{***}	-0.035^{***}	-0.053^{***}	-0.068^{***}
	(0.007)	(0.012)	(0.011)	(0.012)	(0.008)	(0.008)	(0.005)	(0.011)	(0.012)
				log(Median House Value)					
No Police	-0.102	0.007	0.010	0.038	0.009	0.005	0.001	0.038	0.042
	(0.100)	(0.090)	(0.075)	(0.195)	(0.124)	(0.110)	(0.189)	(0.148)	(0.117)
With Police	0.088^{**}	0.118^{***}	0.115^{**}	-0.004	0.063^{*}	0.086^{*}	0.089^{**}	0.159^{***}	0.197^{***}
	(0.043)	(0.045)	(0.047)	(0.045)	(0.033)	(0.047)	(0.039)	(0.044)	(0.045)
Observations without Police	171	308	400	171	308	400	171	308	400
Observations with Police	1,779	4,085	6,354	1,779	4,085	6,354	1,779	4,085	6,354

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Police is defined as zero if not person in the 1930 census is a police men in that county. Out of 321 counties, 84 counties have no policemen. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.33: Distance to closest city as a proxy for public good provision

	(1)	(2)	(3)	(4)	(5)
	log(NDVI)		Log Median Family Income	Share below Poverty Line	Log Median Value Housing
Access rights	0.105***	0.115***			
	(0.027)	(0.031)			
× Distance to closest City	0.010				
	(0.017)				
× Distance to closest City with Pop≥10,000		−0.032			
		(0.023)			
Private rights	0.104***	0.109***			
	(0.024)	(0.028)			
× Distance to closest City	0.025				
	(0.017)				
× Distance to closest City with Pop≥10,000		−0.006			
		(0.020)			
			Distance to Closest City		
Inside Grazing District			0.146***	−0.033***	0.089*
			(0.053)	(0.012)	(0.049)
× Distance to closest City			0.110*	−0.018	0.093*
			(0.059)	(0.014)	(0.050)
			Distance to Closest City with Pop≥10,000		
Inside Grazing District			0.126***	−0.029***	0.073*
			(0.044)	(0.011)	(0.044)
× Distance to closest City with Pop≥10,000			−0.026	0.009	−0.000
			(0.030)	(0.008)	(0.029)
Observations	56,667	56,667	4,325	4,325	4,325

In this Table I use the distance to the closest city as a proxy for quality of governance. Larger cities are more likely to have police, and thus if the estimated effect in Table 1.10 is driven by the size of the city, it should also show up in this Table. In the two columns the distance of every pixel to the closest city with civil service reform is calculated. Distance is standardized to give the interaction an "one standard deviation increase" interpretation. One standard deviation is 80 miles. A census-block is treated if its center is within the grazing districts with control observations being blocks outside the grazing districts. All columns control for the size of the census-block and year fixed effects. Standard Deviations for the AVHRR data: 11 miles and 33 miles, and for the census blocks 1.8 and 18 miles. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.34: Heterogeneous wealth effect of property rights: Financial access

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
				log(Median Family Income)					
No Bank	0.156	0.129	0.145*	−0.211	0.112	0.130	0.046	0.229*	0.498**
	(0.141)	(0.085)	(0.074)	(0.237)	(0.104)	(0.118)	(0.185)	(0.119)	(0.189)
With Bank	0.126***	0.166***	0.186***	0.025	0.093***	0.097**	0.132***	0.197***	0.239***
	(0.039)	(0.045)	(0.044)	(0.038)	(0.033)	(0.047)	(0.035)	(0.034)	(0.032)
				Share below Poverty Line					
No Bank	−0.012	−0.050*	−0.070**	0.114	−0.018	−0.058	0.005	−0.032	−0.116*
	(0.038)	(0.027)	(0.027)	(0.073)	(0.031)	(0.035)	(0.028)	(0.044)	(0.069)
With Bank	−0.026***	−0.035***	−0.046***	−0.026**	−0.023***	−0.021**	−0.035***	−0.048***	−0.061***
	(0.008)	(0.012)	(0.012)	(0.011)	(0.008)	(0.009)	(0.005)	(0.011)	(0.011)
				log(Median House Value)					
No Bank	0.056	0.166*	0.119	−0.422***	0.037	0.150	−0.673***	0.136	0.733***
	(0.078)	(0.094)	(0.099)	(0.133)	(0.143)	(0.159)	(0.135)	(0.374)	(0.267)
With Bank	0.069*	0.108**	0.105**	0.003	0.060**	0.083*	0.083**	0.140***	0.163***
	(0.038)	(0.045)	(0.047)	(0.041)	(0.029)	(0.044)	(0.037)	(0.042)	(0.040)
Observations without Bank	128	279	430	128	279	430	128	279	430
Observations with Bank	1,766	3,964	6,091	1,766	3,964	6,091	1,766	3,964	6,091

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Bank is defined as zero if the county had no bank in 1934. Downloaded from the Federal Deposit Insurance Corporation (2001). Out of 321 counties, 60 counties had no bank. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.35: Heterogeneous wealth effect of property rights: Newspaper

	Baseline			Interacted with Distance			Boundary specific productivity		
	(1) 1 mile	(2) 2 miles	(3) 3 miles	(4) 1 mile	(5) 2 miles	(6) 3 miles	(7) 1 mile	(8) 2 miles	(9) 3 miles
				log(Median Family Income)					
No Newspaper	−0.054	−0.038	−0.017	0.064	0.051	0.033	0.022	0.049	0.061
	(0.040)	(0.044)	(0.043)	(0.070)	(0.064)	(0.062)	(0.065)	(0.056)	(0.047)
With Newspaper	0.127***	0.167***	0.214***	−0.031	0.068*	0.107**	0.085***	0.164***	0.219***
	(0.032)	(0.039)	(0.040)	(0.040)	(0.034)	(0.042)	(0.018)	(0.040)	(0.048)
				Share below Poverty Line					
No Newspaper	0.013	0.004	−0.006	−0.014	0.008	−0.001	0.001	−0.015	−0.021
	(0.011)	(0.010)	(0.012)	(0.017)	(0.015)	(0.013)	(0.022)	(0.020)	(0.018)
With Newspaper	−0.031***	−0.038***	−0.055***	−0.019	−0.025***	−0.027***	−0.029***	−0.033***	−0.051***
	(0.006)	(0.012)	(0.013)	(0.015)	(0.010)	(0.009)	(0.006)	(0.011)	(0.012)
				log(Median House Value)					
No Newspaper	−0.081*	−0.008	−0.011	0.046	0.092	0.103	0.026	0.125**	0.099*
	(0.044)	(0.055)	(0.050)	(0.088)	(0.070)	(0.067)	(0.080)	(0.060)	(0.053)
With Newspaper	0.051	0.096*	0.113*	−0.073	0.003	0.056	0.037	0.106	0.164**
	(0.042)	(0.057)	(0.059)	(0.046)	(0.034)	(0.053)	(0.034)	(0.068)	(0.070)
Observations without Newspaper	545	1,200	1,776	545	1,200	1,776	545	1,200	1,776
Observations with Newspaper	1,383	3,125	4,882	1,383	3,125	4,882	1,383	3,125	4,882

Wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the model in the header. A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Newspaper is defined as zero if the county had no newspaper in 1932 (Gentzkow et al., 2014). Out of 321 counties, 230 counties have no Newspaper. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.36: Spread of Information: comparing local newspaper to radio

	Newspaper			Radio share, standardized		
	(1) Average Effect	(2) without Newspaper	(3) with Newspaper	(4) Average Effect	(5) without Newspaper	(6) with Newspaper
			log(Median Family Income)			
Access rights	0.161***	−0.038	0.167***	0.147***	−0.086***	0.158***
	(0.044)	(0.044)	(0.039)	(0.035)	(0.032)	(0.050)
Access rights x Radio				0.054***	−0.101***	0.023
				(0.019)	(0.024)	(0.029)
			Share below Poverty Line			
Access rights	−0.035***	0.004	−0.038***	−0.032***	0.009	−0.037***
	(0.012)	(0.010)	(0.012)	(0.010)	(0.007)	(0.014)
Access rights x Radio				−0.011	0.014*	−0.003
				(0.008)	(0.008)	(0.009)
			log(Median House Value)			
Access rights	0.110**	−0.008	0.096*	0.098**	−0.063	0.093
	(0.043)	(0.055)	(0.057)	(0.046)	(0.052)	(0.075)
Access rights x Radio				0.036	−0.088**	0.013
				(0.027)	(0.043)	(0.046)
Std. Dev.				0.140	0.147	0.134

Log(NDVI) from the AVHRR data series and wealth effects using census-block Groups in 1990, 2000, and 2010. Every cell constitute a separate regression of the treatment indicator on the variable in the first using the restriction and model in the header. Sample is split into presence of local newspapers from Gentzkow et al. (2014). A census-block is treated if its center is within the historical grazing districts. Control observations census-blocks outside the historical grazing districts. All columns control for the size of the census-block and a year fixed effect. Radio share in 1930 is standardized within sample. Data from Strömberg (2004). I abstain from instrumenting the radio share in 1930 with ground conductivity due to the low F-stats. F-stat: 5.36 within the nine states, F-stat: within the 2 mile sample: 0.30 , and F-stat for the entire US: 35.42. Standard errors clustered by the boundary segments shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.B Appendix: Agricultural Census

Here, I exploit two versions of a standard DID estimation equation with county (α_c), year (α_t), and state $\times$ year ($\alpha_{st} \times \alpha_t$) fixed effects. The inclusion of α_c captures all unobservable county characteristics, and $\alpha_{st} \times \alpha_t$ captures any change in state policy that might affect the outcome. In its most basic form, I estimate:

$$\log \mathrm{Y}_{c,t} = \sum_{s=1910}^{T=2007} \beta_s \text{Access rights}_c \times \mathbf{I}[t = s] + \alpha_c + \alpha_t + \alpha_{st} \times \alpha_t + \varepsilon_c \quad (1.4)$$

where I regress farm values per county and survey period ($Y_{c,t}$) on an indicator variable of whether any part of the grazing district is within the county borders (Access rights$_c$). I allow the coefficient β_s to vary by time to verify the assumption of similar pre-trends. As selection into treatment is potentially endogeneous, identifying a causal effect requires that any unobserved characteristics are linearly additive. Testing this linearity assumption requires $\beta_{1910} = \beta_{1920} = \beta_{1925} = \beta_{1930} = \beta_{1935} = 0$ and ensures that selection is not based on differential pre-trends and any post-treatment difference is due to treatment. In this setup, the point estimates β_{1940} and β_{1945} capture the immediate effects of the reform on farm values.

In this section, I regress the dependent variable on an indicator for post-treatment, assuming no pre-trends.[86]

$$\log \mathrm{Y}_{c,t} = \beta \text{Access rights}_c \times \mathbf{I}[t > 1935] + \alpha_c + \alpha_t + \alpha_{st} \times \alpha_t + \varepsilon_c \quad (1.5)$$

The results in Table A.37 confirm the initial results using the regression discontinuity design and suggest larger farm values following the implementation of the Taylor Grazing Act. Since I control for county and year

[86] Full results of (1.4) available upon request.

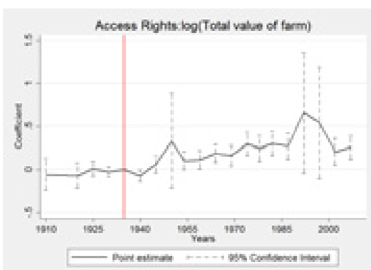
Table A.37: Income effect for farmers

	(1) log(Total value of farm)	(2) log(Average value of farm)	(3) log(Average value of farm land)	(4) log(#Farms)	(5) log(Average Farm Size)	(6) log(Acres)	(7) log(Average farm income)	(8) log(Average farm expenditures)	(9) log(Average farm profits)
Access Rights $\times \mathbf{I}[t > 1935]$	0.272***	0.115***	0.209***	0.075*	0.313***	0.394***	0.295**	0.155	0.183*
	(0.063)	(0.040)	(0.060)	(0.042)	(0.060)	(0.060)	(0.121)	(0.101)	(0.110)
Observations	5,628	5,619	5,608	5,641	5,625	5,628	4,177	4,684	3,455
Adjusted R^2	0.903	0.926	0.887	0.948	0.876	0.858	0.853	0.870	0.885

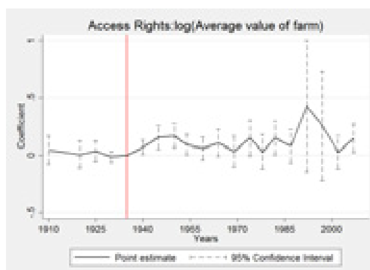
Wealth effects using for farmers using the agricultural census 1910–2007. Standard errors are clustered by county and are shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

fixed effect, a violation of the identification strategy is unlikely and indeed, Figure A.11 suggests that the assumption of common pre-trends are fulfilled.

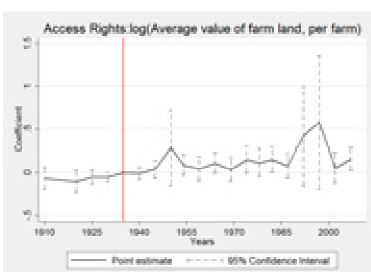
Figure A.11: Lead-lag graph: Access rights treatment



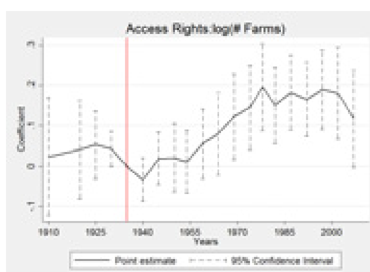
(a) log(Total value of farm)



(b) log(Average value of farm)



(c) log(Average value of farm land)



(d) log(#Farms)

Lead-lag graph for the main outcomes from the agricultural census using equation (1.4).

1.B.1 Instrumenting Access Rights

To counteract the possibility that counties endogenously select themselves into access rights regimes, I employ identification strategy based on soil erosion. As soil erosion in October 1934 was crucial in determining the location of the grazing borders, soil erosion strongly predicts whether a county was selected to be part of the Taylor Grazing Act. I trace the extend of soil erosion back to the Palmer Drought Severity Index [PDSI], which in turn is affected by rainfall in 1934. I show that the standardized rainfall in October 1934 strongly predicts selection into the Taylor Grazing Act and confirm a downward bias on the initial point estimates.

The natural instrument is soil erosion in October 1934, the point in time when the maps were drawn. Then, the exclusion restriction in equation (1.5) requires that soil erosion in 1934 only affected farm values through the policies of the Taylor Grazing Act. As soil erosion is greatly affected by weather fluctuations, it is likely that soil erosion maps in any other year would have been drawn to an entirely different extent. However, much like rainfall, soil erosion follows some historical average. As local erosion is influenced by differential rainfall with an unknown functional form, I cannot know the correlations between the average historical soil erosion and its temporary realization in 1934. Thus, I assume that the 1934 version of soil erosion was particularly severe, since it followed a period of relative drought (Figure 1.5). As rainfall was more beneficial in every year thereafter, even absent the Taylor Grazing Act, soil erosion would not have been as severe as in 1934, and is thus not likely to have influenced other policies or farm values.

As I cannot rule out that soil erosion in 1934 only affected farm values through the Taylor Grazing Act, I document a strong correlation between erosion and treatment status using the standard Palmer Drought Severity Index in October 1934, the month the erosion maps were drawn. I standardize the PDSI using the historical mean and standard deviation of each county and create a standardized index in October 1934 where larger values indicate less drought.[87] I show the first-stage relationship in Figure A.12a and plot all possible months-year combinations to highlight the importance for treatment assignment in Figure A.13a and severity of erosion in Figure A.13c.

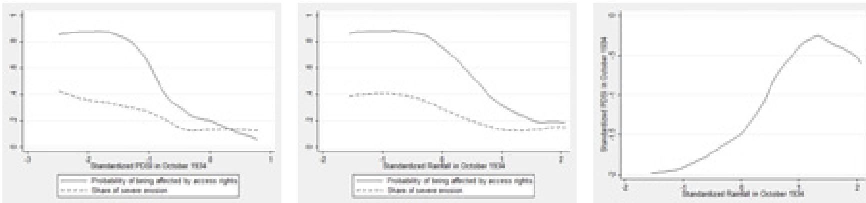
Since droughts are the consequence of missing precipitation, I trace back the PDSI in October 1934 to the standardized amount of rainfall during the same month per county. I first show the first-stage relationship in Figure A.12b and the relationship between rainfall and the PDSI in Figure A.12c. Both figures suggest significant relationships, and indeed the placebo estimates in Figure A.13b suggest high relevance for treatment assignment and the severity of erosion as measured by the erosion maps (Figure A.13d).

The results are shown in Table A.38. Instrumenting treatment assignment with rainfall in October 1934 results in significant effects on farm values and fewer farms in the last panel. The first stage F-Statistic is always large, and the point estimates from the instruments $\times$ Post TGA suggest significant impacts on the variables of interest. Column (1), (6) and (7) show my preferred estimates, the OLS, the reduced form estimates and the results from instrumenting treatment assignment with having rain in October 1934.

The estimated effects support the initial results from the regression discontinuity design in a different setting with different identification assumptions. In the RDD setting, identification relies on the exact measurement of the boundaries and their exogeneity to local characteristics. In the differences-in-differences design, we could allow for differential selection, as long as this selection is not based on time trends. By showing the point estimates for all years and instrumenting treatment assignment with rainfall, we show that these identification assumptions are indeed valid.

[87] Data description `https://climatedataguide.ucar.edu/climate-data/palmer-drought-severity-index-pdsi`.

Figure A.12: First stage relationship to Access Rights



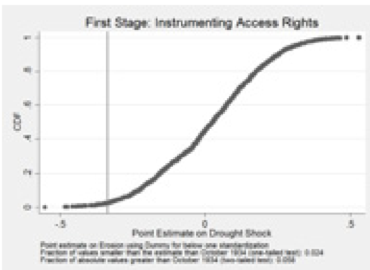
(a) Palmer Drought Severity Index (PDSI) in October 1934

(b) Standardized Rainfall in October 1934

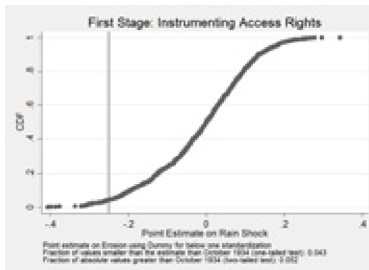
(c) Rainfall and the Palmer Drought Severity Index

The left figure show the first stage relationship between the Palmer Drought Severity Index (PDSI) and the severity of erosion and having access rights. The Figure in the center shows the relationship between rainfall during drawing the grazing districts in October 1934 and having access rights. The right figure shows the relationship between the PDSI and rainfall in October 1934. A one standard deviation increase in rain, increases the PDSI by 0.49 standard deviations.

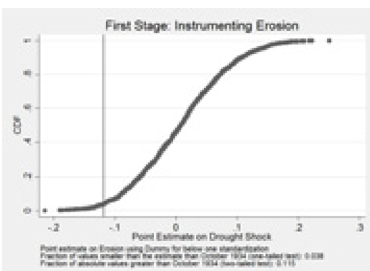
Figure A.13: Placebo graphs for the Palmer Drought Severity Index and Rainfall



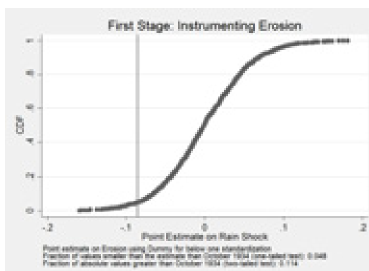
(a) PDSI and Access Rights



(b) Rainfall and Access Rights



(c) PDSI and Severe Erosion



(d) Rainfall and Severe Erosion

Placebo estimates using all months and years between 1900-2015 for the Palmer Drought Severity Index (PDSI) and 1915-2011 for the rainfall. Point estimates show the regression of the instrument in a given month-year combination on the access rights treatment (upper panel) or the severity of erosion (lower panel) including state fixed effects. The line marks the first stage point estimate. Two statistics are show below each figure. The fraction of values smaller than the first stage and the fraction of values that are greater in absolute terms than the first stage.

Table A.38: The effect of access rights on farm values using the Agricultural census

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log(Total value of farm)			
Access Rights × Post TGA	0.272***		0.582**		−0.015		0.250*
	(0.063)		(0.289)		(0.142)		(0.132)
Share of Severe Erosion × Post TGA		0.207**					
		(0.102)					
PDSI × Post TGA				0.005			
				(0.048)			
Rainfall × Post TGA						−0.063*	
						(0.032)	
				log(Average value of farm)			
Access Rights × Post TGA	0.115***		0.558**		0.150		0.375***
	(0.040)		(0.271)		(0.098)		(0.113)
Share of Severe Erosion × Post TGA		0.199**					
		(0.082)					
PDSI × Post TGA				−0.051			
				(0.034)			
Rainfall × Post TGA						−0.094***	
						(0.024)	
				log(Average value of farm land)			
Access Rights × Post TGA	0.209***		1.058***		0.262*		0.435***
	(0.060)		(0.404)		(0.137)		(0.133)
Share of Severe Erosion × Post TGA		0.377***					
		(0.116)					
PDSI × Post TGA				−0.088*			
				(0.047)			
Rainfall × Post TGA						−0.109***	
						(0.031)	
				log(# Farms)			
Access Rights × Post TGA	0.075*		−0.313		−0.235**		−0.217**
	(0.042)		(0.237)		(0.096)		(0.105)
Share of Severe Erosion × Post TGA		−0.112					
		(0.073)					
PDSI × Post TGA				0.079***			
				(0.030)			
Rainfall × Post TGA						0.055**	
						(0.025)	
	OLS	RF	IV	RF	IV	RF	IV
First stage F-Statistic			12.821		90.546		41.881
Observations	5,641	5,641	5,641	5,641	5,641	5,641	5,641

In this table I present the main outcomes from the Agricultural Census using the OLS estimates (column 1), and three instruments based on the erosion status in 1934 (column 2 and 3), the standardized Palmer Drought Severity Index in October 1934 (column 4 and 5), and the standardized rainfall in October 1934 (column 6 and 7). Standard errors clustered by county shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.C Appendix: Theory

1.C.1 Farm prices under open-access and access rights

Consider a farmer with private land, of which she derives a per period utility of one. The value of her farm is derived from the discounted sum of all future payoffs and given by $\frac{1}{1-\beta}$, where β is the per period discount factor. Additionally, the farmer has access to the public domain δ^i. The public domain δ^i is indistinguishable before the implementation of the reform, but are split into public lands δ^T and open-access δ. It is assumed that the public domain is less productive than private land $\delta^i < 1$, such that no new farmer has an incentive to buy it.[88] The sum of all future payoffs is thus given by:

$$V^{AR} = \sum_{t}^{\infty} \beta^t \delta^i = \frac{\delta^i}{1-\beta}, \tag{1.6}$$

where AR stands for access rights. With no entry, her farm value, determined by the discounted sum of all future payoffs from her private farm and her profits from the public domain. A farmer thus sells her farm if the price of private rights (P^{PR}) and access rights (P^{AR}) plus the discounted expected income in another occupation $\mathbb{E}[I]$ is larger than the value of the farm.

$$P^{PR} + P^{AR} + \mathbb{E}[I] \geq V^{PR} + V^{AR} \tag{1.7}$$

Assuming that the price of private rights is fair ($V^{PR} \approx P^{PR}$), a farmer thus sells her farm if the price of the off-farm income is priced to compensate for the loss in income.[89] For the seller the problem is similar, but the expectation is over the value of the off farm income. A seller buys if:

[88] Existing farmers have an incentive to buy it as acquiring this land would provide them with economies of scale, but due to homesteading, the public domain was hard to acquire before the reform.

[89] Note that someone with expected income $\mathbb{E}[I]$ larger than the value of access rights would have sold the farm prior to the reform.

$$P^{PR} + P^{AR} \leq V^{PR} + \mathbb{E}\left[V^{AR}\right] \tag{1.8}$$

Hence, in equilibrium with fair pricing a farm is sold if the expected valuation of the buyer matches the income loss of the seller:

$$V^{AR} - \mathbb{E}\left[I\right] \leq P^{AR} \leq \mathbb{E}\left[V^{AR}\right] \tag{1.9}$$

In the extreme, moving from no valuation ($\mathbb{E}\left[V^{AR}\right] = 0$) to full realization of profits ($\mathbb{E}\left[V^{AR}\right] = V^{AR}$), increases the sales of farms.

1.C.2 With entry into the public domain

The above framework can be extended to include shared resources, uncertain entry into the public domain and price bargaining. The farmer shares access to the public land with R other farmers in a stable collusive agreement. That is:

$$V^C(R) = \sum_t^{\infty} \beta^t \frac{\delta^i}{R} = \frac{\delta^i}{(1-\beta)R} \tag{1.10}$$

However, in every period t with probability $1-p$, $N-R$ new entrants compete in the public domain δ if no entry occurred before.[90] If in any previous period $N-R$ farmers entered, no future entry is possible.[91] If no entry occurs, the collusion payoff with R farmers is realized. After entry, all N ranchers decide to collude and obtain the collusion payoff $V^C(N)$. The uncertainty of new competition in the public domain reduces the expected future profits from the public domain for each farmer:

[90]For simplicity, assume that entry does not reduce the size of the public domain.

[91]The underlying assumption is that once N farmers are in the market for δ, no farmer has an incentive to enter the public domain.

$$\begin{aligned}
\mathbf{E}\left[V^C(R)\right] &= \sum_t^{\infty} \beta^t \left[p^t(1-p)V^C(N) + p^{t+1}V^C(R)\right] \\
&= \sum_t^{\infty} (p\beta)^t \left(\frac{(1-p)}{1-\beta} \times \frac{\delta^i}{N} + p\frac{\delta^i}{R}\right) \\
&\Rightarrow \frac{1}{1-p\beta}\frac{(1-p)}{1-\beta}\frac{\delta^i}{N} + \frac{p}{1-p\beta}\frac{\delta^i}{R}
\end{aligned}$$

If the farmer now tries to sell her land, the price for her land is determined by the future payoffs of her own land, plus the expected payoff for the share of public domain. Since the farmer cannot guarantee the latter, the price for the share of profits from the public domain lies in the interval $[0, \mathbf{E}\left[V^C(R)\right]$. If the price P is uniformly distributed the average price is given by:[92]

$$P(p, \delta^i, R) = \frac{1}{2}\left(\frac{1}{1-p\beta}\frac{(1-p)}{1-\beta}\frac{\delta^i}{N} + \frac{p}{1-p\beta}\frac{\delta^i}{R}\right) \tag{1.11}$$

which depends on the probability of no entry p, the land quality δ^i, and the number of farmers active.

The implementation of the reform had two distinct effects on this price. First, it divided lands into public lands δ^T and open-access lands δ. Both lands could be grazed and used by adjacent farmers. Secondly, for Taylor lands δ^T it decreased the probability of new entry to zero, as it assigned access rights to the R farmers that used the lands before. Since access rights were transferable, farmers could price their lands accordingly. For farmers with open-access lands, the probability of new entry decreased as well such that with probability $\bar{p} \in (p, 1]$, farmers of the public land did not face entry. Here the baseline model can be obtained by setting $\bar{p} = 1$.

[92]For analytical convenience, any other distribution, or the extremes would yield the exact same result. The same results holds when the price is an outcome of a bargaining process where the bargaining power for the seller increases with the number of farmers active.

Differences-in-difference estimates compare the prices obtained, within those lands that were to be come Taylor lands, minus the difference in prices of those lands that remained public:

$$\Delta(\bar{p}, \delta, \delta^T, R) = \left[P(1, \delta^T, R) - P(p, \delta^T, R)\right] - \left[P(\bar{p}, \delta, R) - P(p, \delta, R)\right] \tag{1.12}$$

Since land quality has been shown to be the same in the previous section, we can assume that there is no difference in the average price before the reform $P(p, \delta^T, R) \approx P(p, \delta, R)$. Moreover, since farmers inside the Taylor lands could contract on δ^T, they were able to obtain $V^C(R)$ for their farm, reducing Δ to:

$$\Delta(\bar{p}, \delta, \delta^T, R) = V^C(R) - P(\bar{p}, \delta, R) \tag{1.13}$$

$$= \frac{\frac{\delta^T}{R}}{1-\beta} - \frac{1}{2}\left(\frac{1}{1-\bar{p}\beta}\frac{(1-\bar{p})}{1-\beta}\frac{\delta}{N} + \frac{\bar{p}}{1-\bar{p}\beta}\frac{\delta}{R}\right) \tag{1.14}$$

Since farmers inside Taylor lands could contract on the output δ^T, and enforce the mechanism, any distributional assumption on the price outside the Taylor lands, leads to a gain in price, and thus farm value.

Chapter 2

State Repression, Exit, and Voice: Living in the shadow of Cambodia's Killing Fields[1]

2.1 Introduction

Over the last century, state repression has cost the lives of millions, and several more have suffered from various forms of political persecution.[2] While there has been recent progress in our understanding of what influences repressive behavior (Besley and Persson, 2011; Yanagizawa-Drott, 2014), we know much less about whether state-led violence has been suc-

[1]Co-authored with Andreas Madestam. We are grateful to Gerard Padró i Miguel, Nathan Nunn, Stelios Michalopolous, David Yanagizawa-Drott, and seminar participants at Barcelona GSE Summer Forum - Advanced in Micro Development Economics, CEPR Politicial Economy of Conlfict and Development, CEMFI, European University Institute (Florence), Universitat Pompeu Fabra (Barcelona), MIT, Bonn University, Brown University, University of Bergen, Goethe University (Frankfurt), Stockholm University and The Economic History Association Meeting for valuable comments.

[2]The estimated deaths following China's Cultural Revolution, Stalin's terror in the Soviet Union, the genocide in Cambodia, the Rwandan genocide, the Holocaust, and the massacre of suspected communists in Indonesia, together exceed 11 million. Moreover, the number of surviving victims affected by repression during the Cultural Revolution alone accounts for over 22 million people (Walder, 2014).

cessful at silencing dissent or changed political beliefs.[3] The lack of empirical evidence on this question is surprising, given that the answer is critical to understanding why repression exist at all, and why it is so common.

In fact, it is an open question if government coercion leads to a generation of politically passive citizens or whether it motivates increased political participation. It is also not clear if people, who do mobilize, act in favor of strong authoritarian leaders that promise stability or if they embrace pluralism to avoid the concentration of power that may have caused repression in the first place. Understanding whether and how state repression influences the way citizens exercise political power is particularly important as the threat of political violence still matters in many post-conflict societies.

In this paper, we examine the political legacy of state coercion. Using evidence from one of history's most severe episodes of state-led repression, the genocide in Cambodia under the Khmer Rouge, we estimate the effects of political violence on political behavior in Cambodia four decades later. During their four-year rule, 1975-1979, the Khmer Rouge killed between 1.7-3 million people or over 20 percent of the population (Kiernan, 2008). Forty years after the genocide, Cambodia is a democracy but power has been in the hands of the same party, the Cambodia People's Party [CPP], and leader, Hun Sen, since the introduction of multiparty elections in 1993. The CPP often refer to its role as the guarantor of stability, keeping Cambodia from slipping back into the abyss of violence (Strangio, 2014; Giry, 2015). However, corruption is endemic and key elements of democracy such as civil liberties, a free press, and rule of

[3] See Yanagizawa-Drott (2014) for the role of mass media and state-sponsored violence and Davenport and Inman (2012) for an overview of the literature on state repression. It is important to distinguish civil war and two-sided violence between insurgents from one-sided political violence, with citizens suffering state repression (Besley and Persson, 2011). Whereas the political consequences of two-sided violence have received growing attention, work on the political effects of one-sided state repression is still in its infancy (see Bauer et al., 2016 for an overview).

law, have been repeatedly compromised since multiparty elections were introduced (Norén-Nilsson, 2016a).[4] Despite the scale of suffering caused by the genocide, there is no systematic evidence quantifying the effects of the Khmer Rouge's repressive regime on subsequent political outcomes.

To understand how repressive state behavior affects the way citizens exercise political power, we develop a simple model inspired by Hirschman (1970) work on exit and voice. Voters have preferences over pluralism, where more pluralism is illustrated by their support for the opposition as opposed to the long-term authoritarian incumbent. To capture that milder forms of repression are present in contemporary Cambodia, we assume that it is costly to openly express preferences against the incumbent. Voters can take two political actions, vote and engage in local civil society, with the important difference that preferences remain anonymous when citizens vote but are revealed when they participate in civil society. In the model, the median voter decreases her local civic engagement and exits civil society if the experience of state repression raises the expected cost of dissent.[5] If repression also changes the voter's preferences in favor of more pluralism, she uses exit and voice: she engages less in civil society but is more likely to vote for the opposition. The underlying mechanism is that the median voter expresses her preferences for pluralism in elections because she does not run the risk of detection; stating these preferences openly in civil society is, however, costly. In short, the experience of political violence makes voters more convinced about the need for opposing views but more cautious in expressing them.

A challenge when estimating the effect of state repression is that political violence most often occurs nationwide without any credible counterfactual. Even if the intensity of coercion varies, selective targeting of

[4]Cambodia ranks as one of the most corrupt country in the world (161 of 180) according to Transparency International latest corruption perceptions' index (`www.transparency.org`, accessed April, 2018).

[5]While Hirschman (1970) interpreted exit more literally, in the sense of physically leaving a location, we follow recent work where exit can take the form of abstaining from different political activities (see Clark et al., 2017).

specific regions or groups based on pre-war political views may confound estimates of post-repression beliefs and behavior. We address this problem by relying on the Khmer Rouge's desire to create an agrarian socialist society, where the regime displaced large parts of the population to labor camps to increase rice production. Many areas close to the camps became known as Killing Fields as laborers were executed or died of starvation and overwork (Chandler, 2008; Kiernan, 2008). We investigate how these Killing Fields affect the local population today. To establish causality, we explore the movement of forced labor to areas experiencing higher agricultural productivity. Local rainfall variation during the Khmer Rouge era generates exogenous variation in rice productivity and, hence, variation in the size of camps and the subsequent casualties. Conditional on the likelihood of rain, rainfall is a random event, arguably uncorrelated with other factors that affect political behavior in today's Cambodia.

We assemble unique commune- and individual-level data from Cambodia using information from a large number of historical and contemporary sources to measure the influence of the atrocities under the Khmer Rouge. We first show that significantly more people died in communes experiencing higher productivity during the Khmer Rouge era using geo-coded data on genocide casualties. We then estimate the effect of higher productivity under Khmer Rouge on a range of political outcomes to examine our hypotheses on citizens' use of exit and voice.

The results show that state repression leads to the use of voice in the form of political mobilization and stronger preferences in support of pluralism. Communes with higher productivity and more killings during the Khmer Rouge experience larger turnout, primarily favoring the opposition parties compared to the authoritarian incumbent in the three most recent elections. These communes also exhibit higher levels of political competition. Using election survey data, we corroborate the findings by showing that individuals living in these communes are more supportive of democratic principles and more politically informed.

At the same time, our analysis indicates that repression increases the use of exit from civil society as citizens are more cautious in their interactions with the local community. The individual-level survey data shows a decline in measures capturing membership and participation in local community organizations as well as a display of lower trust in communes with higher productivity under the Khmer Rouge era. We also present evidence that people in these places are less supportive of the local state and more likely to avoid local state interactions, as captured by lower local tax contributions and a lower likelihood of being a state employee. Therefore, our overall results show that state repression has made politics less personal and more competitive. Also, while the effects are more pronounced for people that were alive under the Khmer Rouge, we cannot statistically separate the difference for those born later suggesting that the legacy of political violence can have a persistent impact on society.

We conduct a number of robustness tests to assess the sensitivity of our identification strategy. Using US Army maps from the early 1970s, we show that Khmer Rouge era rainfall is orthogonal to important predetermined characteristics such as population density, geographic proximity, and state infrastructure. We also use variation in rainfall to assess the statistical significance of our results. Comparing the effect of rainfall during months that matter for rice production under Khmer Rouge to the distribution of placebo estimates of rainfall in the same months in all other three-year periods in 1951-2017, shows that our findings are clear outliers. We further address concerns regarding statistical inference (following Anderson, 2008), given that we test multiple hypotheses with our individual-level survey data. Together with several other tests, these findings demonstrate the reliability and significance of the results.

What are the underlying channels behind our findings? We contrast three possible explanations. First, people residing in areas more exposed to the atrocities during the Khmer Rouge are more likely to have been directly affected by the killings, suffering, and breakdown of trust, and also have parents, relatives, and neighbors with similar experiences. In addition, the memory of the violence is kept salient by annual ceremonies at some of the grave sites and by the use of these sites for political meetings during election years (Bennett, 2015). Second, political violence could have changed the demographics of the survivors resulting in compositional differences in population, age, gender, and education explaining some of the results. Related to this, differential migration patterns subsequent to the Khmer Rouge regime might also play a role. Third, assets, land, and earnings could have changed directly as an outcome of the labor camps, or indirectly, following post-Khmer Rouge investments in public infrastructure in places experiencing more political violence.

To investigate these hypotheses, we examine if contemporary population, age structure, gender ratios, assets, consumption, poverty indicators, market access, and public infrastructure are driven by productivity during the Khmer Rouge. We further study if there is evidence of differential migration just after the genocide. None of these characteristics turn out to be systematically and significantly explained by our measure of productivity. These findings are further corroborated by other historical evidence. While the Khmer Rouge singled out previous government supporters, suspected Khmer Rouge dissidents, and more educated individuals, the selection occurred across Cambodia's communes. Many people had to relocate, forcibly moving from one cooperative to the other (Rice and Tyner, 2017; Tyner, 2017a). Following the genocide, a majority returned to the villages they had occupied before 1975 (Desbarats, 1995). As victims came from across Cambodia, people residing near the Killing Fields today are more likely to have experienced the atrocities up close since a significant fraction of people were allowed to remain in their villages. Using our data, we also find that areas experiencing more violence as captured by higher productivity during the Khmer Rouge are more likely to have constructed war memorials to commemorate the political violence, further facilitating the persistence of beliefs at the local level. Together this suggest that our evidence is more consistent with the first hypothesis, where people's political preferences and behavior change as a result of experiencing state repression and because of the Killing Fields' presence today, acting as salient markers of past violence.

This paper advances economic research on state repression. In recent years, there has been some progress in our understanding of the causes of one-sided mass violence, where the state represses its citizens (Besley and Persson, 2011; Yanagizawa-Drott, 2014).[6] However, we know much less about the political consequences of government coercion and political violence. Existing literature focuses on the effects of civil war and two-sided violence between insurgents (see e.g. Bellows and Miguel (2009); Blattman (2009); Voors et al. (2012); Bauer et al. (2016) for an overview). A central finding of this work is that two-sided violence fosters cooperation as a result of increased pro-sociality toward in-group members. We contribute to this literature by showing how violence, in the absence of a group-cohesion component, can lead to less cooperation within civil society while still increasing political engagement in elections.[7] The result that state repression and one-sided violence induces withdrawal from local civic community interactions aligns with other examples from South-East Asia, where citizens purposely avoid relations with a coercive state (Scott, 2009).

Our paper adds to work examining the long-term consequences of conflict on trust (Nunn and Wantchekon, 2011), anti-Semitism (Voigtländer and Voth, 2012), and on social structure (Acemoglu et al., 2011). More broadly, it connects to papers emphasizing the persistence of political preferences and behavior generated via the experience of political ideology (Alesina and Fuchs-Schündeln, 2007), economic fluctuations (Giuliano and Spilimbergo, 2014), stock market participation (Malmendier and Nagel, 2011), and patriotic events (Madestam and Yanagizawa-Drott, 2011).

Section 2.2 provides background information on the Khmer Rouge era and the contemporary political setting in Cambodia and Section 2.3 presents a conceptual framework. Section 2.4 introduces our data and

[6]There is more work examining the determinants of civil war, see Blattman and Miguel (2010) for an overview.

[7]We also share the link between the climate and conflict literature insofar that our rainfall-induced productivity measure predicts more deaths (see e.g. Miguel et al., 2004; Burke et al., 2009; Dell, 2012; Ciccone, 2013; Hsiang et al., 2013; Harari and Ferrara, 2018).

2.5 deals with the empirical strategy. Section 2.6 discusses our results and robustness tests with Section 2.7 discussing alternative hypotheses. Section 2.8 concludes.

2.2 Historical Background

This section provides a brief overview of the Khmer Rouge era, the contemporary political situation, and the presence of the Killing Fields in Cambodia today.

2.2.1 The Khmer Rouge

Cambodia gained independence in 1953, with King Norodom Sihanouk dominating political life until the late 1960s (Chandler, 1988). In 1970, Sihanouk was removed through a coup by General Lon Nol. Nol in turn lost his power to the Khmer Rouge in April 1975, after a civil war where the US had supported the Nol regime via heavy carpet bombings of the country.[8] The four years to follow marks history's worst genocide, with 1.7-3 million or over 20 percent of the population dying, an era that ended when Vietnam invaded Cambodia and defeated the Khmer Rouge in early 1979 (Kiernan, 2008).[9]

Immediately after taking power, the Khmer Rouge set out to create an agrarian socialist society, collectivizing the economy by banning money, markets, and private property (Chandler, 2008). The aim was to leapfrog development through successive "four-year plans" that increased the national production of rice, allowing the regime to generate a surplus that

[8] Chandler (2008) argues that the bombings were the most important factor explaining the rise of the Khmer Rouge. From 1965 until 1973, the US dropped 2.7 million tons of ordnance on Cambodia, more than the Allies dropped during the entire WW2 (Owen and Kiernan, 2006).

[9] There is some disagreement over the exact number of people that died during the Khmer Rouge regime. Kiernan (2008) estimates a national toll of between 1.67 and 1.87 million people whereas other estimates reach as high as 3 million dead (see discussion in Heuveline, 1998).

could finance industrialization (Chandler et al., 1988; Twining, 1988).[10] To succeed, the Khmer Rouge displaced large parts of the population and forced people to live and work in labor camps across the country. In these camps, supporters of the old regime, former state officials, Khmer Rouge dissidents, and the educated were labelled "new" people whereas farmers who had lived in the insurgency areas made up the "base". While base people initially enjoyed better conditions (seen as more loyal and more trustworthy), both groups worked together in the camps (Twining, 1988; Kiernan, 2008). The cooperatives included several villages up to entire communes and laborers were organized into work groups, kemlang ping (full strength) and kemlang ksaoy (weak strength), where the former consisted mostly of adults and the latter of small children and the elderly (Tyner, 2017a).

The country in general, and camps in specific, was governed through a hierarchical military command (Heder and Tittemore, 2001). Each province, district, and commune had committees in charge of politics, security, and economics, respectively. Internal Khmer Rouge documents describe how provincial committees were responsible for organizing production, focusing on places where productivity was higher: "...attack wherever [we are] strongest" (Chandler et al., 1988, p. 20). To achieve this goal, special mobile work committees were responsible for the deployment of mobile work brigades (consisting of workers from the kemlang ping) to undertake specific projects, such as harvesting the fields (Rice and Tyner, 2017; Tyner, 2017a). The committees governed every aspect of life. People were required to attend "livelihood meetings" that served as a propaganda sessions about communist ideals and as confessions, with people admitting past political and ideological sins and informing on other

[10]Internal party documents reveal detailed accounts of how agriculture would lead the transformation of the economy. Specifically, the documents show an obsession with raising productivity, with Khmer Rouge cadre repeating the mantra of increasing rice production to three tons per hectare. By comparison, pre-Khmer rouge productivity averaged one ton per hectare (Chandler et al., 1988).

camp members. People who either expressed the wrong ideas or were accused of differing opinion ran the risk of being escorted from the camp and executed later on (Chandler, 1988; Thion, 1993). Children were also targeted by the Khmer Rouge to spy on their parents, creating a system where neighbors were rewarded for informing on neighbors, friends for informing on friends, and children for informing on parents (Yimsut, 2011; Bennett, 2015).

Despite the planning, rice production remained low. One reason was that the Khmer Rouge cadres lacked farming experience and were unfamiliar with the local conditions (Vickery, 1999; Ledgerwood and Vijghen, 2002). As the harvests failed, people were pushed even harder leading to further purges, not only of labors but also of local Khmer Rouge cadre for failing to meet production targets. By the end of 1978, the explosion of violence had completely upended collectivized agriculture across Cambodia (Hiebert, 2017). When Vietnam defeated the Khmer Rouge in early 1979, people who had been displaced returned back to the villages they had occupied before 1975 (Desbarats, 1995; Kiernan, 2008). Left in the rice fields were the remains of those who had been executed or died of starvation and overwork (Chandler, 1988; Kiernan, 2008).

2.2.2 Contemporary politics in Cambodia

Cambodia has been an electoral democracy since 1993. In the country's first multi-party elections, the current incumbent party, Cambodian People's Party [CPP], shared power with the Royalist party, FUNCINPEC. However, in 1997 the CPP ousted the FUNCINPEC Prime Minister and has gone on to win all subsequent elections after that. The CPP has been headed by the same leader, Hun Sen, since 1985, making him the longest serving Prime Minister in Asia (Baaz and Lilja, 2014; Strangio, 2014; Norén-Nilsson, 2016b).

CPP and Hun Sen were part of the Vietnamese coalition that liberated Cambodia from the Khmer Rouge in 1979 and CPP often refer to its role as the guarantor of stability, keeping Cambodia from slipping back into the abyss of violence (Strangio, 2014; Giry, 2015). Hun Sen explicitly refers to the horrors of the Khmer Rouge during electoral campaigns. In the 2013 elections, CPP trucks drove around Cambodia showing films including documentary footage of the Khmer Rouge era as well as the 1984 Hollywood-blockbuster "The Killing Fields" (Norén-Nilsson, 2016a).

While the economy has grown at almost 7% annually since the mid-1990s, corruption is endemic with Cambodia ranking as one of the most corrupt countries in the world and political patronage governs business, military, and state relations with CPP at the center of power (Un, 2015; Norén-Nilsson, 2016b).[11] Moreover, key elements of democracy such as civil liberties, a free press, and rule of law have been repeatedly compromised since multiparty elections were introduced (Norén-Nilsson, 2016b). Partly in response to Hun Sen's authoritarian rule, the two largest opposition parties formed an alliance, Cambodia National Rescue Party (CNRP), in 2012. While the policy platforms of CPP and CNRP share many elements, the CNRP has made stronger appeals to combat corruption and improve the legal system. At the same time, CNRP also resorts to a strong nationalist anti-Vietnamese rhetoric (Norén-Nilsson, 2016a). In the analysis that follows, we will focus on the electoral outcomes involving the CPP and CNRP in the three most recent election taking place in 2012, 2013, and 2017.

2.2.3 The Killing Fields today

Mass graves tracing back to the Khmer Rouge are still an important part of the landscape in contemporary Cambodia. Figure 2a shows the location of more than 300 known sites spread across the country. Not only are they physically present but the sites have also shaped post-Khmer Rouge political culture. Annual ceremonies are held at the grave sites to

[11] See also `http://www.worldbank.org/en/country/cambodia/overview` and `https://www.transparency.org/news/feature/corruption_perceptions_index_2017`.

remember the violence and CPP used the sites frequently during 1980s to legitimize the new government. Hun Sen has stated that "...the remains of those killed during Democratic Kampuchea will not be cremated because they remain the only evidence of the Khmer Rouge regime" (Bennett, 2015, p. 224). More recently, the sites have been used for political meetings, by the CPP and the opposition, during election years (Bennett, 2015; Tyner, 2017b). In many locations, memorials have also been constructed to commemorate the dead, leaving them as salient markers of past violence.

2.3 Conceptual Framework

In his seminal work, Hirschman (1970, 1978) suggests that in democracies, dissatisfied citizens have two options to voice their discontent. Either, they participate in civic society and actively try to change politics, or use elections to elect a new leader. In oppressive regimes, however, taking political actions might result in persecution. Here, citizens with opposing views are left with the choice to either exit civic society and decrease the possibility of being detected, or use the veil of anonymity and vote against the incumbent. In this section we describe a model that conceptualizes our findings in the spirit of 'exit' and 'voice' in oppressive regimes. Specifically, we contrast an observable action, civic participation, and an unobservable action, casting a vote, to differentiate the effects of state repression on preferences for pluralism and the expected cost of dissent. In contrast to the looming memory of Killing Fields increasing the cost of dissent, changed preferences might generate persistent effects across generations. More educated and informed voters likely increase the competitiveness of elections and elect more competent leaders. Politicians then find their ability to extract rents restricted, leading to better policies being implemented (Ades and Di Tella, 1999; Besley et al., 2010).

Individual preferences θ_i are uniformly distributed over an interval $[\theta_L, \theta_H]$, where higher values of θ indicate stronger support for pluralism. Every individual obtains a benefit from voting $B(\theta)$ with $B'(\theta) > 0$. Individuals with positions close to the authoritarian incumbent θ_L have weaker preferences for democracy, and hence obtain less utility from voting. At the other extreme, voters obtain large benefits when signaling their preferences for democracy. In autocratic regimes, individual preferences are revealed to the authority with probability $f(\theta)$ capturing the idea that extreme positions are easier to observe than nuanced differences in preferences.[12] Given that preferences are revealed, a voter faces a cost of dissent c which are uncertain at the beginning of a period. Given the expectation of the cost of dissent $\mathbb{E}[c]$, a voter decides whether to conduct an unobservable action V or to participate in civic society P.

To analyze the impacts of state repression we simplify the entire distribution of voters in relation to the cutoff where voters are indifferent:

$$\max_{V\in[0,1]} \quad V \times [B(\theta) - \gamma f(\theta)\mathbb{E}[c]] \quad \text{s.t.} \quad \mu_V(1-V) = 0 \tag{2.1}$$
$$\max_{P\in[0,1]} \quad P \times [B(\theta) - f(\theta)\mathbb{E}[c]] \quad \text{s.t.} \quad \mu_P(1-P) = 0 \tag{2.2}$$

Here, the Kuhn-Tucker-conditions μ_V and μ_P allow for absence from the ballot box $\mu_V = 0$ or civic society $\mu_P = 0$. This maximization defines two cutoffs $\{\theta^P, \theta^V\}$ where voters are indifferent between participation θ^P and voting for the opposition θ^V. Since $\gamma \in [0, 1]$ captures that casting a vote is less observable than civic participation, we know that every participant is also voting $\theta^V < \theta^P$. We derive three testable hypotheses from this setup that allow us to estimate the impact of state repression on preferences θ and the expected costs of dissent $\mathbb{E}[c]$.

[12]Both $B(\theta)$ and $f(\theta)$ are continuous and increasing in their arguments. To generate interesting cases we can assume $\frac{\partial^2 B(\theta)}{\partial\theta\partial\theta} < 0$ and $\frac{\partial^2 f(\theta)}{\partial\theta\partial\theta} = 0$. Then, both benefits and probability of detection increase with θ, but with decreasing rates for the benefits and constant rates for the probability we obtain a θ^P such that $f(\theta)\mathbb{E}[c] = B(\theta)$. Here, individuals with $\theta < \theta^P$ exit civic society.

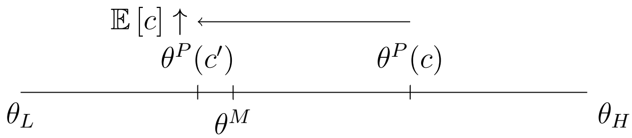
State repression and the cost of dissent In the first hypothesis, individuals who suffered under state repression have more accurate expectations about the cost of dissent $\mathbb{E}[c]$. In our model, this is reflected by an increased cost of dissent $c' > c$ which unambiguously decreases the respective cutoffs for voting $\theta^V(c') < \theta^V(c)$ and participation $\theta^P(c') < \theta^P(c)$.[13] Given unchanged preferences, every voter faces higher costs of detection which leads to exit from civic participation and fewer votes for the opposition. This shift is shown graphically in Figure 1a, where we depict the location of the median voter θ^M on the spectrum of preferences $[\theta_L, \theta_H]$. Initially, the median voter takes part in civic society as the expected cost of detection is lower than the benefit $B(\theta^M) \geq f(\theta)\mathbb{E}[c]$. As the cost of dissent increase to c', only voters with preferences close to θ^L remain in civic society, as for all others the cost outweigh the benefits.

Given that for voting $\gamma \approx 0$, we expect to see no changes in the voting behavior as preferences are unchanged, but strong responses in exiting local civic society due to the increased cost of detection.

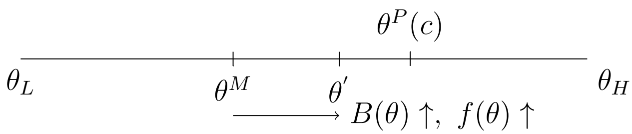
Hypothesis 1 'Exit': *If state repression increases the cost of dissent, vote shares for the opposition are unaffected and people exit local civic society.*

State repression and support for pluralism The second hypothesis captures the idea that individuals who suffered under state repression have stronger preferences for pluralism. In this case the cutoffs θ^V and θ^P remain unchanged and we can focus on the decisions of voters. If voting is less detectable $\gamma < 1$ and autocratic regimes allow for some voters to remain $\theta^V > \theta_L$, voting for the opposition increases. The effect for the observable action is ambiguous and depends on functional form assumptions. If the expected benefits of participation increase slower than

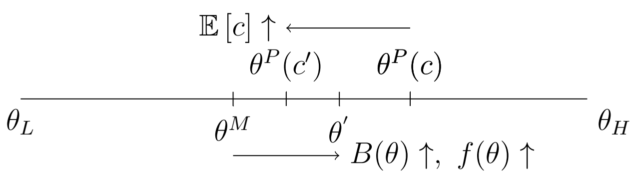
[13] The same prediction holds with concave utility functions if the variance of $\mathbb{E}[c]$ is decreasing, that is, individuals have a more precise idea of the cost of dissent.



(a) Increased expected cost of dissent



(b) Stronger support for pluralism



(c) Stronger support for pluralism and increased cost of detection

Figure 1: Mechanisms for participation in our model. An increase in the cost of dissent moves the cutoff for participation to $\theta^P(c')$ and hence the median voter at θ^M exits local civic society (a). An increase in the support for pluralism increases the benefits from participation, but also the risk of detection $f(\theta)$. Hence, the median voter is only continuing to participate if $\theta^{'} \leq \theta^P(c)$ (b). Combining the two effects in (c), the median voter exits civic participation due to the increased expected cost of detection.

the expected cost of participation $\frac{\partial B(\theta)}{\partial \theta} d\theta \leq \frac{\partial f(\theta)}{\partial \theta} d\theta$, previously indifferent voters exit local civic society.[14] In Figure 1b we show the case for the median voter, who remains active in the civic society as the increased risk of detection does not outweigh the increased benefits from participation. Hence, the predictions from an increase in preferences depend on the position of the median voter and the functional form assumptions on the benefits and detection probability.

Hypothesis 2 'Voice': *If state repression increases the support for pluralism in autocratic societies, vote shares for the opposition increase and the effects on local civic participation are ambiguous.*

State repression, the cost of dissent, and support for pluralism If state repression affects both the expected cost of dissent $\mathbb{E}[c]$ and voters preferences θ, two countervailing forces are at work. The increase of the cost of dissent decreases voting and participation which is partially offset by the increase in preferences. If votes are unobservable ($\gamma \approx 0$), vote shares unambiguously increase while previously indifferent voters exit local civic society if benefits increase less than costs $\frac{\partial B(\theta)}{\partial \theta} d\theta \leq \mathbb{E}[c] \frac{\partial f(\theta)}{\partial \theta} d\theta + \theta \frac{\partial \mathbb{E}[c]}{\partial c} dc$. As the benefits from participation increase, they are offset by an increase in the probability of detection due to increased preferences and the increasing cost of dissent making the voter less likely to participate. In Figure 1c, the increased benefits move the position of the median voter to $\theta^{'}$ where she continues to participate in the civic society. However, as the cost of dissent is increasing to c', the expected cost of detection are larger than the benefits, the cutoff moves to $\theta^P(c')$ and she exits civic society.

[14] It is important to note, that our aggregate predictions hold unambiguously for a uniform distribution of θ. If the distribution of voters is extreme value distributed, calculations of the average effect need to take into account the density of voters. Standard probabilistic voting models however assume uniform distributions which encourage us to make these aggregate predictions.

Hypothesis 3 'Exit and Voice': *If state repression affects both the preferences for pluralism and the cost of dissent, unobservable vote shares for the opposition increase, while local civic participation decreases.*

Combining the hypotheses, our model yields clear predictions for the channels at work. If state repression only increases the cost of dissent and votes are undetectable, vote shares are unaffected and people exit local civic society. If state repression increases preferences for pluralism, vote shares for the opposition increase and depending on the position of the median voter, she exits or continues to participate in local civic society. If both forces are affected by state repression, we should observe strong effects on vote shares and decreased local civic participation.

2.4 Data

We combine high resolution data on the genocide, electoral outcomes, and individual data to differentiate the hypotheses from our model. At first, we use data on genocide intensity to establish a first stage identifying communes that suffered more under the Khmer Rouge regime. Second, we use electoral outcomes to estimate the impact of genocide intensity on the preferences for pluralism. We then combine data from electoral surveys and large repeated cross-sectional surveys to identify individuals' preferences, local state avoidance, as well as potential confounding hypotheses.

Cambodia Genocide Project We use data from the Cambodian Genocide Program database to measure the intensity of violence and capture pre-genocide characteristics. First, the database contains 309 locations with 18,953 mass graves containing 974,734 bodies, which we combine with current commune boundaries to identify treated communes (Figure 2a). Second, we digitize commune characteristics from the US Army map series L7016 covering the entire country of Cambodia in 1970. These maps were used during the American bombing campaign and are of exceptional detail, containing relevant information on population density, temples, post offices, telephones, and agricultural productivity.[15] We calculate the area of each commune that is covered by forests, rice fields or is partially inundated, to capture the underlying productivity of each commune before the Khmer Rouge intervened. Third, the database contains 113,716 sites of bombing during the 1965–1973 campaign with a total volume of 2.7 million tons of explosives. Such bombings were targeted and likely indicate areas of stronger support for the Khmer Rouge. As we are interested in estimating a differential impact on preferences, we use this data to establish pre-treatment similarity of communes and argue for a causal effect of our instrument. Jointly, this database provides the most accurate description of violence in Cambodia and 13 pre-genocide characteristics that might have influenced the placement of labor camps (Table 1).

Election results Information on the national election in 2013 was obtained from Open Development Cambodia and constitutes our base map of communes. Information on communal elections in 2012 and 2017 were digitized from the official website of the national election office. Since the Cambodian National Rescue Party [CNRP] was formed in 2013 to unify the opposition, we aggregate votes of the 'Sam Rainsy Party' and 'Human Rights Party' in the 2012 commune elections to match the coalition from the 2013 national election.

Survey of the electorate The Asia Foundation conducted limited surveys on the political preferences of 2,008 individuals in 2003 and 2013 which we use to test the hypotheses from our model. We identify 24 questions that unambiguously identify voter informedness, support for democracy, local civic participation, or trust. Only questions that speak

[15] In Figure 3 we show that the maps are strongly correlated with aerial photographs from the same time.

strongly in favor of these channels and clearly distinguish an supporter of the Cambodians Peoples Party [CPP] from an CNRP supporter are included.[16] We exclude questions that aim at current beliefs about the direction of Cambodia or the commune.[17] The full set of selected questions is provided in Appendix 2.B and their summary statistics in Table A.7. Following Anderson (2008), we standardize each question in our four categories and sum the standardized outcomes weighting each outcome by the inverse of the covariance matrix of the standardized outcomes.[18] These four indexes help reducing the threat of multiple hypothesis testing and capture changes in preferences that individual questions only measure imperfectly. We follow Cantoni et al. (2017) and provide the results on the individual questions with the estimated p-values and FDR adjusted p-values in Table A.7.[19]

Cambodia Socio-Economic Survey We provide additional evidence on our mechanisms and test competing theories using repeated cross-sectional information on 393,607 individuals from 12 surveys in the years 1996–2014. This survey contains measures of local state avoidance based on paying local property taxes and working for the government. Additionally, the individual data allows us to reject alternative mechanisms based on population, age, education, migration, and assets. The village questionnaires from these surveys enable us to estimate the impacts on

[16]Sample question: 'How interested are you in politics?' or 'How often do you discuss politics with friends?' as both questions could be equally applicable to either supporters.

[17]Sample question: 'Now let's talk about the commune where you live. Generally speaking, do you think things in your commune today are going in the right direction, or do you think they are going in the wrong direction?'

[18]By taking into account the covariance between individual questions we obtain a more accurate measure than alternative standardizations that use an equally-weighted average. With the exception of our election results, we present standardized scores for all outcome categories where single regressions are significant. This procedure excludes results on competing hypotheses, which we present individually for disclosure.

[19]P-values adjusted for False Discovery Rates (FDR) are computed using the procedure outlined in Anderson (2008).

illegal rent extraction of deforestation and reject hypotheses based on differential state investment or public infrastructure.

School Census To further address competing theories based on differential education investments by the government, we use the school census 1997–2002. It contains school-level information about classes, teachers, students, and parents, which we aggregate to the commune level. Conceived to inform a policy change towards free education, this data additionally contains data on school income and measures of inequality at the commune level which we use to reject a Malthusian hypothesis based on income differences.

2.5 Empirical Strategy

The Khmer Rouge strategically placed labor camps around the country. This procedure might introduce an upward bias if these camps were placed in communes with larger initial dissent. Conversely, a downward bias arises if labor camps were built in areas with stronger support for the Khmer Rouge. Our identification strategy is thus based on temporal productivity differences during the Khmer Rouge that influence the size and location of labor camps. We argue that temporal productivity differences during the Khmer Rouge regime are uncorrelated to confounding factors and hence identify a causal effect on genocide intensity as well as preferences.

Our empirical strategy exploits the regime's desire to create an agricultural empire in the spirit of the great Angkor. Internal leadership documents reveal an extensive plan to increase productivity across Cambodia and sell the excess rice for foreign currency (Chandler et al., 1988). The central party ordered a three ton per hectare yield in all communes (Figure 4) and gave considerable freedom to provincial commanders who were instructed to "attack wherever the opportunities are greatest" and "attack in places where we are strong" (Chandler et al., 1988, p.20). This explicitly included using additional labor as failure was linked to "a lack of forces" (Chandler et al., 1988, p.15).

In order to determine which communes were more productive during the Khmer Rouge regime, we use temporal variation in rainfall to predict productivity. We combine historical precipitation data from Aphrodite at a 0.25 degree resolution covering the periods 1951–2007 with data from NOAA for the years 2002–2017 to obtain a long panel of precipitation in Cambodia. To account for commune specific variation in rainfall, we standardize the average rainfall of the three harvest seasons during the Khmer Rouge regime using the historical mean and standard deviation for each commune.[20] As excessive rain during the harvest season drowns the crop, we document a negative relationship between the standardized harvest season rainfall and rice yields using contemporaneous data in Figure 5. In a final step, we exploit the Khmer Rouge's decision to allow provincial leaders to allocate labor inside their provinces to relatively more productive areas. To match this strategy, we calculate the average productivity in every province and identify above average productive communes as our treated sample.[21] Formally,

$$\text{Productive during KR}_c = \mathbb{I}\left[\frac{\text{Rain during KR}_c - \mu_c}{\sigma_c} \leq \mu_p^{KR}\right] \tag{2.3}$$

where μ_c and σ_c is the communes historical mean and standard deviation used to standardize the rainfall in the commune during the Khmer Rouge period. Our binary treatment then defines productive communes

[20]The harvest season is defined as September, October, and the first two weeks of November according to Nesbitt (1997).

[21]Our procedure is a two-step standardization. First, we use the historical mean and standard deviation of each commune to determine how productive this commune was relative to its history. Then, we standardize again using the mean and standard deviation of all communes in a given province during the Khmer Rouge period, and define treated observations as those who were more productive relative to its own mean and the mean of its surrounding communes.

as experiencing rainfall below the province mean. Although loosing potentially interesting continuous variation, this procedure is closest to the leaderships plan which we aim to recreate and retains a considerable amount of variation across Cambodia (Figure 2b).[22]

We estimate the impact of being productive during the genocide using ordinary least squares, controlling for a second-order polynomial in latitude and longitude and pre-genocide commune characteristics:

$$Y_c = \beta \text{ Productive during KR}_c + \Gamma_c + X_c + \varepsilon_p \tag{2.4}$$

We report standard errors clustered at the province level as well as corrected for spatial correlation for all results to account for spatially correlated rainfall. β identifies the causal effect of state repression on preferences if the temporary production shock during the genocide is uncorrelated with observable characteristics at the time. In the last column of Table 1 we show that all pre-determined commune characteristics are uncorrelated with being relatively more productive during the Khmer Rouge regime. We document large p-values for all variables including the area of rice fields as an indicator of productivity, suggesting that underlying productiveness is uncorrelated to our instrument. Moreover, large p-values for population density and having a school as proxies for social capital suggest no pre-existing differences in preferences between productive and non-productive communes before the Khmer Rouge came to power. We thus argue that our production shock identifies a causal effect of state repression and test the hypothesis of exit and voice.

[22] In the appendix we document the robustness of our results to three additional definitions. First, we calculate the standard deviation within each province σ_p and define the continuous, within province productivity $R_{KR,c,p}$ as the standardized version of (2.3). Second, we use this continuous version and define more productive communes as $R_{KR,c,p} < -0.5$ and less productive as $R_{KR,c,p} > 0.5$. Third, we only use the standardized rainfall using the historical rainfall of the commune: $\frac{\text{Rain during KR}_c - \mu_c}{\sigma_c}$. The results are robust in all specifications of our productivity shock.

2.6 Results

State repression likely affects voters preferences for pluralism and democratic values as well as the perceived cost of dissent. Our model predicts that increasing democratic values strongly affect voting outcomes, while an increase in the perceived cost of dissent actively discourages voters to participate in local society. Only when both, preferences and cost of dissent, are affected by state repression do we expect more exit from civic society and more votes for the opposition in arguably free elections.

In this section, we present a sequence of results providing evidence for a causal effect of state repression on exit and voice in the Cambodian context. First, having established that our rain instrument predicts productivity, we show that more productive communes during 1975–1978 are not statistically different in any pre-genocide characteristic that could predict differential preferences. Second, we provide evidence that despite being similar prior to the genocide, more productive communes have been differentially targeted by the Khmer Rouge and have significantly higher rates of mass murder as measured by the death count in mass graves. Third, using data on recent election outcomes, we show how individuals from historically more productive communes use democratic institutions to voice their discontent at the voting booth. We corroborate this finding with results from various individual surveys that suggest differential preferences, more informedness about the political process, and exit from civic society in a fourth step. We conclude this section by arguing that the combination of results is consistent with our model of exit and voice. Voters have stronger preferences for democracy and are more informed about the political process, reducing the possibility of legal and illegal rent extraction for individuals and state officials alike.

State Repression To identify exogenous variation in state repression, we first establish that contemporaneous yields correlate negatively with excessive rainfall during the harvest season (Figure 5). We thus define a commune to be productive during the Khmer Rouge, if it experienced below average rainfall during the period 1975-1977 compared to other communes in the same province. In Table A.2 we show that our instrument for productivity increases standardized yields by 0.08 standard deviations or 0.213 tons per hectare in our preferred specification.[23] Importantly, the qualitative relationship is robust in all alternative specifications and shock definitions validating our identification of the Khmer Rouge's targeting of communes during the genocide.

In the next step, we provide evidence that more productive communes are not statistically different in terms of pre-genocide characteristics (Table 1). Our instrument is uncorrelated with important productivity characteristics such as the size of rice fields in 1970 or the FAO estimate of productivity 1960-1990 for low-input rain-fed rice.[24] Productiveness during the genocide is also uncorrelated with measures of market access such as the distance to Phnom Penh or roads, as well as the total bomb load of the American bombing campaign which could indicate areas of stronger support for the Khmer Rogue regime (Kiernan, 2008).

In a final step, we document the relationship between productiveness during the Khmer Rouge and indicators of violence in Table 2. Controlling for commune characteristics, a more productive commune has 387 more dead bodies in 8 more mass graves and a 62% higher probability of having a war memorial marking a Killing Field. Since increased violence indicates larger labor camps, these outcomes are indicative of increased state repression during the Khmer Rouge's reign. By the same token, these measures are highly correlated and to mitigate concerns of multiple hypotheses testing we standardize each violence measure and sum

[23]Our data suggest that in 1970, Cambodia had about 2.6 million hectares of rice which is corroborated by other sources that give a figure of 2.4 million hectares (`http://ricepedia.org/cambodia`, accessed April, 2018). Today, Cambodia has 3.1 million hectares of land producing 9.3 million tons of rice and a 0.2 ton increase in production is worth about 260 million USD in March 2018 prices.

[24]In addition, we report the balance test for individual characteristics we include in our individual regressions in the same Table 1. Except for interview circumstance (p=0.074), non are remotely significant with p-values above 0.253.

the standardized outcomes weighting each outcome by the inverse of the covariance matrix of the standadized outcomes (Anderson, 2008). The results in column (7) and (8), suggest that our instrument increases violence by 0.135 standard deviations.

The advantage of using precipitation as an instrument for productivity is that the same data can be used to validate the identification assumption. We argue that rainfall during the harvest season 1975–1977 affected the movement of people across Cambodia and ultimately the size and location of Killing Fields. Then, rainfall in any other period should be uncorrelated with measures of violence, except for chance. To test our identification, we employ two methods of randomization inference in Figure 6. Since our rainfall data only allows for 66 placebo harvest seasons, we first randomly allocate productiveness within each province. The point estimates from 1,000 repeated draws are shown in the left panel of Figure 6. Here, p-values for two-tailed tests range from 0.008 for the standardized violence measure to 0.051 for the probability of having a war memorial. Instead, using the 66 placebo years in the right panel of Figure 6, we obtain p-values in a range of 0.014 and 0.044 suggesting a highly significant first stage estimate.[25]

Having identified that productiveness during the Khmer Rouge is highly predictive of violence, we continue and establish the robustness of this result in different specifications and dependent variables in Table A.3. Here, we vary the definition of our shock in rows, and the dependent variable in columns. All measures, including per-capita or per-square-kilometer as well as log transformations of bodies and mass graves are robustly predicted by all shock definitions. Moreover, even though we show that our instrument is uncorrelated with population density, violence measures are likely to be positively correlated with population. Hence, we document that our point estimates remain unchanged if we omit the first, fifth, or tenth percentile of largest communes in 1970 (Table A.4).

[25]In an additional step, we verify in Table A.1 that growing season shocks are not correlated with our measures of violence.

In summary, our rainfall instrument strongly predicts productivity of rice fields and indicators of violence during the genocide. We document the robustness of this relationship using methods of randomization inference using random assignment of treatment, placebo estimates in any three-year period from 1951–2017, and varying shock definitions. As violence indicators are correlated with state repression, we argue that we have identified exogenous variation in state repression to test the implications of our model in terms of exit and voice.

Voice In our model, we hypothesize that increased preferences for pluralism lead to more people turning up to vote for the opposition party if this action is unobservable to the incumbent or society as a whole. We test this hypothesis using election results from the national election in 2013 and two communal elections in 2013 and 2017 (Table 3). In columns (1)-(6) of the upper panel, we document a strong relationship between our instrument, vote shares for the opposition (β: 4.766, s.e.: 1.049) and voter turnout (β: 2.939, s.e.: 1.292). We identify a similar, albeit smaller, effect in the communal elections suggesting that voters' party preferences have changed systematically in response to the memory of state repression.

Changed preferences for pluralism also affect the competitiveness of elections, as the likelihood of obtaining an absolute majority decreases significantly (column 8). To obtain a measure of competitiveness, we follow Besley et al. (2010) and show that since our instrument decreases the absolute win margin of the incumbent, both communal and national elections are more competitive in historically more productive communes (column 10).

To solidify the link to state repression during the Khmer Rouge era, we present the placebo estimates for all outcomes and elections in Figures 7-9. Vote shares for the opposition in the, arguable most free, national election 2013 are an extreme outlier with no other placebo or randomly selected treatment having a greater effect. The placebo p-values from the communal elections are slightly lower for two reasons. First, in the communal election of 2012, the opposition parties were not united and we sum their vote shares for comparability. During the second communal election cycle, CNRP and its representatives faced serious pressure and defamation likely decreasing their appeal for many voters.[26] For completeness, we report the robustness of our results in alternative shock definitions in Table A.5 and without large communes in Table A.6.

We conclude that state repression during the Khmer Rouge reign had a strong effect on voters' tendency to use voice as their political action. In line with our model, this result suggests that preferences for pluralism and democratic values were positively affected in response to a period of severe state repression.

Exit In our model, voters choice to exit from civic society depends on preferences and the expected cost of being detected as a dissident. The second hypothesis predicts that voters unambiguously reduce their participation in society if they face higher costs of being detected as a dissident. We address this 'exit'-hypothesis using two distinct data sets. First, we use a survey of the electorate from the Asia Foundation covering questions directly related to preferences, informedness, local civic participation, and trust. Second, we employ the Cambodian Socio-Economic Survey [CSES] and use the principle of revealed preferences which predicts changes in the decisions taken by individuals.

As multiple questions are targeted to elicit correlated information, we construct a standardized index accounting for the correlation between the variables in each category to avoid false discovery rates [FDR] due to multiple hypothesis testing. Following Cantoni et al. (2017), we addi-

[26]Pressure included, amongst others, suggesting that the opposition was bought by Vietnam, their leader is even more corrupt, and violent threats by Hun Sen himself: *"Words can cause war if the CPP loses patience and goes to your homes and burns down your homes."* `https://www.ft.com/content/3894454c-4681-11e7-8519-9f94ee97d996`, accessed April, 2018. For the 2018 national election, the main opposition party CNRP has been dissolved as the supreme court banned the party and more than 100 CNRP lawmakers from politics.

tionally provide the results on the individual questions in each category together with FDR adjusted p-values in Table A.7. We present our main estimates using the standardized scores on voter informedness, support for democratic values, civic participation and trust from the survey of the electorate in Table 4 and Figure 10. Voters in our sample are significantly more informed and show more support for democracy, corroborating our hypothesis that as preferences change, voters increasingly voice their discontent at the voting booth. As individuals develop stronger preferences for democracy, they inform themselves and vote for a more democratic party. The results on civic participation and trust suggest that voters additionally retreat from civic society. According to our hypothesis, this is unambiguously true if the memory of state repression increases the cost of openly voicing discontent.

Again, the results are highly robust using all placebo years (Figure 11) and various shock definitions (Table A.8). As the sample size is decreased to only 189 communes in 24 provinces, we report all estimates with zone fixed effects that split the provinces of Cambodia into four, roughly geographical, zones. However, the results using province fixed effects remain unchanged (Table A.9). Splitting the results by age group in Figure 10, it is clear that the average effect is driven by the cohort who survived the genocide, but important factors such as trust, informedness, and support for democracy are transmitted across generations.

In the second part, we use the CSES to estimate the effect of state repression on revealed preferences for paying local property taxes and working for the government (Table 5). Both are straightforward choices to identify 'exit' in general surveys as property taxes are easily observable and locally collected taxes and government employment directly measures daily interactions with the government. Across all generations, people living in communes that were historically more productive are paying less property tax in total, as a share of consumption, or per square-meter of housing. Standardizing these variables, the effect is about 0.03 standard deviations, with the effect being more pronounced for cohorts born after the genocide. Similarly, people are less likely to work for the government (column 10) and are more likely to be self employed (column 12).[27]

Combining the results from electoral results on voters voicing discontent under the anonymity of elections, and results from electoral surveys as well as the CSES on local civic participation, our findings suggest that voters use exit and voice as a result of state repression. In our model, these findings support our third hypothesis that state repression affects both preferences for pluralism and democratic values, as well as the expected cost of dissent. We do find direct evidence supporting this hypotheses as voters report a shift in preferences towards more democratic values in our election surveys. While shifting preferences can affect a voter's decision to participate in local civic society, the strong results we report on exit are indicative of an increase of an individual's perceived cost of dissent. In short, our findings suggest that the experience of political violence makes voters more convinced about the need for opposing views, but more cautious in expressing them.

Rent Extraction The combination of a more informed electorate with stronger democratic values and a low trust environment likely has implications for the ability to extract rents for politicians in developing countries (Pande, 2011). As voters become more informed about the political process, they likely elect more competent leaders who find their possibilities to legally extract rents restricted. Additionally, the low trust environment could affect the possibility to coordinate between individuals to illegally extract rents.

In developing countries, rents are most commonly extracted by legally or illegally selling timber from rain forests. As a case in point, Cambodia had the worlds third largest deforestation rate between 2000 and 2005

[27]The findings from revealed preferences using the CSES are again robust using all placebo years and various shock definitions (Figure 12 and Table A.10). In Table A.13 we use the 2008 census to show that while there is no detectable difference in the number of establishments, people in historically more productive communes tend to work in manufacturing, suggesting a sectoral shift away from agriculture.

with 29.4% of primary forest lost.[28] In Table 6, we combine data using the Hansen et al. (2013) deforestation measures between 2000 and 2014, data on land concessions for mining, and village level data on illegal logging and overuse from the CSES. Controlling for the size of forests in 1970 as calculated by the US Army maps, historically more productive communes have drastically lower rates of deforestation. The point estimate suggest about a 50% reduction in deforestation, a result highly significant in the placebo distribution (Figure 13a). Similarly, land concessions in affected communes decrease by 15% and illegal activities contributing to deforestation decrease by 12.6 and 15% in columns (4), (6), and (8).[29] Importantly, while the first two measures of rent extraction can be directly linked to a politician or party, the latter two suggest that the decreased trust observed in the election surveys affected extractive cooperation of citizens.

Jointly, the results suggest that individuals who suffered from state repression both exit civic society and voice their discontent if votes are undetectable. We establish in our model, that for this result both preferences for pluralism and democratic values, as well as the cost of dissent need to be affected by traumatic experiences during the genocide. Using electoral outcomes and individual surveys we suggest that the observed increased vote share for the opponent has its roots in more informed, more democratic, voters. The same voters, however, shun civic society and report lower levels of trust that appear to affect local cooperation, as we observe fewer illegal activities and decreased rates of deforestation. We observe that communes with larger vote shares for the more democratic opposition have significantly less deforestation and land concession, which suggests that politicians are less able to extract rents from office.

[28] `https://news.mongabay.com/2005/11/nigeria-has-worst-deforestation-rate-fao-revises-figures/`, accessed April, 2018.

[29] All measures are robust to alternative shock definitions and dropping large communes (Tables A.11 and A.12).

2.7 Alternative Hypotheses

Up until this point we presented evidence consistent with our third hypothesis of changed preference and cost of dissent. The traumatic experience of state repression during the Khmer Rouge's reign over Cambodia led people to exit civil society and voice their discontent at the voting booth if that action is unobservable. Arguably, our findings could be explained exclusively by changes in preferences which then allows for alternative hypotheses. In this section, we provide evidence against hypotheses based on differential survival rates, incomes, migration rates, and state investments influencing preferences.

Differential survival Differential survival rates between productive and non-productive communes could partially explain our results, since higher population density, younger voters, or more educated voters likely have stronger preferences for pluralism. Moreover, if these differences exist contemporaneously, preferences might have been different before the genocide, and are only imperfectly captured by our commune characteristics from the US Army maps. To address this potential channel, we combine census data from 1962, 1998, and 2008 with data from the CSES to provide evidence against systematic differences between productive and non-productive communes.

We digitized the 1962 census that allows us to compute the age distribution of Cambodia prior to the genocide.[30] In Figure 14 we contrast the resulting distribution with the computed distribution from the CSES in productive and non-productive communes. While it is apparent that many prime aged citizens lost their lives during the genocide (40-60 today), there appears to be no correlation with productive communes. In fact, testing the differences between the distributions in Figure 15, we doc-

[30]Unfortunately, the data is only available at the country level as the commune data only contains names and provinces which are nearly impossible to match to contemporaneous communes.

ument no systematic difference between productive and non-productive communes today.[31] This result is corroborated in Table 7, where census data from 1998 and 2008 reveals no significant differences between productive and non-productive communes in the upper panel. Population sizes across various cohorts and population density is uncorrelated to our rain instrument in both census waves. Using data from the CSES, we group ages into decades to increase power and again find no systematic differences in any age group. Finally, using the information on educational achievements in the bottom panel, we find no systematic differences in educational attainment between productive and non-productive communes.

While compared to the 1962 distribution of age in the population, prime aged citizens seem to have disproportionately suffered under the Khmer Rouge's regime. Census and survey data, however, show no contemporaneous differences in population, age, or education that could explain our results. While we reject the hypothesis of differential survival rates, we continue and investigate whether the observed death rates among the prime age population influenced the asset distribution in productive communes.

Income and wealth A classical Malthusian argument provides a second explanation for our findings on vote shares and turnout. People dying on the Killing Fields, or potentially even the sole presence of labor camps, likely changed the asset distribution in communes. Hence, people in productive communes likely have different consumption possibilities and assets than their non-productive counterparts.

In Table 8, we use the CSES survey data for household heads and test for systematic differences between communes. We find no differential endowments of house sizes or farm values and sizes between productive and non-productive communes. Further, households in productive communes

[31]The same result is seen in Figure A.1 when estimating the differences in distributions for men and women separately.

do not spend more on durable or non-durable consumption goods. Importantly, consumption of alcohol and tobacco products, which are likely correlated to post traumatic stress disorders, also show no significant differences. Here, as asset or consumption differences might be pre-existing or build up in the years after the genocide, we report the differences for the average in the upper panel, and the differences for the subset of residents that never moved from their current village. However, even for the subset of individuals that lived their entire lives under the shadow of the nearby killings, we do not find significant differences in consumption behavior or assets.

To provide additional evidence against a Malthusian argument, we extend our analysis of income differences towards income inequality in Table 9. If non-productive communes are more unequal, they might favor leaders that offer quick solutions for their economic problems. However, no measure of poverty or inequality is correlated to our rainfall instrument. Thus, we argue that differences in income, assets, or inequality are unlikely to explain our findings.

Migration Using information on migration status from the CSES we are able to identify differential population movements just after the genocide. If people would systematically move to either type of communes, such migration streams could be indicative of differential income expectations which could explain our results.

In Table 10 we focus on individuals who were alive during the Khmer Rouge reign and provide evidence against differential migration rates. People do not systematically migrate out of communes with labor camps in 1979 or 1980 (columns 2 and 4). Moreover, for the 6% of individuals who willingly respond to have returned in 1979 after having been displaced by the Khmer Rouge, there exists no differential movement between communes.

By confining our sample to the survey from 1996, we can also identify the district of origin for every migrant in our sample. By collapsing our instrument to the district level, we show in Figure A.2 that people who returned just after the genocide in 1979 or 1980, disproportionately arrive from districts that were more productive during the genocide. While in later years between 40 and 50% came from non-productive districts, more than 80% of migrants in 1979 came from districts that were targeted by the Khmer Rouge.

In short, migration destinations in 1979 are uncorrelated to our rain instrument suggesting that the Khmer Rouge forced people from all over Cambodia to migrate to labor camps. Notably, in our sample of adults, 20% of the people alive during the genocide moved in 1979 and then stayed in the commune and 42% did not move during the genocide. Hence, while we observe high migration rates two facts stand out. First, 62% of the population did not move since 1979 and second, 41% of those who migrated between 1979 and today did so immediately after the genocide. Moreover, we show that the source districts of people who migrated in 1979 are disproportionately more likely to have been productive during the genocide. Taken together, these findings provide additional evidence for our identification assumption, as well as evidence against the hypothesis of differential migration.

State investment The last alternative hypothesis we explore is the possibility of differential investment into communes. As governments distribute funds, communes benefit differentially from programs that increase access to food, capital, or health. If, in return, voters cast their ballot for the incumbent, we should expect to see such investments in communes with a larger CPP share. However, as these investments could also be used to persuade potential CNRP voters, both positive and negative correlations with productivity during the Khmer Rouge era are possible.

Using data from village survey part of the CSES, we show no significant correlation between productiveness and market access or public infrastructure (upper panel of Table 11). Additionally, we use the EMIS data from 1997-2002 to capture potential education programs, and show that neither school characteristics nor class characteristics are correlated to productiveness during the genocide. We complement the evidence with null results using recent data on night time lights, the number of food markets, health centers, and access to information in Table A.14.

In this section we have explored, and provided evidence against, four competing hypotheses that could explain our results. We have shown that while the population of Cambodia as a whole was affected by the killings of more than 1.5 million people during the genocide, citizens were not differentially affected as population, education, income, assets, and migration is balanced across productive and non-productive communes today. Moreover, we find no evidence for differential investments into communes that would bias our estimates. Reviewing our evidence in its entirety we conclude that our model of changing preferences and costs of dissent is in line with our empirical findings.

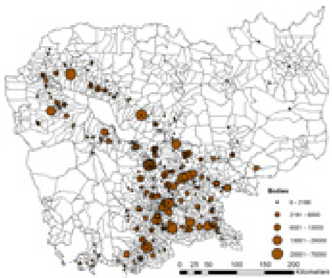
2.8 Conclusion

We show that state repression makes politics less personal and more competitive. Using evidence from history's most severe episode of political violence, the genocide in Cambodia under the Khmer Rouge, we find that state coercion leads to more votes in favor of the opposition over the authoritarian incumbent and increased support for democratic principles 4 decades after the genocide. At the same time, citizens become more cautious in their interactions with the local community as captured by lower participation in community organizations and less trust. While the results are more pronounced for people that were alive under the Khmer Rouge, effects persist across generations suggesting that the legacy of political violence can have a long-term impact on society. In addition, we also provide evidence that the changes in people's political preferences and behavior are driven by experiences of state repression rather than altered demographics of the survivors or direct economic effects.

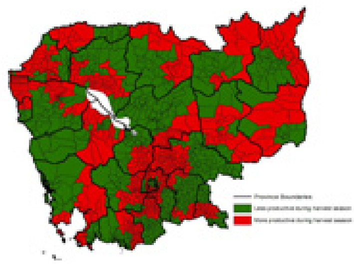
The results are relevant for the policy debate on democratic development, contributing to our understanding of political participation in post-conflict societies where citizens still live under the threat of political violence. Even in authoritarian states, such as Cambodia, elections matter as a source of legitimacy and corrective feedback (Brownlee, 2007; Magaloni, 2006) or as a way to allow for a credible power sharing among the elites (Bidner et al., 2015). Our findings also open up for additional questions. First, do prisons or labor camps induce similar effects on preferences and behavior in other contexts, such as Nazi Germany's concentration camps or the US WW2 internment of Japanese Americans? Second, given our findings on the link between state repression and political competition, more research in needed to understand the implications for the theory of electoral competition in political economics.

2.9 Figures and Tables

Figure 2: Killing fields in Cambodia

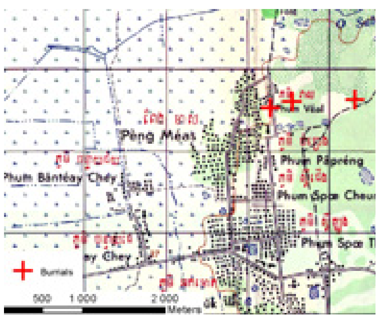


(a) Cambodia's Killing Fields, 309 sites with mean no. of killed: 3,154

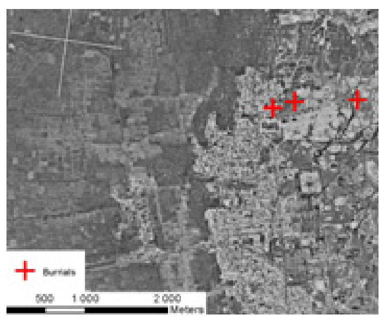


(b) More and less productive communes during the genocide

Figure 3: Pre-genocide covariates from US Army L7016



(a) Example of L7016 map



(b) Aerial photography 1976

Figure 4: Production plans of the Khmer Rouge leadership

TABLE 3

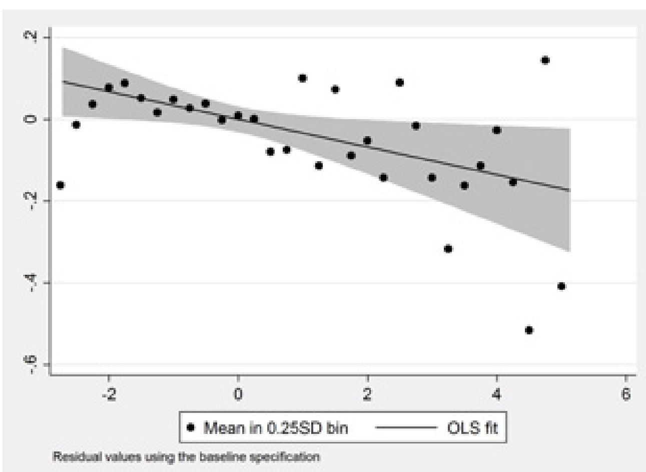
PLAN FOR RICE PRODUCTION THROUGHOUT THE COUNTRY DURING THE PERIOD 1977 - 1980

Zone and Region	1977	1978	1979	1980	Total For Four Years
1. NW	1,620,000T	1,900,000T	2,250,000T	2,600,000T	8,370,000T
2. East	1,290,000T	1,410,000T	1,510,000T	1,620,000T	5,830,000T
3. SW	1,140,000T	1,210,000T	1,320,000T	1,440,000T	5,110,000T
4. North	695,000T	758,000T	935,000T	912,000T	3,200,000T
5. West	432,000T	450,000T	480,000T	510,000T	1,872,000T
6. NE	73,000T	78,000T	84,000T	90,000T	335,000T
7. Region 106	306,000T	336,000T	366,000T	384,000T	1,392,000T
8. Region 103	42,000T	48,000T	54,000T	60,000T	204,000T
9. Centre Armed Forces	18,000T	24,000T	30,000T	36,000T	108,000T
10. Zone Armed Forces	39,000T	54,000T	66,000T	90,000T	249,000T
Total:	5,555,000T	6,268,000T	6,995,000T	7,742,000T	26,560,000T[a]

[a] Total rice produced. Total production for fields harvested twice per year is figured as 6 tons per hectare; ordinary fields harvested once per year is estimated at 3 tons per hectare.

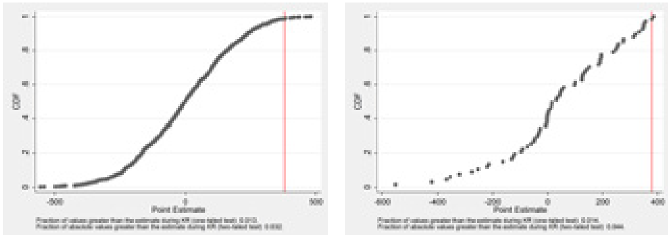
Example of a production plans across different regions of Cambodia (Chandler et al., 1988).

Figure 5: Rice yields

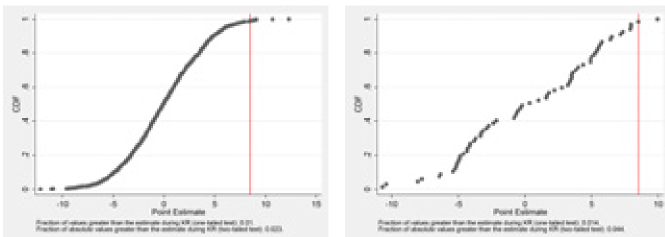


Rice yields as a function of standardized rainfall during the harvest season. Data taken from the Cambodian socio economic survey 1996–2014. More rain is associated with lower yields as it drowns the rice. 95% confidence intervals shown. Province fixed effects and a second-degree polynomial in latitude in longitude included in the regression. Commune characteristics included and defined in Table 1.

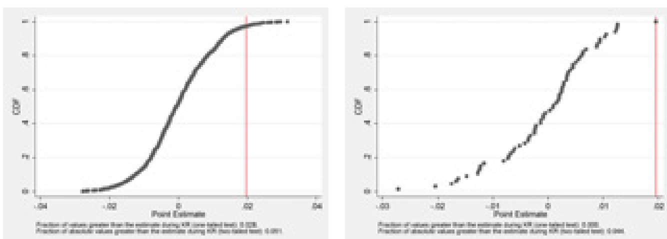
Figure 6: Placebo estimates for violence



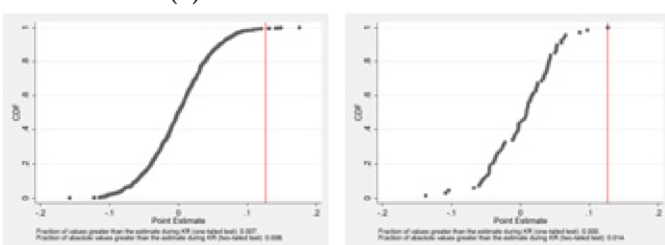
(a) # Bodies in commune



(b) # Mass graves in commune



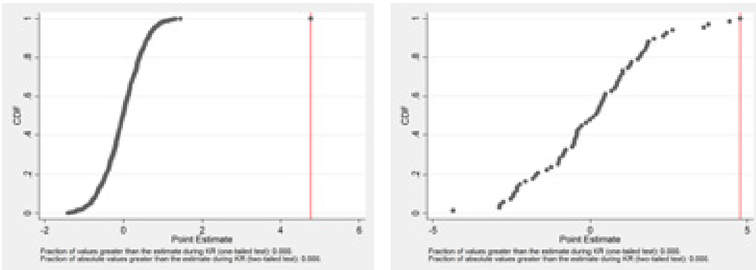
(c) War memorial in commune



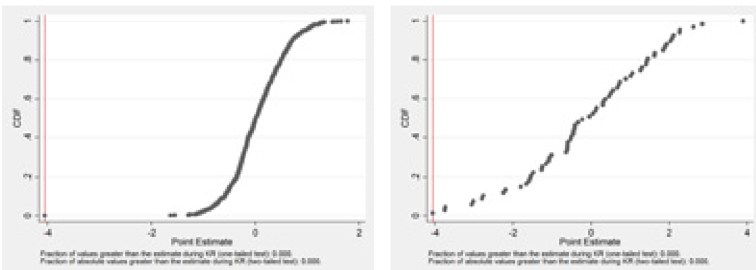
(d) Standardized violence

The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Randomization inference (left) and placebo seasons (right) for the main violence indicators. The randomization procedure assigns 50% of the communes within a province to treatment using 1,000 draws. In the placebo estimations (right), treatment is assigned based on the within province productivity in the harvest season in all three-year windows from 1951 until 2017. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions.

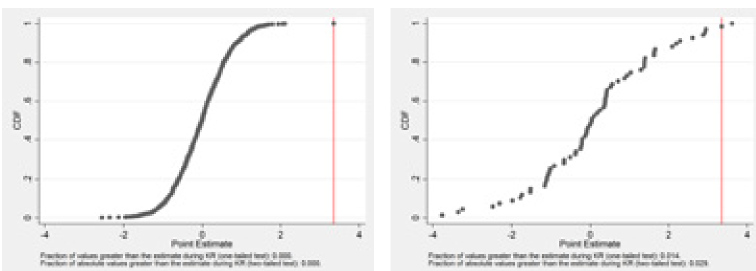
Figure 7: Placebo estimates for political mobilization: National election



(a) Vote share opposition in commune (CNRP)



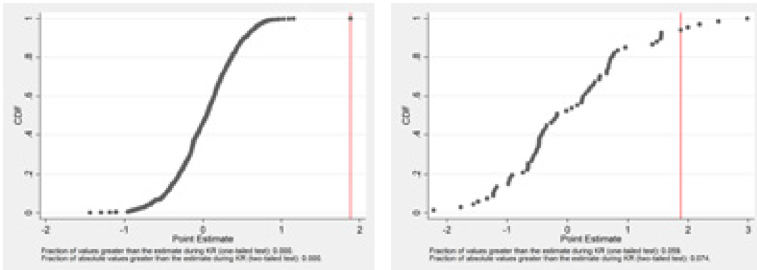
(b) Vote share incumbent in commune (CPP)



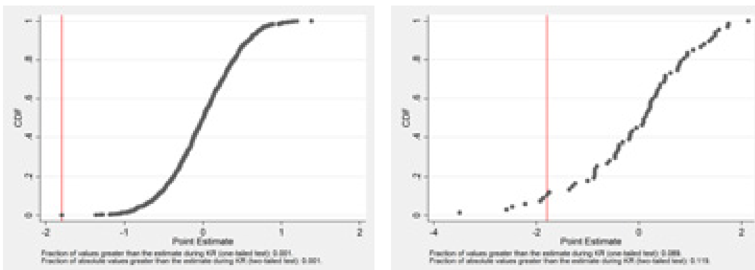
(c) Turnout in commune

The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Randomization inference (left) and placebo seasons (right) for the main violence indicators. The randomization procedure assigns 50% of the communes within a province to treatment using 1,000 draws. In the placebo estimations (right), treatment is assigned based on the within province productivity in the harvest season in all three-year windows from 1951 until 2017. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions.

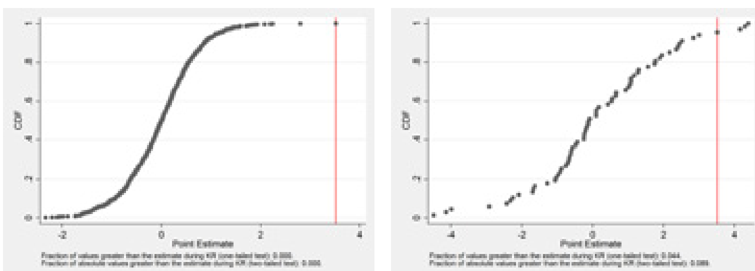
Figure 8: Placebo estimates for political mobilization: Communal elections



(a) Vote share opposition in commune (CNRP)



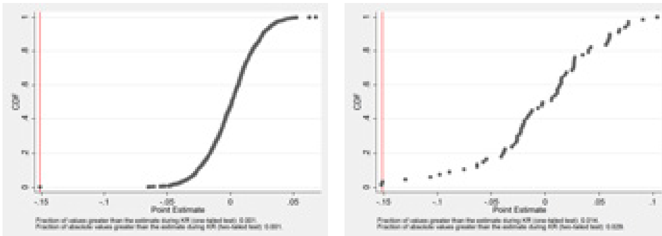
(b) Vote share incumbent in commune (CPP)



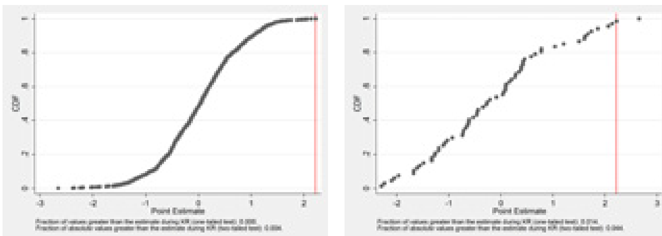
(c) Turnout in commune

The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Randomization inference (left) and placebo seasons (right) for the main violence indicators. The randomization procedure assigns 50% of the communes within a province to treatment using 1,000 draws. In the placebo estimations (right), treatment is assigned based on the within province productivity in the harvest season in all three-year windows from 1951 until 2017. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions.

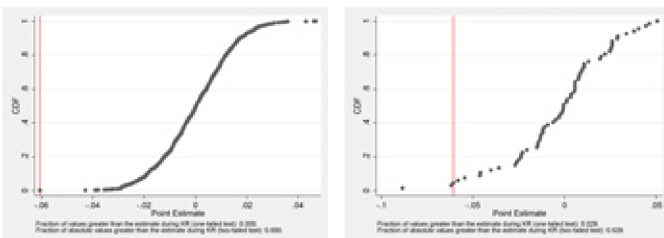
Figure 9: Placebo estimates for political competition



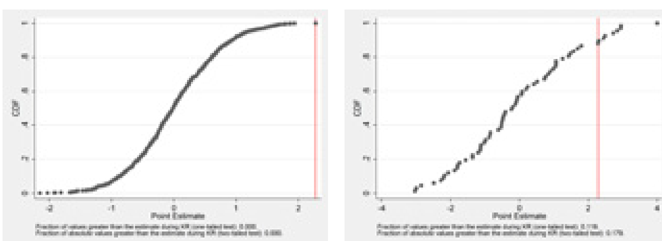
(a) National election: Probability of CPP having a majority in commune



(b) National election: Margin -|CPP-CNRP|



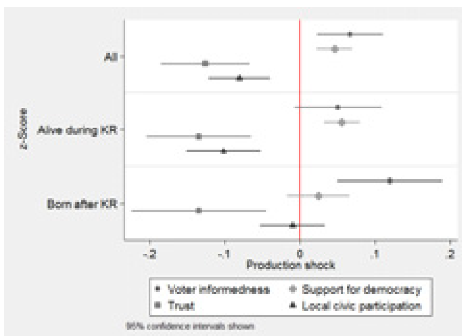
(c) Commune election: Probability of CPP having a majority in commune



(d) Commune election: Margin -|CPP-CNRP|

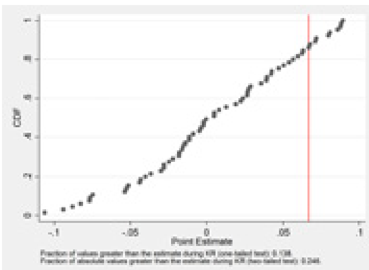
The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Randomization inference (left) and placebo seasons (right) for the main violence indicators. The randomization procedure assigns 50% of the communes within a province to treatment using 1,000 draws. In the placebo estimations (right), treatment is assigned based on the within province productivity in the harvest season in all three-year windows from 1951 until 2017. 'Margin -|CPP-CNRP|' is calculated as the vote share of CPP minus CNRP and a variation of the competitiveness measure by Besley et al. (2010). Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions.

Figure 10: Exit and voice

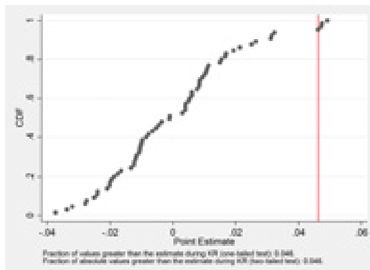


Standardized scores on voter informedness, support for democracy, local civic participation, and trust from the Asia Foundation 2003 and 2013 survey. Zone fixed effects and a second-degree polynomial in latitude in longitude included in all communes. 95% confidence intervals shown.

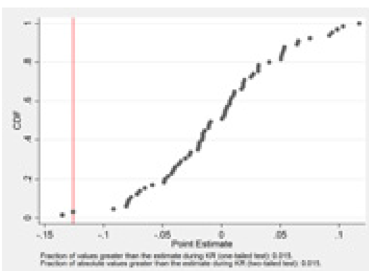
Figure 11: Placebo estimates for exit and voice



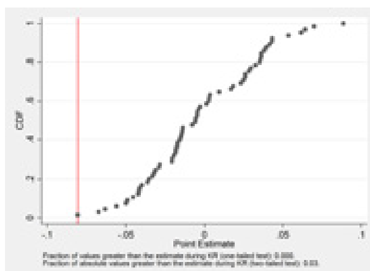
(a) Voter informedness



(b) Support for democracy



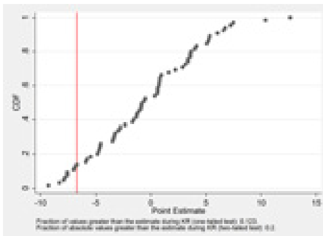
(c) Trust



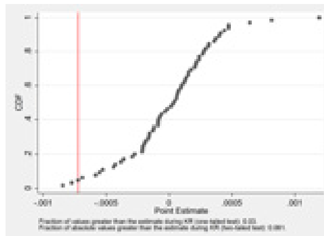
(d) Civic participation

The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Placebo estimations for the average effects. Treatment is assigned based on the within province productivity in the harvest seasons in all three year windows from 1951 until 2017. Zone fixed effects and a second-degree polynomial in latitude in longitude included in all communes. Commune characteristics included and defined in Table 1.

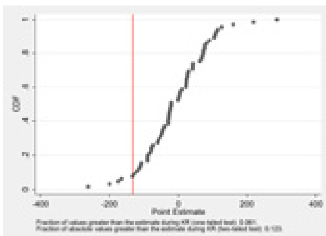
Figure 12: Placebo estimates: Paying property taxes and government employment



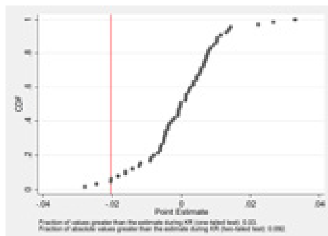
(a) Property taxes paid



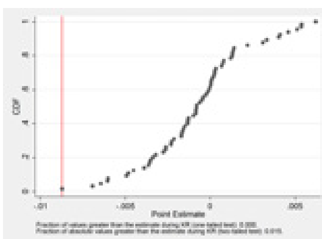
(b) Share of property tax



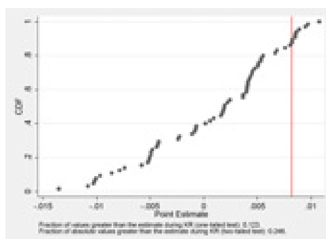
(c) Property tax per sqm



(d) Standardized score



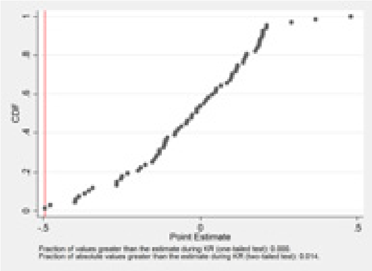
(e) Government employee



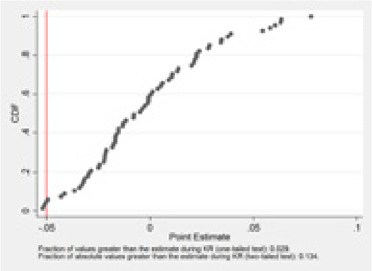
(f) Self employed

The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Placebo estimations for the average effects. Treatment is assigned based on the within province productivity in the harvest seasons in all three year windows from 1951 until 2017. Source for all variables: Cambodia socio economic survey 1996–2014. 'Share of property tax' is defined as the amount of property tax paid, relative to all non-food expenditures. 'Property tax per sqm' is defined as the amount of property tax paid, relative to the floor area of the individuals home. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics included and defined in Table 1.

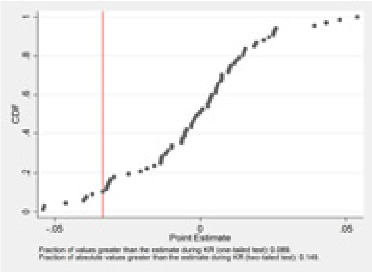
Figure 13: Placebo estimates: Deforestation and illegal land use



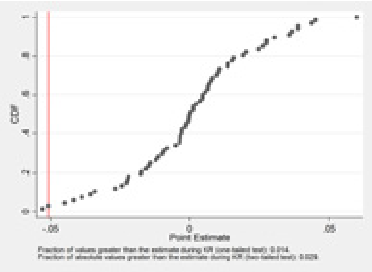
(a) Forest loss



(b) Land concessions



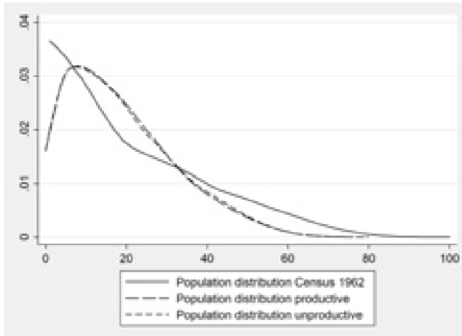
(c) Illegal logging



(d) Illegal overuse

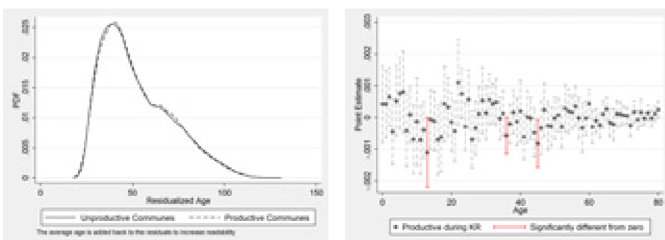
The graphs show the main effect of the production shock in the harvest seasons during the Khmer Rouge reign, compared to the cumulative distribution of estimates of a production shock in placebo years. The line indicates the estimated coefficient during the Khmer Rouge. Under every graph two statistics indicating the p-value of a one-sided and two-sided test are presented. Placebo estimations for the average effects. Treatment is assigned based on the within province productivity in the harvest seasons in all three year windows from 1951 until 2017. Source for all variables: Cambodia socio economic survey 1996–2014. 'Forest loss' is defined as the square kilometers of forest lost between 2000–2014 and provided by Hansen. Source for all other variables: The village data set from the Cambodia socio economic survey 1996–2014. 'Land concessions' is defined as one if a commune sold land for mining of forest operations. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics included and defined in Table 1.

Figure 14: Survival composition: Age



The distribution of age in the 1962 census (solid line) and the Cambodian socio economic survey 1996–2014. The dashed line represents the age distribution within communes that where more productive during the Khmer Rouge and the dotted line those that were less productive.

Figure 15: Distributional effects: Age



The distribution of age in the Cambodian socio economic survey 1996–2014, separated by the productiveness of the commune during the Khmer Rouge regime. Histogram on the residualized distributions (left) and point estimates on the difference between the distributions for every age between 0–80. Differences based on whether the commune was productive during the genocide. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics included and defined in Table 1.

Table 1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-productive communes		Productive communes			Exogeneity test		
	Mean	S.D.	Mean	S.D.	β	s.e.	T-Stat	p-value
Violence indicators:								
#Bodies in commune	407.873	2724.575	792.152	4514.115				
#Mass graves in commune	7.086	46.580	16.237	97.032				
War memorial in commune	0.035	0.183	0.053	0.224				
Political mobilization:								
Vote share for CNRP, national election	37.512	15.710	41.814	16.315				
Vote share for CNRP, commune election	33.683	12.887	35.808	13.438				
Vote share for CPP, national election	54.782	14.639	50.852	14.967				
Vote share for CPP, commune election	61.664	14.304	59.405	14.852				
Turnout, national election	77.274	18.361	80.430	17.012				
Turnout, commune election	75.427	20.175	78.799	18.768				
CPP≥50%, national election	0.593	0.492	0.468	0.499				
CPP≥50%, commune election	0.782	0.413	0.708	0.455				
Margin: -|CPP-CNRP|, national election	−27.889	20.337	−25.271	20.001				
Margin: -|CPP-CNRP|, commune election	−31.053	22.807	−28.509	22.705				
Local state avoidance:								
Property taxes paid	54.737	306.353	22.398	150.791				
Share of property tax	0.005	0.020	0.002	0.011				
Property tax per sqm of housing	906.715	12980.324	351.728	2276.594				
Working for the government	0.098	0.298	0.065	0.247				
Self employment	0.216	0.412	0.248	0.432				
Deforestation:								
log Forest loss	3.846	3.104	3.093	2.959				
Land concession	0.317	0.466	0.244	0.430				
Illegal logging	0.252	0.434	0.274	0.446				
Illegal overuse	0.304	0.460	0.326	0.469				
Commune characteristics:								
Commune with school	0.670	0.471	0.705	0.456	0.026	0.025	1.046	0.296
Commune with telephone	0.004	0.061	0.006	0.078	0.002	0.003	0.579	0.563
Commune with commune office	0.383	0.486	0.386	0.487	0.001	0.029	0.048	0.961
Commune with post office	0.017	0.131	0.016	0.125	−0.003	0.005	−0.529	0.597
log Population density	5.189	1.521	5.096	1.576	−0.024	0.133	−0.182	0.856
log Rice field area	5.691	2.841	6.239	2.430	0.392	0.349	1.123	0.261
log Area partially inundated	3.250	3.246	2.894	3.085	−0.125	0.247	−0.504	0.614
log Area covered by dense forests	4.081	3.941	3.911	3.594	−0.281	0.469	−0.599	0.549
log Commune area	3.864	1.619	3.814	1.152	−0.134	0.114	−1.173	0.241
log Distance to Phnom Penh	4.448	1.450	4.549	0.937	−0.067	0.069	−0.967	0.334
log Distance to closest road	0.397	1.416	0.387	1.465	0.032	0.116	0.272	0.786
log Distance to province capital	2.440	2.851	2.810	2.125	−0.003	0.103	−0.032	0.974
log Bomb load 1965-1973	4.932	3.356	4.630	3.188	0.095	0.236	0.402	0.688
log Potential yields (FAO, 1960-1990)	1.013	0.014	1.015	0.013	0.000	0.000	0.850	0.395
Individual characteristics, Asia foundation 2003 and 2013:								
Ethnicity	0.038	0.335	0.088	0.592	0.041	0.042	0.968	0.344
Year of birth	1969.798	15.256	1970.949	14.963	0.288	0.620	0.464	0.647
Male	0.488	0.500	0.501	0.500	0.013	0.017	0.779	0.445
Education	2.369	1.298	2.244	1.192	−0.131	0.112	−1.173	0.253
Income	2.990	1.839	3.003	1.829	−0.085	0.128	−0.664	0.513
Interview circumstance	1.125	1.220	1.163	1.198	0.111	0.059	1.874	0.074
Urbanity	0.497	0.500	0.506	0.500	0.009	0.062	0.146	0.885
Brick House	0.892	0.311	0.887	0.317	0.011	0.023	0.480	0.636
Individual characteristics, Cambodian socio economic survey 1996–2014:								
Year of birth	1979.462	19.389	1980.076	19.591	−0.058	0.208	−0.277	0.782
Male	0.480	0.500	0.481	0.500	−0.001	0.002	−0.432	0.666
Urbanity	0.338	0.473	0.230	0.421	0.021	0.055	0.377	0.706
Years of education	5.532	5.337	5.010	5.142	−0.033	0.070	−0.047	0.635

Data on violence taken from the Cambodian Genocide Project. Data on Political mobilization taken from the national election offices in Cambodia. Commune characteristics are taken from the L7016 army maps covering Cambodia in 1970 and digitized by the authors if not otherwise noted. 'log Bomb load' taken from the Cambodian Genocide Project. 'Potential yields' are for low input rain fed rice from 1960–1990 and taken from the FAO. For deforestation, 'log Forest loss' is defined as the hectares of forest lost between 2000 and 2014, as calculated by Hansen et al. (2013), and 'land concessions' is a binary variable indicating whether any area in the communes was sold under a land concessions. The remaining variables are taken from the village questionnaires from the Cambodian socio-economic survey 1996–2014. Individual characteristics obtained by the indicated surveys and are included into regressions as fixed effects. Interview circumstance indicates whether the respondent was alone, with family, or a local official when answering the questionnaire.

Table 2: Violence incidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	#Bodies		#Mass graves		War Memorial		Standardized violence	
Productive during KR	377.914***	387.276***	8.501***	8.038***	0.020**	0.022***	0.127***	0.135***
	(171.222)	(150.958)	(3.529)	(3.265)	(0.011)	(0.010)	(0.045)	(0.043)
	[141.584]	[138.934]	[2.909]	[2.856]	[0.008]	[0.008]	[0.033]	[0.031]
Commune characteristics		Yes		Yes		Yes		Yes
Mean non-productive	407.873	407.873	7.094	7.094	0.035	0.035		
Observations	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621

First stage results on how productivity during the Khmer Rouge influenced violence in a commune. 'Standardized violence' is the standardized index of '#Bodies', '#Mass graves, and 'War memorial', taking into account the covariance between these variables. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Political mobilization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Political mobilization						Political competition			
	Vote share CNRP (Opposition)		Vote share CPP (Incumbent)		Turnout		Absolute majority for CPP		Margin -|CPP-CNRP|	
	National election in 2013									
Productive during KR	4.766***	4.890***	−4.054***	−4.218***	3.351**	2.939**	−0.151***	−0.155***	2.223*	1.754
	(1.530)	(0.879)	(1.463)	(0.876)	(1.767)	(1.474)	(0.055)	(0.038)	(1.276)	(1.147)
	[1.049]	[0.600]	[0.979]	[0.614]	[1.583]	[1.292]	[0.031]	[0.026]	[1.341]	[1.305]
	Communal elections in 2012 and 2017									
Productive during KR	1.882***	2.100***	−1.794**	−2.016***	3.514*	3.080**	−0.061**	−0.069***	2.280**	2.383***
	(1.053)	(0.680)	(1.108)	(0.811)	(2.157)	(1.686)	(0.032)	(0.024)	(1.484)	(1.048)
	[0.720]	[0.493]	[0.724]	[0.565]	[2.077]	[1.478]	[0.024]	[0.019]	[1.040]	[0.851]
Commune characteristics		Yes		Yes		Yes		Yes		Yes
Observations national election	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621
Mean national election	37.512	37.512	54.782	54.782	77.274	77.274	0.593	0.593	−27.889	−27.889
Observations commune elections	3,230	3,230	3,230	3,230	3,230	3,230	3,230	3,230	3,230	3,230
Mean commune elections	33.683	33.683	61.664	61.664	75.427	75.427	0.782	0.782	−31.053	−31.053

Every cell constitutes a separate regression of the instrument on the dependent variable in the header. Data taken from the official election results. 'Vote share CNRP' in the communal elections 2012 is calculated as the combined votes of the 'Sam Rainsy Party' and the 'Human Rights Party'. 'Turnout' is calculated using the electorate information from the national elections in 2013. 'Vote share CPP $\geq 50\%$' is a binary variable indicating an absolute majority for the incumbent party CPP. 'Margin -|CPP-CNRP|' is calculated as the vote share of CPP minus CNRP and a variation of the competitiveness measure by Besley et al. (2010). Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Exit and Voice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Voter informedness		Support for democracy		Local civic participation		Trust	
Average effect	0.060^{**}	0.067^{***}	0.043^{***}	0.046^{***}	-0.081^{***}	-0.081^{***}	-0.131^{***}	-0.126^{***}
	0.027	0.027	0.008	0.009	0.025	0.022	0.034	0.035
	[0.025]	[0.023]	[0.012]	[0.012]	[0.023]	[0.021]	[0.030]	[0.030]
Alive during KR	0.039	0.049^{*}	0.050^{***}	0.054^{***}	-0.096^{***}	-0.099^{***}	-0.128^{***}	-0.131^{***}
	0.031	0.032	0.009	0.009	0.026	0.023	0.037	0.036
	[0.028]	[0.028]	[0.012]	[0.012]	[0.026]	[0.024]	[0.034]	[0.033]
Born After KR	0.108^{***}	0.110^{***}	0.022	0.022	-0.012	-0.009	-0.124^{***}	-0.123^{***}
	0.037	0.033	0.014	0.014	0.034	0.032	0.054	0.055
	[0.038]	[0.033]	[0.018]	[0.019]	[0.022]	[0.021]	[0.045]	[0.045]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes

Every cell constitutes a separate regression of the instrument on the dependent variable in the header using individual level data. The row names define the sample used based on whether the year of birth is before or after 1978. Results using questions from the Asia Foundation 2003 and 2013. Individual results per category show in Table 4. Individual covariates are ethnicity, year of birth, education, income, interview circumstance, rural status and housing status. Zone fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Zone fixed effects sort provinces in four zones to improve power. Commune characteristics are defined in Table 1.Results with province fixed effects shown in Table A.9. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Local state avoidance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Property taxes paid		Share of property tax		Property tax per sqm of housing		Standardized z-score		Working for the government		Self employment	
Average Effect	-8.637^{**}	-6.902^{**}	-0.001^{*}	-0.001^{*}	-156.334^{***}	-133.906^{***}	-0.024^{**}	-0.021^{**}	-0.008^{***}	-0.008^{***}	0.006	0.007^{*}
	(4.325)	(3.390)	(0.000)	(0.000)	(77.477)	(65.947)	(0.012)	(0.009)	(0.002)	(0.003)	(0.005)	(0.005)
	[3.541]	[2.699]	[0.000]	[0.000]	[52.485]	[41.728]	[0.011]	[0.011]	[0.002]	[0.002]	[0.004]	[0.004]
Alive during KR	-7.682^{*}	−4.665	-0.001^{*}	−0.001	-159.334^{***}	-118.912^{***}	-0.023^{**}	-0.018^{*}	-0.008^{***}	-0.007^{**}	0.008^{**}	0.008^{**}
	(4.568)	(3.679)	(0.000)	(0.000)	(75.710)	(60.812)	(0.011)	(0.008)	(0.003)	(0.004)	(0.006)	(0.006)
	[4.111]	[3.233]	[0.000]	[0.000]	[57.099]	[42.413]	[0.011]	[0.010]	[0.002]	[0.003]	[0.004]	[0.004]
Born After KR	-6.192^{*}	-4.669^{*}	−0.000	−0.000	−47.867	−35.720	-0.051^{*}	-0.044^{*}	-0.007^{***}	−0.004	−0.001	0.000
	(3.218)	(2.679)	(0.000)	(0.000)	(61.250)	(55.796)	(0.026)	(0.022)	(0.002)	(0.003)	(0.004)	(0.004)
	[3.211]	[2.613]	[0.000]	[0.000]	[66.946]	[60.245]	[0.027]	[0.024]	[0.002]	[0.003]	[0.003]	[0.003]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes		Yes		Yes
Observations average	19,671	16,513	19,671	16,513	19,671	16,513	19,671	16,513	136,027	118,849	136,027	118,849
Mean control average	61.386	61.386	0.005	0.005	974.410	974.410			0.098	0.098	0.216	0.216
Observations alive during KR	14,916	11,758	14,916	11,758	14,916	11,758	14,916	11,758	85,635	68,457	85,635	68,457
Mean control alive during KR	61.039	59.020	0.005	0.005	961.806	930.878			0.126	0.126	0.302	0.302
Observations born after KR	4,755	4,755	4,755	4,755	4,755	4,755	4,755	4,755	47,996	47,996	47,996	47,996
Mean control born after KR	63.861	63.861	0.005	0.005	1012.070	1012.070			0.050	0.050	0.063	0.063

Every cell constitutes a separate regression of the instrument on the dependent variable in the header using household level data. The row names define the sample used based on whether the year of birth is before or after 1978. Results using expenditure data from the Cambodian socio-economic survey 1996–2014. 'Property taxes paid' denotes the amount of property tax paid in Cambodian Riel.'Share of property tax' is defined as the share of nonfood consumption in the last 12 months of the household. 'Property tax per sqm of housing' relates the property taxes to the size of the households house. The standardized z-score in Columns (7) and (8) combines the previous variables, accounting for the covariance between the variables. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Included individual covariates are: Urbanity and highest education. Year of birth and male fixed effects included in all estimations. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Rent extraction: Deforestation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	log Forest loss		Land concessions		Illegal logging		Illegal overuse	
Productive during KR	-0.827^{***}	-0.501^{***}	-0.074^{***}	-0.048^{*}	-0.037^{*}	-0.032^{**}	-0.051^{***}	-0.051^{***}
	(0.207)	(0.151)	(0.029)	(0.026)	(0.023)	(0.016)	(0.027)	(0.017)
	[0.178]	[0.131]	[0.025]	[0.027]	[0.019]	[0.016]	[0.018]	[0.016]
Commune characteristics		Yes		Yes		Yes		Yes
Year fixed effects					Yes	Yes	Yes	Yes
Mean non-productive	3.846	3.846	0.317	0.317	0.252	0.252	0.304	0.304
Observations	1,621	1,621	1,621	1,621	3,027	3,027	3,027	3,027

Commune level results using various data sources. log Forest loss is defined as the hectares of forest lost between 2000 and 2014, as calculated by Hansen et al. (2013). 'Land concessions' is a binary variable indicating whether any area in the communes was sold under a land concessions. The remaining variables are taken from the village questionnaires from the Cambodian socio-economic survey 1996–2014. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Population, age, and education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Population: Census 1998				Population: Census 2008			
	log Population ≤ 15	log Population $\in [10,19]$	log Population $\in [15,64]$	log Population density	log Population ≤ 15	log Population $\in [10,19]$	log Population $\in [15,64]$	log Population density
Productive during KR	0.013	0.002	0.001	0.002	0.038	0.027	0.026	0.031
	(0.036)	(0.040)	(0.042)	(0.039)	(0.038)	(0.040)	(0.044)	(0.042)
	[0.031]	[0.033]	[0.034]	[0.034]	[0.035]	[0.036]	[0.039]	[0.037]
	Age: Cambodia socio economic survey 1996–2014							
	Age $\in [0,9]$	Age $\in [10,19]$	Age $\in [20,29]$	Age $\in [30,39]$	Age $\in [40,49]$	Age $\in [50,59]$	Age $\in [60,69]$	Age $\in [70,79]$
Productive during KR	0.002	-0.003	0.001	0.001	-0.002	0.000	-0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
	Education: Cambodia socio economic survey 1996–2014							
	Can read	Can write	Speaking English	Speaking French	Lower secondary school	Upper secondary school	Bachelor	Years of education
Productive during KR	0.003	0.004	-0.004	-0.001	0.000	-0.001	-0.003^{*}	0.003
	(0.007)	(0.006)	(0.005)	(0.001)	(0.001)	(0.002)	(0.002)	(0.070)
	[0.003]	[0.004]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.033]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations population	1,570	1,570	1,570	1,570	1,614	1,614	1,614	1,614
Mean population	7.822	7.307	8.039	4.870	7.716	7.378	8.276	4.906
Observations age	393,591	393,591	393,591	393,591	393,591	393,591	393,591	393,591
Mean age	0.208	0.237	0.181	0.128	0.103	0.074	0.042	0.020
Observations education	266,586	266,600	347,794	347,794	289,062	289,062	289,062	289,062
Mean education	0.710	0.736	0.065	0.019	0.017	0.027	0.020	5.762

Data taken on population taken from commune level censuses in 1998 and 2008. Remaining data taken from the Cambodian socio economic survey 1996–2014. Regressions on age feature a binary variable if the age of the individual is within the indicated interval as the dependent variable. Point estimates then reflect differences in the distributions of productive and non-productive communes. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Assets and consumption

	(1) Rooms p.c.	(2) log Farm value	(3) log Size of farm	(4) log Consumption p.c.	(5) log Food expenditure p.c.	(6) log Non-food expenditure p.c.	(7) log Expenditure p.c.	(8) log Alcohol & tobacco
Average	−0.001	0.069	−0.050	0.003	0.011	0.011	0.007	−0.054
	(0.004)	(0.271)	(0.152)	(0.019)	(0.016)	(0.028)	(0.018)	(0.093)
	[0.004]	[0.220]	[0.123]	[0.016]	[0.014]	[0.026]	[0.016]	[0.095]
Never movers	−0.008	0.266	0.051	0.016	0.029	0.037	0.021	−0.014
	(0.006)	(0.238)	(0.153)	(0.024)	(0.022)	(0.040)	(0.022)	(0.279)
	[0.006]	[0.209]	[0.119]	[0.025]	[0.022]	[0.047]	[0.023]	[0.219]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control average	0.378	8.329	4.852	8.361	7.870	6.735	8.259	0.700
Observations average	52,222	68,938	68,938	77,201	77,105	77,119	77,205	49,336
Mean control never movers	0.32	12.446	7.477	7.766	7.318	5.869	7.636	1.129
Observations never movers	11,241	13,659	13,659	18,745	18,735	18,720	18,747	6,153

Every cell constitutes a separate regression of the instrument on the dependent variable in the header using individual data from the Cambodian socio-economic survey 1996–2014. The row names define whether the individual ever moved and has been in that village since birth. Variabels with 'p.c.' are denominated by household size. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Poverty and income inequality

	(1) Poverty Rate (Head Count Ratio)	(2)	(3) Poverty gap	(4)	(5) Poverty severity	(6)	(7) Gini coefficient	(8)
Productive during KR	−0.009	−0.006	−0.005	−0.004	−0.003	−0.002	0.001	0.001
	(0.016)	(0.011)	(0.007)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
	[0.015]	[0.010]	[0.007]	[0.005]	[0.004]	[0.003]	[0.003]	[0.004]
Commune characteristics		Yes		Yes		Yes		Yes
Observations	1,470	1,470	1,470	1,470	1,470	1,470	1,470	1,470
Mean non-productive	0.388	0.388	0.119	0.119	0.052	0.052	0.304	0.304

Data about poverty taken from Cambodian EMIS census data on enrollment and school characteristics in 1997. Head count ratio is the proportion of a population that lives below the poverty line. Poverty gap is defined as the ratio by which the mean income of the poor falls below the poverty line. Poverty severity is defined as the squares of the poverty gaps relative to the poverty line. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Migration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Returned 1979/1980		Returned 1979		Return after displacement		In village during KR	
Alive during KR	0.004	0.011	−0.004	0.001	0.001	0.004	0.021	0.011
	(0.013)	(0.012)	(0.008)	(0.008)	(0.008)	(0.007)	(0.016)	(0.016)
	[0.012]	[0.013]	[0.009]	[0.009]	[0.007]	[0.007]	[0.014]	[0.014]
Older than 18 during KR	0.007	0.018	−0.004	0.002	0.007	0.012	0.027*	0.008
	(0.015)	(0.016)	(0.010)	(0.010)	(0.010)	(0.009)	(0.017)	(0.016)
	[0.014]	[0.016]	[0.010]	[0.011]	[0.008]	[0.009]	[0.014]	[0.014]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes
Mean control alive during KR	0.219	0.205	0.163	0.150	0.071	0.062	0.426	0.415
Observations alive during KR	75,112	60,707	75,112	60,707	75,112	60,707	75,112	60,707
Mean control older than 18 during KR	0.281	0.271	0.209	0.194	0.092	0.082	0.421	0.399
Observations older than 18 during KR	33,245	23,671	33,245	23,671	33,245	23,671	33,245	23,671

Every cell constitutes a separate regression of the instrument on the dependent variable in the header using individual data from the Cambodian socio-economic survey 1996–2014. The row names define the sample used based on whether the individual had reached adulthood in 1978. 'Returned 1979/1980' defines whether an individual returned in either of these years and stayed until the survey. 'Returned 1979' narrows this down to the 15% of individuals who returned directly after the genocide. 'Returned after displacement' is a variable that asked whether an individual returned in 1979 and gave the reason that you were displaced. The last two columns estimate the probability that an individual was in the commune during the genocide. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Market access or public infrastructure

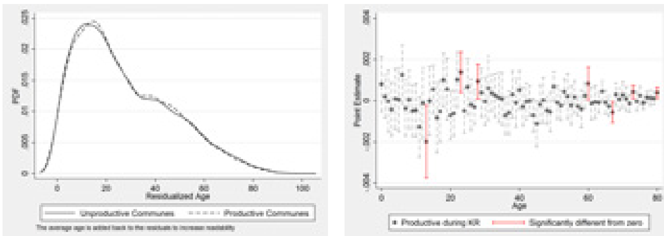
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Market access and public infrastructure							
	Distance to food store	Distance to bank	Distance to extension worker	Distance to market	Distance to agricultural market	% Pop with electricity	% Pop with piped water	Public hospital
Productive during KR	−0.337	−0.136	−1.159	−0.385	−0.217	0.789	−0.252	0.028
	(0.493)	(0.675)	(1.100)	(0.666)	(0.653)	(1.680)	(1.983)	(0.019)
	[0.493]	[0.645]	[1.010]	[0.620]	[0.591]	[1.384]	[1.275]	[0.019]
	School characteristics							
	Distance to school	Village with school	Director with degree	log School income p.c.	Enrollment rate	# Teachers	Student-teacher-ratio	Number of classes
Productive during KR	0.060	0.081	0.002	0.041	0.881	0.573	0.601	−0.085
	(0.059)	(0.229)	(0.002)	(0.069)	(1.004)	(3.526)	(1.627)	(0.314)
	[0.074]	[0.170]	[0.002]	[0.058]	[0.941]	[3.286]	[1.647]	[0.274]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean market access	6.272	10.698	18.123	7.060	7.190	37.027	27.236	0.119
Observations market access	3,593	3,665	3,724	3,684	3,614	3,812	3,812	3,027
Mean school characteristics	1.370	6.404	0.002	8.529	39.705	53.023	41.727	7.908
Observations school characteristics	1,593	1,621	1,543	1,436	4,518	1,592	1,592	1,592

Data on market access and public infrastructure taken from the village survey of the Cambodian socio economic survey 1996–2014. Data on School characteristics taken from Cambodian EMIS census data on enrollment and school characteristics in 1997–2002. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

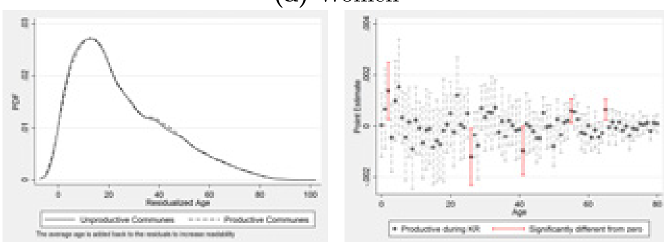
2.A Appendix

2.A.1 Figures

Figure A.1: Distributional effects: Sex ratio



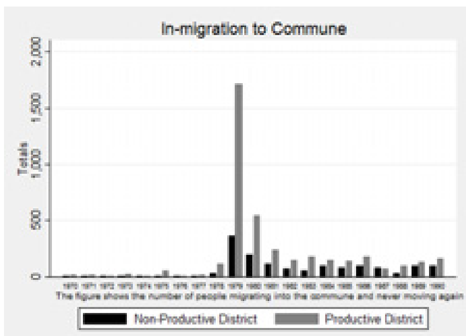
(a) Women



(b) Men

The distribution of age in the Cambodian socio economic survey 1996–2014, separated by the productiveness of the commune during the Khmer Rouge regime and sex of the respondent. Histogram on the residualized distributions (left) and point estimates on the difference between the distributions for every age between 0–80. Differences based on whether the commune was productive during the genocide. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics included and defined in Table 1.

Figure A.2: In Migration



In migration into commune, based on productiveness status of district during the Khmer Rouge. Source: Cambodian socio economic survey 1996

2.A.2 Tables

Table A.1: Growing season shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	#Bodies		#Massgraves		Memorial		Violence Index	
Productive during harvest season	391.671***	480.267**	7.728***	13.211***	0.023***	0.018	0.136***	0.147***
	(149.841)	(272.957)	(3.174)	(5.781)	(0.011)	(0.012)	(0.046)	(0.054)
	[133.149]	[243.396]	[2.822]	[3.767]	[0.008]	[0.012]	[0.033]	[0.053]
Productive during growing season	30.288	120.672	−2.136	3.458	0.003	−0.002	0.005	0.016
	(246.852)	(154.951)	(3.636)	(4.238)	(0.014)	(0.014)	(0.064)	(0.060)
	[166.737]	[159.584]	[3.922]	[3.959]	[0.008]	[0.012]	[0.037]	[0.050]
Interaction harvest × growing		−183.105		−11.332*		0.010		−0.022
		(466.051)		(7.378)		(0.015)		(0.086)
		[345.908]		[5.862]		[0.016]		[0.082]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean non-productive	407.873	407.873	7.094	7.094	0.035	0.035	−0.063	−0.063
Observations	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621

Robustness to including growing season controls. We include a binary variable indicating less rain during the growing season May–August, and its interaction in odd columns. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Rice yields using different shock definitions

	(1)	(2)	(3)	(4)	(5)	(6)						
	Shock in survey year						Average over three previous years					
	Yield in tons per hectar			Standardized yields			Yield in tons per hectar			Standardized yields		
Productive during harvest season	0.173***	0.149**	0.213***	0.077***	0.066***	0.083***	0.201**	0.176**	0.185**	0.075**	0.064**	0.059*
	(0.087)	(0.080)	(0.095)	(0.027)	(0.025)	(0.028)	(0.098)	(0.089)	(0.086)	(0.036)	(0.031)	(0.032)
	[0.063]	[0.062]	[0.056]	[0.019]	[0.017]	[0.018]	[0.085]	[0.077]	[0.092]	[0.032]	[0.028]	[0.030]
Continuous	0.088**	0.068*	0.078**	0.039***	0.031***	0.035***	0.085**	0.062*	0.063*	0.041***	0.032**	0.032**
	(0.052)	(0.044)	(0.048)	(0.015)	(0.013)	(0.015)	(0.045)	(0.037)	(0.039)	(0.017)	(0.015)	(0.015)
	[0.040]	[0.035]	[0.032]	[0.012]	[0.010]	[0.011]	[0.036]	[0.032]	[0.038]	[0.014]	[0.013]	[0.015]
Continuous, one SD	0.178***	0.183***	0.194***	0.067***	0.069***	0.074***	0.107**	0.085*	0.112	0.049**	0.040**	0.050**
	(0.065)	(0.073)	(0.088)	(0.020)	(0.023)	(0.026)	(0.059)	(0.054)	(0.056)	(0.021)	(0.019)	(0.018)
	[0.051]	[0.059]	[0.057]	[0.019]	[0.021]	[0.021]	[0.050]	[0.047]	[0.068]	[0.019]	[0.019]	[0.022]
Raw continuous variation	0.079**	0.067*	0.112***	0.039***	0.034***	0.053***	0.072	0.045	0.080	0.058***	0.048**	0.058
	(0.030)	(0.033)	(0.047)	(0.012)	(0.012)	(0.019)	(0.052)	(0.066)	(0.085)	(0.019)	(0.023)	(0.033)
	[0.035]	[0.036]	[0.041]	[0.013]	[0.013]	[0.017]	[0.073]	[0.077]	[0.117]	[0.022]	[0.022]	[0.037]
Commune characteristics		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
Year × province			Yes			Yes			Yes			Yes
Mean non-productive	3.041	3.041	3.041				3.041	3.041	3.041			
Observations	3,772	3,772	3,772	3,772	3,772	3,772	3,772	3,772	3,772	3,772	3,772	3,772

Every cell constitutes a separate regression of the row instrument on the dependent variable in the header from the Cambodian socio-economic survey 1996–2014. More productive implies relatively less rain during the harvest season. Yields are calculated at the individual level and then aggregated to the commune. In columns (1)–(6), the shock in the survey year is used while in columns (6)–(12) the average of the previous three years is taken. The variable 'Standardized yields' indicate standardized yields across Cambodia in a given year and the point estimates can be interpreted as standard deviation increases in productivity. One standard deviation is 3.213. 'Productive during harvest season' is a binary variable indicating whether the commune was above average productive during the harvest season. 'Continuous' is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders' allocation process. 'Continuous, one SD' takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes, introducing a spread of one standard deviation between productive and unproductive communes. 'Raw continuous variation' is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects, year fixed effects, and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Communes that were more productive during the Khmer Rouge do not have higher contemporaneous yields. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Alternative shock and violence definitions

	(1) #Bodies	(2) #Mass graves	(3) Memorial	(4) Violence Index	(5) Bodies per capita	(6) Bodies per sqkm	(7) Mass graves per capita	(8) Mass graves per sqkm	(9) log Bodies	(10) log Bodies, per capita	(11) log Bodies, per sqkm	(12) log Mass graves	(13) log Mass graves, per capita	(14) log Mass graves, per sqkm
Productive during KR	387.276***	8.038***	0.022***	0.135***	1.245***	7.459	0.026*	0.262***	0.183**	0.072***	0.148***	0.111***	0.015**	0.055***
	(150.958)	(3.265)	(0.010)	(0.043)	(0.595)	(4.109)	(0.014)	(0.085)	(0.079)	(0.027)	(0.044)	(0.042)	(0.006)	(0.015)
	[138.934]	[2.856]	[0.008]	[0.031]	[0.458]	[4.876]	[0.013]	[0.077]	[0.077]	[0.024]	[0.039]	[0.038]	[0.006]	[0.012]
‘Continuous’	0.012***	150.140***	4.006***	0.065***	0.565***	2.652	0.014*	0.119***	0.082*	0.028*	0.070***	0.045*	0.007**	0.025***
	(0.004)	(57.207)	(1.606)	(0.016)	(0.209)	(2.399)	(0.007)	(0.044)	(0.051)	(0.014)	(0.025)	(0.025)	(0.003)	(0.008)
	[0.004]	[50.660]	[1.547]	[0.013]	[0.199]	[2.307]	[0.008]	[0.040]	[0.047]	[0.015]	[0.022]	[0.024]	[0.004]	[0.007]
‘Continuous, one SD’	0.012**	343.592***	7.056***	0.093***	1.181***	5.431	0.023**	0.193***	0.190***	0.069***	0.138***	0.091***	0.012***	0.039***
	(0.004)	(128.034)	(2.760)	(0.022)	(0.302)	(4.430)	(0.009)	(0.070)	(0.063)	(0.017)	(0.033)	(0.034)	(0.004)	(0.012)
	[0.005]	[121.771]	[2.396]	[0.020]	[0.312]	[4.571]	[0.009]	[0.063]	[0.056]	[0.018]	[0.031]	[0.027]	[0.004]	[0.009]
Raw continuous variation	0.032*	1025.012***	24.017***	0.281***	3.554***	31.724***	0.078**	0.679***	0.326	0.152**	0.307***	0.242**	0.043***	0.133***
	(0.017)	(411.779)	(8.154)	(0.091)	(1.095)	(10.391)	(0.031)	(0.189)	(0.212)	(0.057)	(0.105)	(0.124)	(0.014)	(0.034)
	[0.018]	[382.599]	[7.081]	[0.074]	[1.149]	[9.948]	[0.033]	[0.180]	[0.209]	[0.063]	[0.104]	[0.119]	[0.015]	[0.029]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean non-productive	0.035	407.873	7.094		1.693	23.189	0.027	0.162	0.714	0.189	0.312	0.317	0.018	0.057
Observations	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621	1,621

Various indicators of violence, denominated by population or area, and taken the logarithm. ‘Productive during KR’ is a binary variable indicating whether the commune was above average productive during the genocide. ‘Continuous’ is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders’ allocation process. ‘Continuous, one SD’ takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes. ‘Raw continuous variation’ is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Dropping large communes: Violence definitions

	(1) #Bodies	(2) #Mass graves	(3) Memorial	(4) Violence Index	(5) Bodies per capita	(6) Bodies per sqkm	(7) Mass graves per capita	(8) Mass graves per sqkm	(9) log Bodies	(10) log Bodies, per capita	(11) log Bodies, per sqkm	(12) log Mass graves	(13) log Mass graves, per capita	(14) log Mass graves, per sqkm
All communes	387.276***	8.038***	0.022***	0.135***	1.245***	7.459	0.026*	0.262***	0.183**	0.072***	0.148***	0.111***	0.015**	0.055***
(1,621 communes)	(150.958)	(3.265)	(0.010)	(0.043)	(0.595)	(4.109)	(0.014)	(0.085)	(0.079)	(0.027)	(0.044)	(0.042)	(0.006)	(0.015)
	[138.934]	[2.856]	[0.008]	[0.031]	[0.458]	[4.876]	[0.013]	[0.077]	[0.077]	[0.024]	[0.039]	[0.038]	[0.006]	[0.012]
All communes$\leq$ 99th percentile	384.081***	7.821***	0.022***	0.132***	1.253***	6.488	0.026*	0.244***	0.165**	0.070***	0.136***	0.103***	0.015**	0.051***
(1,605 communes)	(154.943)	(3.381)	(0.010)	(0.042)	(0.600)	(4.283)	(0.014)	(0.093)	(0.084)	(0.027)	(0.046)	(0.041)	(0.007)	(0.016)
	[141.230]	[2.894]	[0.007]	[0.030]	[0.461]	[5.011]	[0.013]	[0.081]	[0.081]	[0.024]	[0.039]	[0.038]	[0.006]	[0.012]
All communes$\leq$ 95th percentile	418.539**	7.405**	0.026***	0.147***	1.336**	14.497***	0.024	0.247**	0.185**	0.078***	0.144***	0.116***	0.015**	0.053***
(1,540 communes)	(197.018)	(3.963)	(0.012)	(0.051)	(0.736)	(6.069)	(0.018)	(0.111)	(0.093)	(0.033)	(0.053)	(0.041)	(0.008)	(0.019)
	[175.937]	[3.309]	[0.009]	[0.040]	[0.578]	[5.470]	[0.017]	[0.096]	[0.086]	[0.028]	[0.044]	[0.037]	[0.007]	[0.015]
All communes$\leq$ 90th percentile	315.459***	8.130**	0.020**	0.121***	1.203**	13.246***	0.026	0.267**	0.157*	0.071**	0.125***	0.114***	0.016**	0.055***
(1,459 communes)	(140.213)	(4.106)	(0.009)	(0.043)	(0.754)	(5.886)	(0.019)	(0.121)	(0.092)	(0.036)	(0.055)	(0.039)	(0.009)	(0.020)
	[114.167]	[3.469]	[0.008]	[0.034]	[0.545]	[4.951]	[0.018]	[0.105]	[0.080]	[0.028]	[0.044]	[0.036]	[0.008]	[0.016]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Every row drops communes that have a pre-genocide population above the 99th, 95th, or 90th, percentile of the distribution. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Alternative shock definitions: Election results

	National election 2013					Commune elections 2012 and 2017				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Vote share CNRP	Vote share CPP	Turnout	Vote share CPP $\geq$ 50%	Margin -$\vert$CPP-CNRP$\vert$	Vote share CNRP	Vote share CPP	Turnout	Vote share CPP $\geq$ 50%	Margin -$\vert$CPP-CNRP$\vert$
Productive during KR	4.890***	−4.218***	2.939**	−0.155***	1.754	2.100***	−2.016***	3.080**	−0.069***	2.383***
	(0.879)	(0.876)	(1.474)	(0.038)	(1.147)	(0.680)	(0.811)	(1.686)	(0.024)	(1.048)
	[0.600]	[0.614]	[1.292]	[0.026]	[1.305]	[0.493]	[0.565]	[1.478]	[0.019]	[0.851]
Continuous	2.992***	−2.611***	1.282	−0.092***	1.534*	1.242***	−1.259***	1.201	−0.043***	1.526***
	(0.489)	(0.487)	(0.795)	(0.020)	(0.805)	(0.380)	(0.463)	(0.878)	(0.012)	(0.610)
	[0.330]	[0.328]	[0.795]	[0.011]	[0.826]	[0.316]	[0.349]	[0.905]	[0.012]	[0.478]
Continuous, one SD	3.368***	−2.934***	1.715*	−0.104***	1.809*	1.286***	−1.237***	1.661	−0.040***	1.404***
	[0.362]	[0.352]	[0.918]	[0.014]	[0.934]	[0.348]	[0.370]	[1.029]	[0.015]	[0.513]
	[0.362]	[0.352]	[0.918]	[0.014]	[0.934]	[0.348]	[0.370]	[1.029]	[0.015]	[0.513]
Raw continuous variation	13.085***	−10.846***	2.716	−0.393***	4.522	4.265***	−4.584***	2.611	−0.176***	5.165**
	(2.502)	(2.724)	(3.685)	(0.117)	(4.090)	(1.862)	(2.157)	(3.316)	(0.053)	(2.772)
	[2.023]	[1.986]	[3.458]	[0.078]	[3.893]	[1.564]	[1.669]	[3.385]	[0.046]	[2.374]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects						Yes	Yes	Yes	Yes	Yes
Mean non-productive	37.512	54.782	77.274	0.593	−27.889	33.683	61.664	75.427	0.782	−31.053
Observations	1,621	1,621	1,621	1,621	1,621	3,230	3,230	3,230	3,230	3,230

Data taken from the official election results. ‘Vote share CNRP’ in the communal elections 2012 is calculated as the combined votes of the ‘Sam Rainsy Party’ and the ‘Human Rights Party’. ‘Turnout’ is calculated using the electorate information from the national elections in 2013. ‘Vote share CPP $\geq$ 50%’ is a binary variable indicating an absolute majority for the incumbent party CPP. ‘Margin -|CPP-CNRP|’ is calculated as the vote share of CPP minus CNRP and a variation of the competitiveness measure by Besley et al. (2010). ‘Productive during KR’ is a binary variable indicating whether the commune was above average productive during the genocide. ‘Continuous’ is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders’ allocation process. ‘Continuous, one SD’ takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes. ‘Raw continuous variation’ is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: **Dropping large communes**: Election results

	National election 2013					Commune elections 2012 and 2017				
	(1) Vote share CNRP	(2) Vote share CPP	(3) Turnout	(4) Vote share CPP ≥ 50%	(5) Margin -|CPP-CNRP|	(6) Vote share CNRP	(7) Vote share CPP	(8) Turnout	(9) Vote share CPP ≥ 50%	(10) Margin -|CPP-CNRP|
All communes	4.890***	−4.218***	2.939**	−0.155***	1.754	2.100***	−2.016***	3.080**	−0.069***	2.383***
(1,621 communes)	(0.879)	(0.876)	(1.474)	(0.038)	(1.147)	(0.680)	(0.811)	(1.686)	(0.024)	(1.048)
	[0.600]	[0.614]	[1.292]	[0.026]	[1.305]	[0.493]	[0.565]	[1.478]	[0.019]	[0.851]
All communes≤ 99th percentile	4.908***	−4.237***	2.847**	−0.154***	1.830	2.104***	−2.027***	2.420**	−0.070***	2.027***
(1,605 communes)	(0.890)	(0.888)	(1.482)	(0.039)	(1.144)	(0.686)	(0.821)	(1.698)	(0.024)	(1.072)
	[0.605]	[0.618]	[1.294]	[0.027]	[1.295]	[0.496]	[0.571]	[1.486]	[0.019]	[0.866]
All communes≤ 95th percentile	4.090***	−3.511***	1.770	−0.123***	2.110	1.853***	−1.642***	1.606	−0.054***	2.095**
(1,540 communes)	(0.879)	(0.905)	(1.611)	(0.040)	(1.230)	(0.696)	(0.839)	(1.782)	(0.022)	(1.138)
	[0.625]	[0.634]	[1.467]	[0.026]	[1.382]	[0.519]	[0.586]	[1.543]	[0.017]	[0.908]
All communes≤ 90th percentile	3.899***	−3.276***	1.897	−0.119***	1.869	1.751***	−1.529**	1.754	−0.056***	2.020**
(1,459 communes)	(0.950)	(0.973)	(1.644)	(0.043)	(1.335)	(0.782)	(0.937)	(1.821)	(0.024)	(1.283)
	[0.658]	[0.662]	[1.478]	[0.027]	[1.428]	[0.565]	[0.638]	[1.565]	[0.017]	[1.004]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects						Yes	Yes	Yes	Yes	Yes

Data taken from the official election results. 'Vote share CNRP' in the communal elections 2012 is calculated as the combined votes of the 'Sam Rainsy Party' and the 'Human Rights Party'. 'Turnout' is calculated using the electorate information from the national elections in 2013. 'Vote share CPP ≥ 50%' is a binary variable indicating an absolute majority for the incumbent party CPP. 'Margin -|CPP-CNRP|' is calculated as the vote share of CPP minus CNRP and a variation of the competitiveness measure by Besley et al. (2010). Every row drops communes that have a pre-genocide population above the 99th, 95th, or 90th, percentile of the distribution. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Results from Asia Foundation: Adjusting for multiple hypothesis testing for the average effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Without individual controls				With individual controls				Summary Statistics			
									Non-productive communes		Productive communes	
	beta	s.e.	p-value	FDR adj. p-value	beta	s.e.	p-value	FDR adj. p-value	Mean	S.D.	Mean	S.D.
	Category: Voter informedness											
Frequency: Listen to radio	0.036	(0.033)	0.269	0.506	0.047	(0.033)	0.149	0.271	1.798	0.873	1.818	0.876
Frequency: Watch TV	0.131	(0.061)	0.031	0.129	0.132	(0.056)	0.018	0.069	2.870	1.411	3.028	1.451
Know parties are different	0.010	(0.051)	0.839	0.887	0.004	(0.053)	0.956	0.682	2.715	1.239	2.744	1.255
Can name representative	0.047	(0.022)	0.033	0.129	0.047	(0.020)	0.018	0.069	0.116	0.321	0.182	0.386
Know whether representative visited	0.027	(0.021)	0.211	0.506	0.028	(0.020)	0.171	0.271	0.194	0.396	0.229	0.420
Know role of parties in assembly	0.018	(0.024)	0.470	0.887	0.021	(0.023)	0.378	0.426	0.305	0.461	0.347	0.476
Understands purpose of democracy	0.014	(0.028)	0.609	0.887	0.025	(0.025)	0.325	0.426	0.580	0.494	0.542	0.499
z-Score	0.069	(0.025)			0.067	(0.023)						
	Category: Support for democracy											
Democracy preferred to strong leader	0.001	(0.014)	0.938	0.350	0.002	(0.033)	0.337	0.252	0.904	0.295	0.892	0.311
One can vote against the government	−0.035	(0.019)	0.069	0.137	−0.033	(0.018)	0.068	0.099	0.854	0.353	0.839	0.368
Not voted because told to vote	0.028	(0.019)	0.144	0.177	0.028	(0.018)	0.124	0.115	0.941	0.236	0.967	0.180
Government and people are equals	0.053	(0.023)	0.024	0.078	0.052	(0.027)	0.051	0.090	0.366	0.482	0.434	0.496
All Political parties should hold events	0.030	(0.013)	0.020	0.078	0.029	(0.013)	0.031	0.088	0.905	0.293	0.924	0.265
Democracy empowers People	0.035	(0.020)	0.089	0.144	0.050	(0.021)	0.016	0.077	0.141	0.348	0.137	0.345
Women make own choice in voting	0.035	(0.012)	0.002	0.026	0.035	(0.012)	0.004	0.043	0.858	0.349	0.889	0.314
Women as a representative	0.046	(0.036)	0.207	0.208	0.049	(0.036)	0.172	0.115	1.080	0.901	1.115	0.893
Would like to see more women	0.011	(0.012)	0.355	0.247	0.015	(0.011)	0.181	0.115	0.944	0.230	0.954	0.209
Reserved top list place for women	0.038	(0.026)	0.150	0.177	0.048	(0.024)	0.040	0.088	0.542	0.499	0.570	0.496
z-Score	0.041	(0.012)			0.046	(0.012)						
	Category: Local civic participation											
Member of #civil associations (CA)	−0.112	(0.043)	0.009	0.017	−0.101	(0.039)	0.009	0.013	0.416	0.910	0.346	0.825
Took part in a meeting of a CA	−0.028	(0.022)	0.201	0.088	−0.031	(0.022)	0.152	0.083	0.218	0.413	0.208	0.422
Helped reach a decision of a CA	−0.039	(0.015)	0.010	0.017	−0.034	(0.013)	0.008	0.013	0.129	0.336	0.120	0.325
Local government affects my life	−0.118	(0.044)	0.008	0.017	−0.134	(0.040)	0.001	0.004	0.535	0.499	0.434	0.496
Would report election crime	−0.102	(0.067)	0.127	0.088	−0.068	(0.060)	0.262	0.118	3.228	1.065	3.170	1.129
z-Score	−0.076	(0.023)			−0.081	(0.021)						
	Category: Trust											
Trust in neighborhood	−0.204	(0.053)	0.000	0.001	−0.181	(0.060)	0.002	0.005	0.196	0.397	0.202	0.402
Trust in general	−0.039	(0.020)	0.054	0.028	−0.039	(0.019)	0.040	0.021	2.485	0.702	2.257	0.780
z-Score	−0.123	(0.031)			−0.126	(0.030)						

Individual covariates are ethnicity, year of birth, education, income, interview circumstance, rural status and housing status. 'FDR adj. p-value' represent p-values adjusted for false detection rates following Anderson (2008). Zone fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics included in all estimations and are defined in Table 1. Standard errors and p-values corrected for spatial dependence within 1 degree.

Table A.8: Alternative Shock Definitions: Asia Foundation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Voter informedness			Support for democracy			Local civic participation			Trust		
	All	Alive During KR	Born After KR	All	Alive During KR	Born After KR	All	Alive During KR	Born After KR	All	Alive During KR	Born After KR
Productive during KR	0.067^{***}	0.049^{*}	0.110^{***}	0.046^{***}	0.054^{***}	0.022	-0.081^{***}	-0.099^{***}	-0.009	-0.126^{***}	-0.131^{***}	-0.123^{***}
	0.027	0.032	0.033	0.009	0.009	0.014	0.022	0.023	0.032	0.035	0.036	0.055
	[0.023]	[0.028]	[0.033]	[0.012]	[0.012]	[0.019]	[0.021]	[0.024]	[0.021]	[0.030]	[0.033]	[0.045]
Continuously within province	0.055^{***}	0.042^{***}	0.089^{***}	0.022^{***}	0.025^{***}	0.022	-0.038^{***}	-0.045^{***}	0.011	-0.079^{***}	-0.072^{***}	-0.080^{***}
	0.018	0.018	0.021	0.005	0.005	0.011	0.013	0.012	0.018	0.014	0.011	0.035
	[0.015]	[0.016]	[0.021]	[0.006]	[0.006]	[0.016]	[0.014]	[0.013]	[0.015]	[0.016]	[0.013]	[0.029]
Continuously within province, one SD	0.047^{**}	0.034	0.065^{***}	0.021^{**}	0.027^{***}	0.012	-0.041^{***}	-0.048^{***}	0.006	-0.083^{***}	-0.069^{***}	-0.089^{***}
	0.026	0.029	0.024	0.007	0.007	0.016	0.017	0.021	0.016	0.017	0.017	0.032
	[0.023]	[0.026]	[0.022]	[0.009]	[0.010]	[0.017]	[0.015]	[0.016]	[0.016]	[0.018]	[0.017]	[0.030]
Raw Continuous Variation	0.130^{**}	0.107^{*}	0.106	0.087^{***}	0.128^{***}	-0.027	-0.143^{**}	-0.141^{**}	-0.053	-0.315^{***}	-0.270^{***}	-0.271^{**}
	0.072	0.076	0.076	0.022	0.018	0.035	0.063	0.058	0.093	0.109	0.100	0.165
	[0.054]	[0.060]	[0.065]	[0.029]	[0.024]	[0.046]	[0.059]	[0.063]	[0.067]	[0.090]	[0.093]	[0.135]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,999	1,321	681	1,999	1,321	681	1,999	1,321	681	1,999	1,321	681

Individual covariates are ethnicity, year of birth, education, income, interview circumstance, rural status and housing status. ‘Productive during KR’ is a binary variable indicating whether the commune was above average productive during the genocide. ‘Continuously within province’ is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders allocation process. ‘Continuously within province, on SD’ takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes. ‘Raw continuous variation’ is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: **Exit and voice** Using province fixed effects instead

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Voter informedness		Support for democracy		Local civic participation		Trust	
Average effect	0.051**	0.058**	0.044***	0.049***	−0.043**	−0.038**	−0.094***	−0.096***
	0.031	0.032	0.012	0.012	0.022	0.020	0.023	0.022
	[0.026]	[0.027]	[0.013]	[0.013]	[0.018]	[0.018]	[0.020]	[0.020]
Alive during KR	0.034	0.049*	0.058***	0.064***	−0.067***	−0.065***	−0.080***	−0.090***
	0.035	0.037	0.015	0.015	0.027	0.025	0.023	0.026
	[0.027]	[0.029]	[0.015]	[0.015]	[0.021]	[0.021]	[0.021]	[0.022]
Born After KR	0.069	0.037	0.019	0.023	0.054	0.047	−0.089*	−0.080
	0.043	0.030	0.026	0.028	0.036	0.040	0.050	0.046
	[0.047]	[0.035]	[0.027]	[0.029]	[0.035]	[0.035]	[0.051]	[0.051]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes

Results using questions from the Asia Foundation 2003 and 2013. Individual results per category show in Table 4. Individual covariates are ethnicity, year of birth, education, income, interview circumstance, rural status and housing status. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Zone fixed effects sort provinces in four zones to improve power. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: **Alternative shock definitions:** Local state avoidance

	(1) Property taxes paid	(2) Share of property tax	(3) Property tax per sqm of housing	(4) Standardized tax score	(5) Working for the government	(6) Self employment
Productive during KR	−6.902**	−0.001*	−133.906***	−0.021**	−0.008***	0.007*
	(3.403)	(0.000)	(66.202)	(0.009)	(0.003)	(0.005)
	[2.699]	[0.000]	[41.728]	[0.011]	[0.002]	[0.004]
Continuously within province	−3.739*∗	−0.000	−84.319**	−0.011*	−0.006***	0.007**
	(2.131)	(0.000)	(47.601)	(0.005)	(0.002)	(0.003)
	[1.648]	[0.000]	[34.072]	[0.006]	[0.002]	[0.003]
Continuously within province, one SD	−2.919*	−0.000**	−111.218**	−0.014***	−0.006***	0.006**
	(2.322)	(0.000)	(56.804)	(0.007)	(0.002)	(0.003)
	[1.555]	[0.000]	[45.673]	[0.004]	[0.002]	[0.003]
Raw Continuous variation	−4.844	−0.001	−98.719	−0.025	−0.027***	0.019
	(9.680)	(0.001)	(183.033)	(0.021)	(0.008)	(0.016)
	[7.718]	[0.001]	[128.959]	[0.021]	[0.008]	[0.013]
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Mean non-productive	61.386	0.005	974.410	0.069	0.314	
Observations	16,513	16,513	16,513	16,513	118,849	118,849

'Productive during KR' is a binary variable indicating whether the commune was above average productive during the genocide. 'Continuously within province' is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders allocation process. 'Continuously within province, on SD' takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes. 'Raw continuous variation' is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Alternative shock definitions: Deforestation

	(1) log Forest loss	(2) Land concessions	(3) Illegal logging	(4) Illegal overuse
Productive during KR	-0.501^{***}	-0.048^{*}	-0.032^{**}	-0.051^{***}
	(0.151)	(0.026)	(0.016)	(0.017)
	[0.131]	[0.027]	[0.016]	[0.016]
Continuously within province	-0.297^{***}	-0.035^{***}	-0.021^{**}	-0.028^{***}
	(0.066)	(0.013)	(0.009)	(0.010)
	[0.067]	[0.013]	[0.009]	[0.009]
Continuously within province, one SD	-0.365^{***}	-0.032^{*}	-0.028^{***}	-0.028^{**}
	(0.100)	(0.018)	(0.009)	(0.013)
	[0.097]	[0.018]	[0.011]	[0.011
Raw Continuous variation	-1.282^{***}	-0.147^{*}	-0.055	-0.108^{***}
	(0.270)	(0.066)	(0.035)	(0.039)
	[0.289]	[0.077]	[0.037]	[0.033]
Commune characteristics	Yes	Yes	Yes	Yes
Year fixed effects			Yes	Yes
Mean non-productive	3.846	0.317	0.252	0.304
Observations	1,621	1,621	3,027	3,027

'Productive during KR' is a binary variable indicating whether the commune was above average productive during the genocide. 'Continuously within province' is the continuous version of our standard binary instrument. After standardizing each commune by its mean and standard deviation, we standardize again within each province to match the Khmer Rouge leaders allocation process. 'Continuously within province, on SD' takes the value one if the within-province-standardized rain is larger than 0.5, minus one if it is smaller than -0.5 and zero otherwise, introducing a spread of one standard deviation between productive and unproductive communes. 'Raw continuous variation' is the rainfall in each commune standardized by the commune mean and standard deviation. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Dropping large communes: Deforestation

	(1) log Forest loss	(2) Land concessions	(3) Illegal logging	(4) Illegal overuse
Productive during KR	-0.501^{***}	-0.048^{*}	-0.032^{**}	-0.051^{***}
(1,621 communes)	(0.151)	(0.026)	(0.016)	(0.017)
	[0.131]	[0.027]	[0.016]	[0.016]
All communes≤ 99th percentile	-0.494^{***}	-0.050^{*}	-0.032^{**}	-0.049^{***}
(1,605 communes)	(0.153)	(0.026)	(0.016)	(0.018)
	[0.133]	[0.026]	[0.016]	[0.016]
All communes≤ 95th percentile	-0.376^{***}	-0.032	-0.030^{*}	-0.049^{***}
(1,540 communes)	(0.149)	(0.026)	(0.017)	(0.019)
	[0.106]	[0.024]	[0.017]	[0.017
All communes≤ 90th percentile	-0.402^{***}	-0.035	-0.030^{*}	-0.049^{***}
(1,459 communes)	(0.157)	(0.026)	(0.017)	(0.019)
	[0.105]	[0.024]	[0.017]	[0.017]
Geographic controls	Yes	Yes	Yes	Yes
Commune characteristics	Yes	Yes	Yes	Yes
Year fixed effects			Yes	Yes

Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: Shift in sectoral composition

	(1)	(2)	(3)	(4)	(5)	(6)
	log #Establishments		Employment share: Agriculture		Employment share: Manufacturing	
Productive during KR	0.054	0.063	-2.025	-1.995^{**}	2.807^{***}	2.734^{***}
	(0.085)	(0.057)	(3.040)	(1.181)	(1.020)	(0.649)
	[0.066]	[0.046]	[1.810]	[0.838]	[0.994]	[0.760]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Commune characteristics		Yes		Yes		Yes
Observations	1,611	1,611	1,614	1,614	1,614	1,614
Mean non-productive	5.237	5.237	79.213	79.213	5.578	5.578

Data on the number of establishments taken from the economic census in 2011. Data about the sectoral composition taken from population census in 2008. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.14: Public investments and night time lights

	(1) Maximum night time light	(2) Any night time light 2013	(3) Night time light in 2013	(4) #Markets in commune	(5) Distance to health center	(6) Radio access
Productive during KR	−1.128	0.025	−0.216	−0.020	0.027	0.022
	(0.970)	(0.029)	(0.613)	(0.028)	(0.030)	(0.019)
	[0.805]	[0.018]	[0.458]	[0.032]	[0.033]	[0.018]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Commune characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for 1992 value		Yes	Yes			
Observations	1,621	1,621	1,621	1,621	1,621	1,621
Mean non-productive	9.404	0.409	7.164	0.424	0.688	0.881

Data on the number of markets, access to health facilities, and radio stations obtained from Open Development Cambodia. Night time light data from NOAA covering the years 1992–2013. 'Maximum night time light' indicates the highest observed mean luminosity in the commune. 'Any night time light 2013' is a binary variable indicating whether the mean in 2013 was non-zero. Province fixed effects and a second-degree polynomial in latitude in longitude included in all regressions. Commune characteristics are defined in Table 1. Standard errors clustered by 24 provinces shown in parenthesis and corrected for spatial dependence within 1 degree in brackets. Symbols reflect significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B Questions from the Asia Foundation

Table B.1: Informedness

Frequency: Listen to radio	How frequently do you listen to radio?
Frequency: Watch TV	How frequently do you watch TV?
Know parties are different	What difference do you see, if any, between the different political parties in Cambodia today?
Can name representative	Many people are not sure of the names of their province's representative in the National Assembly. Can you name yours?
Know whether representative visited	As far as you know, have any of the candidates elected to the National Assembly who represent your province visited your area since the last National Assembly election?
Know role of parties in assembly	Different people have different ideas about what the people in the National Assembly do? What do you think they do?
Understands purpose of democracy	If a country is called a democracy, what does this mean to you? (Any answer)

Table B.2: Category: Preferences

Democracy preferred to strong leader	On some occasions, democracy doesn't work. When that happens there are people that say we need a strong leader who doesn't have to be elected through voting. Others say that even if things don't function, democracy is always the best. What do you think?
One can vote against the government	Some people say, "Even if we are not happy with the government, we cannot vote against it. They are the high authority." Other people say, "If you are unhappy with the government, you should vote for another party to let the government know you are unhappy." Which of these is closer to your view?
Not voted because told to vote	What is the most important reason why you want to vote? (Not because she was told to)
Government and people are equals	Here are some different ways people think about the government. The first is that the people and government should be equals, and government should listen to the criticisms voiced by people. The second is that government should be like a father and the people like a child he must look after. The third is that the government is like a boss and the people like a worker who must obey. Which of these is closest to your view of what the government should be?
All Political parties should hold events	Do you think that all political parties, even the ones most people do not like, should be allowed to hold meetings in your area?
Democracy empowers People	If a country is called a democracy, what does this mean to you? (Answer: People are empowered)
Women make own choice in voting	Do you think a woman should make her own choice for voting, or do you think men should advise her on her choice?
Women as a representative	Would you prefer to be represented by a man or a woman in the National Assembly?
Would like to see more women	Would you like to see more women as members of the National Assembly?
Reserved top list place for women	In the National Assembly elections, every party has a list of candidates for the province, but usually only the top two or three people on the list have a chance of being elected. Knowing this, if a woman were included on a list in one of the top three places would you be more likely to vote for the list or less likely to vote for it?

Table B.3: Category: Taking local action

Member of # civil associations (CA)	Here is a list of organizations. As I mention each, please tell me if you belong to it.
Took part in a meeting of a CA	Have you ever participated in a meeting of an association or group you belong to?
Helped reach a decision of a CA	Have you ever helped make a decision at a meeting of an association or group you belong to?
Local government affects my life	Now I'm going to ask you a question about the local commune government. Tell me, whose decisions affect your life more: the national government in Phnom Penh, or the communal government in this town or village?
Would report election crime	If one of these problems were to happen in your area in the election, how likely would you be to report this problem - very likely, somewhat likely, somewhat unlikely or very unlikely?

Table B.4: Category: Trust

Trust in neighborhood	Now, speaking in general terms of the people from here, what do you think about people in this neighborhood are generally:
Trust in general	Generally speaking, do you think that most people can be trusted?

Chapter 3

The Effects of Migration and Ethnicity on African Economic Development[1]

3.1 Introduction

In recent years, researchers have started to scrutinize the historical determinants of economic underdevelopment in Africa. Until recently, African economies underperformed and the arbitrary partitioning of African ethnicities into states was identified to be a contributing factor (Alesina et al., 2016; Michalopoulos and Papaioannou, 2016; Clochard and Hollard, 2018). However, African states are now among the fastest growing nations, which suggests that a model of African growth based on the permanent adverse effects of divided ethnicities on institutions is incomplete.

In contrast to the literature on institutions, research on migration and trade suggests that more connections across borders might promote

[1]I benefited greatly from discussions with Konrad Burchardi, Dorothee Bühler, Serena Cocciolo, Masayuki Kudamatsu, Andreas Madestam, Stelios Michalopolous, Nathan Nunn, Torsten Persson, Thorsten Rogall, Jakob Svensson and Anna Tompsett. Thanks to the participants of the MIT Political Economy Lunch, Harvard History Lunch, Brown Macro Lunch, Nordic International Trade Seminars, the Oxford development workshop, the UCL Enter Seminar and NEUDC.

growth through various channels (Burchardi et al., 2016). However, as bilateral migration is potentially affected by bilateral trade and unobservable factors, causal identification is difficult in most settings. To achieve identification, the literature has resorted to using past migration flows or clusters of migration (Munshi, 2003; McKenzie and Rapoport, 2007). Especially in Africa, this masks considerable heterogeneity as ethnic identification still plays a large role. As migrants identify with their counterparts in the exporting country, this previously neglected population is likely to have a considerable impact on the effectiveness of migration.

Studying the effects of ethnic links across African countries entails many challenges. First, the data quality on African ethnicities is, at best, questionable. Different names of ethnicities, language barriers and historical ethnic conflicts constitute considerable obstacles to research. Second, conflicts or natural catastrophes might cause migration across borders leading to endogenous contemporaneous linkages. Third, if migration today is formed by the same preferences that shape cross-border trade, an omitted variable bias needs to be addressed.

To achieve identification, I employ a spatial identification strategy based on the dispersion of pre-colonial ethnicities and their division into separate countries by the formation of country borders. However, as borders are formed to reflect interests and populations are divided between countries, population shares and trade flows are endogenous to the location of the border. To estimate a causal effect on cross-border trade flows, I rely on the quasi-random formation of country borders in Africa. The use of geographical factors has recently attracted a considerable interest among economists due to their exogeneity to the individual (Nunn, 2008; Nunn and Wantchekon, 2011; Nunn and Puga, 2012; Michalopoulos and Papaioannou, 2013, 2014). I use a map of the pre-colonial distribution of ethnicities in Africa provided by Murdock (1959) to identify ethnic groups and their population shares as an instrument for migration and ethnic linkages today. Due to the randomness of country borders in Africa with respect to ethnicities, I argue that my approach identifies a causal effect of pre-colonial linkages on contemporaneous trade.

In this study, I exploit the exogenous placement of national boundaries with respect to historical ethnic homelands in Africa to evaluate whether ethnic connections across borders are associated with relatively higher rates of bilateral trade. Politically separated ethnic enclaves can be viewed as a network spanning multiple countries, which can constitute a barrier to, as well as an opportunity for, economic development. These networks may exclude non-members from economically profitable actions, but also overcome unfavorable institutions for members. In Africa in particular, members of the same ethnicity have been shown to share information or risks, bargain jointly for preferred policies, or be more productive (Fafchamps and Gubert, 2007; Bates, 2008; de la Sierra and Mutakumura, 2014; Hjort, 2014).

Using recent bilateral trade data and migration data combined with the historical distribution of ethnicities for 46 African countries, I find that migration and exports are positively correlated. I overcome the reversed causality problem by using past migration and pre-colonial ethnic connections between countries as instruments for current migration. Taking into account ethnic heterogeneity, the impact of migration is more than twice as large as previously estimated. The data provides no evidence of potential omitted variable biases from conflict, shared preferences, or linguistic similarity, and suggests a causal impact of the strength of cross-border ethnic connections on exports.

Further, I exploit the uniqueness of African countries and ethnicities to estimate where the gains from trade are located. Using nighttime light data, I show that ethnic connections lead to relative increases in wealth in more ethnically connected regions. Cross-border ethnic connections are likely to compensate for unfavorable institutions within and between African countries, potentially affecting economic development more positively than previously thought.

The magnitude of the effect is at the upper end of the estimates found in the literature for developed countries (Bandyopadhyay et al., 2008), which supports the hypothesis that ethnic linkages in developing countries are especially important due to high levels of corruption (Svensson, 2003; Dunlevy, 2006; Olken and Barron, 2009). I use data on government participation to show that ethnicities that are excluded from political participation or trust their government less, substitute governmental institutions and use their cross-border ethnic connections to facilitate exporting. Using historic variation in political centralization of ethnicities, the evidence presented here is most consistent with information sharing within networks and in line with recent findings on the positive effects of ethnic diversity on cellphone coverage (Clochard and Hollard, 2018).

The heterogeneity in and spatial distribution of ethnicities in African countries has been identified as one reason for the relative underdevelopment of African countries (Alesina et al., 2011; Michalopoulos and Papaioannou, 2014, 2016). The contemporaneous borders in Africa were drawn by European colonial powers in the late nineteenth century. As a result, ethnic borders and country borders rarely coincide, which increased ethnic heterogeneity both within and across countries such that the average African country features more than 10 ethnic groups. As most groups speak their own language, preserve their own historical culture, and potentially share a history of conflict, extractive institutions hindered economic development and harmed smaller ethnicities (Acemoglu et al., 2001; Burgess et al., 2015). I complement this view and suggest that the ethnic division into separate countries created linkages that may have alleviated the negative impact by fostering cross-country trade.

The effects of migration in the US and Canada, as well as Asia, have been extensively studied in the trade literature.[2] With the exception of Felbermayr et al. (2010) and Burchardi et al. (2016), however, most studies suffer from endogeneity concerns. If groups migrate, they use and benefit from clusters of existing migration to settle in their new country (Munshi, 2003; McKenzie and Rapoport, 2007; Battisti et al., 2016). Ac-

[2]Gould (1994); Dunlevy and Hutchinson (1999); Herander and Saavedra (2005); Dunlevy (2006); White (2007); Partridge and Furtan (2008); Burchardi et al. (2016) study the US and Canada and Rauch (1999); Rauch and Trindade (2002); Felbermayr et al. (2010); Felbermayr and Toubal (2012) study Asia.

cordingly, if these clusters are formed by exporting firms hiring workers from the destination country, reverse causality issues arise. These endogeneity concerns are of particular importance in developing countries which account for the majority of bilateral migration. Here, cross border linkages may be more important due to considerable barriers to trade, but also harder to measure causally, such that credible empirical evidence using African countries is missing.[3] I extend the analysis of the importance of ethnic linkages for international trade (Bandyopadhyay et al., 2008; Felbermayr et al., 2010) by using exogenously placed borders and pre-independence ethnic heterogeneity to provide causal evidence on the importance of ethnic linkages across countries for bilateral trade flows in Africa. Additionally, by using both past migration stock and ethnic heterogeneity as instruments, I highlight the potential for biased estimates in highly heterogeneous population groups.

The literature on ethnic identification has highlighted several channels and effects of ethnic fractionalization in Africa. Eifert et al. (2010) show that ethnic identification plays an important role in voting, potentially contributing to post-electoral violence in some countries (Dercon and Gutierrez-Romero, 2012). The importance of ethnicity for trust among local market vendors, price dispersion across borders, and public goods provision has been well documented (Fafchamps, 2003; Aker et al., 2014; Burgess et al., 2015). By focusing on price dispersion, these studies offer credible evidence at the micro level but at the cost of generality for the entire African continent. I contribute to the work on ethnic identification and investigate whether it shapes cross-border trade flows between 46 African countries.[4] However, my estimates reflect the effect of migration

[3]Other papers (Peri and Requena-Silvente, 2010; Felbermayr et al., 2010, e.g.) include some African countries but focus on the links to developed countries.

[4]One example of such cross-border solidarity was the temporary practice of Air Namibia, the major carrier of Namibia, having a stopover in Luanda (Angola) only to refuel due to disputes with the fuel supplier at their main airport. The airline is run by an ethnicity that has strong ties between the two countries and, hence, used its credibility in Angola to buy fuel (`http://www.economist.com.na/headlines/`

on formal trade which is likely to be smaller than the effect on informal trade across border regions between African countries.

In summary, this paper highlights a potential positive effect of ethnic division in Africa and potential channels through which these effects materialize. The paper is structured as follows. In Section 3.2, I discuss my empirical strategy and the data. I present the baseline estimates in Section 3.4 and show robustness in Section 3.5. I identify potential mechanisms in Section 3.6 and conclude the paper in Section 3.7.

3.2 Empirical Strategy

In work on bilateral trade, the value of bilateral exports is modeled in gravity type equations. Here, the value of trade is correlated with the size of the exporter and importer economy as larger economies attract more trade flows. In this framework, adding a stock or flow of migrants estimates the impact of migration on bilateral trade. However, estimating the impact of migration on bilateral trade between two developed and two developing countries is distinctly different. While migrants from developed countries often identify themselves by their nationality, ethnic identification dominates nationality in many African countries. Second, emigration from developing countries is correlated with natural, political or economic factors, leading to severe endogeneity concerns. These distinct features of African countries require a generalization of the standard empirical approach as well as exogenous variation to identify a causal effect.

`2795-air-namibia-increases-frankfurt-flights` and `http://hannamibia.com/uploads/pdf/news/130305093441120.pdf`). Additionally, the main supplier of jet fuel in Namibia, Engen, is South African and the ethnicity is only dispersed in Angola and Namibia.

3.2.1 Empirical framework: trade and migration

Estimating the impact of migration between developed countries, the literature uses gravity type equations derived from most theories of international trade. These gravity equations include a population stock or flow of migrants and take the form (Anderson, 1979):

$$\log(X_{ij,t}) = \beta \log(PS_{j,t}) + B_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t} \tag{3.1}$$

Here the log of exports from the exporting country i to the importing country j $\log(X_{ij,t})$ is correlated with the population share of people from i in j $(PS_{j,t})$. Controlling for exporter (δ_i) and importer (δ_j) fixed effects and bilateral characteristics $(B_{ij,t})$, β identifies the effect of the population share on the log of exports. A larger $\beta > 0$ indicates a stronger response of trade flows to changes in the likelihood of an exporter from country i finding someone with her own nationality in country j.

Implicitly, equation (3.1) assumes that migrants to the importing country j have a population share of one in their exporting country i.[5] While approximately true in developed countries, the population structures in developing countries are more diverse. African countries combine a multitude of ethnicities, each with their own identity and separated into multiple countries. Thus, allowing for multiple ethnicities (e) from the set of ethnicities (E) in each country $e \in E_i \subseteq E$, the general form of equation (3.1) is given by

$$\log(X_{ij,t}) = \beta \log\left(\sum_{e \in E_i \cap E_j}^{E} PS_{i,t,e} \times PS_{j,t,e}\right) + B_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t} \tag{3.2}$$

[5]The underlying equation is of the form $X_{ij,t} = (PS_{i,t} \times PS_{j,t})^{\beta} L_{i,t} L_{j,t}$ where $PS_{i,t}$ reflects the share of people in the exporting country, which is unity in the case of equation (3.1), and raw population $(L_{i,t}L_{j,t})$.

where $PS_{i,t,e} \in [0,1]$ is the population of an ethnicity e that is prevalent in each country pair ij, relative to the population of country i at time t. This equation correlates bilateral exports to the probability of a co-ethnic relationship (match) when randomly drawing two individuals from each country. It captures the idea that it is easier to trade with someone from your own ethnicity, but does not exclude the possibility of trading with other ethnicities if the country is prosperous.

The specific formulation of equation (3.2) is supported by two factors. First, it is the empirical equivalent of an otherwise standard model of international trade by (Melitz, 2003; Chaney, 2008), and when amending the cost function of the exporting firm by an ethnicity specific fixed cost that captures lower entry costs into an export market for ethnically connected firms (Appendix 3.A).[6] Second, the interpretation is equivalent to the search and matching literature if an exporter from country i can export more cheaply if she finds an importer in country j that is of the same ethnicity.[7] Aggregating each firm's exports then yields the gravity type equation (3.2). In the search and matching literature, a match is defined when two individuals with the same characteristics are drawn. Since these characteristics are stochastic, the likelihood of a match is given in probabilities. Here, characteristics are distributed along ethnic lines, and thus the fraction of the population representing an ethnicity in

[6] These costs can be lower information costs, more reliable information about market structures or bribes, and fewer cases of fraud between business partners. In the Appendix, I show that equation (3.2) follows if firms face a fixed cost of exporting

$$PS_{i,e}^{-\eta} f_{ij}$$

with $\eta \in [0,1)$ providing concavity for the impact of fixed costs f_{ij} on the exporting firms' profits. These fixed costs represent costs of setting up a distribution network, informing about markets, administration and paying for permits. A similar model has been suggested by Krautheim (2012), and the model nests the established standard model of Chaney (2008) with $\eta = 0$.

[7] With bilateral trade data at the ethnicity level, this equation would be $X_{ij,e,t} = (PS_{i,t,e} \times PS_{j,t,e})^{\gamma}$ with γ being the elasticity. Aggregating to the exporter-importer pair yields $X_{ij,t} = \sum_{e \in E_i \cap E_j}^{E} (PS_{i,t,e} \times PS_{j,t,e})^{\gamma}$. As long as $\gamma \in [0,1)$, the estimated coefficient β in equation (3.2) underestimates γ due to the concavity introduced by γ.

the importing country is equivalent to the likelihood that an exporting firm from the exporting country finds a match in the importing country. Then, the estimated β can be interpreted as an elasticity that captures the change in match probability of each ethnicity when the population changes on either side of the border.[8]

This interpretation is similar to the standard in equation (3.1), as both can be interpreted as a probability of drawing two connected people in each country. In equation (3.2), however, I incorporate the heterogeneous population structures in African countries and allow for a large amount of subgroups within two countries that are connected. Thus, using the standard empirical approach would identify a 'nationality' effect and overstate the true 'ethnicity' specific effect, as it does not account for the variability in the exporting country.

3.2.2 Empirical specification and data

The empirical equivalent I estimate throughout the paper is given by:

$$\log(X_{ij,t}) = \beta \log \text{Ethnic Match Probability}_{ij} + B_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t} \quad (3.3)$$

Here the 'Ethnic Match Probability' is defined by the sum of all ethnic match probabilities for all ethnic groups that are prevalent in both countries $PS_{i,e} \times PS_{j,e}\ \forall e \in E_i \cap E_j$ and constitutes the measure of ethnic similarity between a country pair. Every regression follows the standard in the trade literature and includes exporter (δ_i) and importer (δ_j) fixed effects and, where applicable, includes exporter-importer pair character-

[8]Note that match probability is defined as the likelihood of randomly drawing two individuals from the same ethnicity. The probability that two randomly drawn individuals are not from the same ethnicity is non-zero, but is captured by the exporter and importer fixed effects in the trade equation (3.2).
I explore the possibility of inter-ethnic trade in Appendix 3.A.1. By assuming an increasing cost of trade for ethnicities that are far away from each other, I confirm the baseline estimates for the entire sample of African countries.

istics ($B_{ij,t}$).[9] A positive point estimate, $\beta > 0$, suggests that a larger population on either side of the border for a connected ethnicity yields larger trade flows.

To identify the current population of each ethnicity in each county, I use the Ethnologue data set with estimates on ethnic populations around the world based on a variety of sources.[10] I obtain exogenous variation in ethnic shares using data containing the distribution of ethnic groups before colonialization. The geographic data provided by Murdock (1959) has been used to study the relationship between slavery and trust (Nunn and Wantchekon, 2011). Matching the spatial extent of every ethnicity with grid cell population data from the United Nations Environment Program in 1960, it approximates the population of every ethnicity in every country in 1960, a time when African countries gradually gained independence.[11] To my knowledge, this is the first paper that combines the Murdock data with population data in the context of international migration and trade.

Due to time invariant population figures, the variation at the country-pair level leads to a Moulton (1986) type problem of inconsistent standard errors. I collapse the data to the exporter-importer observation and report standard errors at the country-pair level.[12] In every country pair, every country is observed once as an exporter and once as an importer, to match the data to observed migration flows. For the dependent variable, the log of bilateral exports, I use UN comtrade data from the World Bank

[9]Exporter-importer pair characteristics include log(Length of border), log(Distance between country centroids), dummies for speaking the same language, number of ethnic connections between the country, sharing a colonial history and a dummy that indicates whether parts of the border are determined by a river or mountains.

[10]`www.ethnologue.com` Sources in the data vary in timing and quality.

[11]France retreated from most of its possessions in 1958–1962, Britain in 1957–1965 and Belgium in 1960–1962. The conclusions in this paper are qualitatively robust to very coarse information on population in 1900 contained in Murdock (1959), but due to its incompleteness and the noise I do not report it here.

[12]As this severely reduces the degrees of freedom and to weight observations by their informativeness, I show robustness to weighting every observation with the number of times I observe trade between that pair.

Integrated Trade Systems from 1989–2014.[13] Since the trade data does not capture unreported and informal trade, the literature has focused on price level differences (Aker et al., 2014). I use reported trade only to attempt to estimate the effects for all countries, taking into account that the point estimates are likely lower bounds on the true extent of exports between countries.[14]

The final sample consists of 46 African countries in 91 country pairs with 182 exporter-importer relationships that share a border, observed over 26 years. Due to non-reported trade, the sample is reduced to 3,287 observations. Since the variation I intend to exploit is at the country-pair level, I follow the conservative choice and cluster the standard errors at this level.[15]

3.3 Identification Strategy

The empirical approach in the trade literature uses flows or stocks of migrants and correlates these with bilateral exports. However, economic activity attracts trade and migration flows in a similar fashion, leading to problems of reversed causality. Additionally, borders are not set randomly and reflect spheres of influence and historical economic activity, such that

[13]In order to have a better match, I download import and export data and cross match imports and exports to generate reliable export measures. The results are robust with either inputs, but for sample-size reasons, I end up using the matched data.

[14]If the data is split up into reported or unreported trade, the true estimate will be $\beta = \left(\beta^{reported} X_{ij}^{reported} + \beta^{unreported} X_{ij}^{unreported}\right) / (X_{ij}^{reported} + X_{ij}^{unreported})$. As long as $\beta^{reported} \leq \beta^{unreported}$, I estimate a lower bound effect. Since unreported trade is much more dependent on trust, I argue that this condition is fulfilled.

[15]The final sample leaves out island territories such as Madagascar or São Tomé and Príncipe, as well as Sudan and South Sudan. In the robustness section, I show that coding the missing observations as zero and applying the standard estimating technique by Santos-Silva and Tenreyro (2006) does not change the results. Dyadic data has specific issues with standard errors, as errors can be correlated across country pairs over time. I explore two-way clustering for 46 exporting and 46 importing countries separately in the robustness section.

the direction of a potential omitted variable bias is unclear. I use the historical dispersion of ethnic groups to address the issue of reversed causation and argue that, contrary to borders between European countries, borders between African countries are exogenously placed. Combined, the historical distribution of ethnicities and exogenous placement of borders allows me to identify a causal effect of cross-border ethnicities on bilateral trade between African countries.

In African countries, ethnic population shares are affected by a multitude of factors. Natural catastrophes, hunger, civil conflicts or past migration contribute to the dispersion of people around the continent. Even without accounting for ethnic heterogeneity, these factors are correlated with economic activity and threaten a causal identification of the ethnicity effect in equation (3.3). In addition, if people migrate following a trade route because it constitutes their best information about potential destinations, any factor that increases trade also increases migration, leading to a problem of reversed causality.

The standard approach in the literature uses past migration to instrument for current migration as it has been shown that migrants follow their networks and settle in clusters in the importing country (Munshi, 2003; McKenzie and Rapoport, 2007). This strategy solves the reversed causality problem if initial migrants were randomly placed in countries. For this approach, I use data on bilateral migration at the country level dating back to 1960 to have some exogeneous variation before the time period of interest 1989–2014.[16] To specifically allow for ethnic heterogeneity and counteract any potentially remaining issues of reverse causality and omitted variable biases, I use the pre-colonial distribution of ethnic tribes in Africa (Murdock, 1959). Here, I combine the geographic location of each ethnicity with detailed grid cell population data from the United Nations Environment Program in 1960 to obtain population estimates of migrants and their home population at the time of independence.

[16] http://databank.worldbank.org/data/reports.aspx?source=global-bilateral-migration

Since the Murdock map shows the pre-colonial distribution of 833 ethnicities in Africa, strategic selective sorting into future countries is relatively unlikely. However, the population figures in Murdock (1959) are estimates combined from different sources and given by ethnicity, as opposed to by country, leading to potentially severe measurement error. Hence, I use detailed grid cell population data at a 4.5 km resolution in 1960 which yields a reliable population estimate for the ethnic homelands just prior to independence.

Having a reliable estimate of pre-independence population does not solve the issue of endogeneity. In European countries, borders reflect spheres of interest and were likely set to encompass a homogeneous population. Thus, the population shares of each group in each country are determined by the border, and governments might have had economic reasons to place a border. If a border between two governments was set to include a territory, this might reflect military considerations, but also the interest of having access to potential future markets. Then, population shares and bilateral trade flows are endogenous, and causal inference is problematic.

In Africa, however, the contemporaneous borders between countries were drawn in 1884 at the Berlin conference. These borders do not reflect the interest of the ethnic groups, but the interest of the colonial powers.[17] The exogeneity of these borders has been extensively used in the recent literature on culture and development, price dispersion across borders as well as ethnic fractionalization (Alesina et al., 2011; Aker et al., 2014; Michalopoulos and Papaioannou, 2014). Most country borders today feature parts that follow either latitudinal or longitudinal lines since the exact geography of Africa was largely unknown at the Berlin confer-

[17]For example, Aker et al. (2014) argue that the border between Nigeria and Niger was set at the Berlin Conference in 1884-1885. It was not a border reflecting geographic features but rather the political interests of France and Britain. The border eventually emerged in 1906 and the resulting mixture of ethnicities shows a similar pattern in 2008.

ence.[18] Where the geography was known and country borders could have been set to follow rivers or mountain ridges, the evidence in Figure 2 still suggests no such pattern. Here, country borders, shown in black, rarely overlap with major rivers shown in blue.

I argue that these borders were arbitrarily drawn, split many ethnic groups into two countries, and do not reflect the interests of a specific ethnicity. In my data, all country borders between African countries divide at least one ethnicity. These split ethnic groups are likely to be different from other ethnicities in terms of size or historical economic activity. In line with Michalopoulos and Papaioannou (2013, 2016), I show in Table 1 that an ethnicity is more likely to be split if it is larger in terms of population or territory. However, population density as a measure of economic activity is negatively correlated with split ethnicities.

Supported by the evidence on historical behavior of ethnicities (columns 4-8), these correlations suggest that more widespread, more nomadic and less economically active ethnicities were split. Using data on historical characteristics of tribes, I show that split ethnicities were more likely to be nomadic (column 4), but neither the size of local communities nor historical institutions predict a future divide into more countries. Estimating all characteristics jointly to account for correlations between variables, the area an ethnicity covers in the Murdock data is the only determinant that robustly predicts the divide between countries (column 8).

However, to address concerns that these correlations influence the results, I only consider country borders where ethnicities have been split, and only consider ethnicities that are split at this border in the heterogeneity analysis. Thus, I abstract from a comparison of influential ethnicities with negligible ethnicities and use a balanced sample across similar ethnic groups. Additionally, this procedure abstracts from selection effects into having a shared ethnicity, and focuses on the intensive margin only.

[18] Alesina et al. (2011) show that 80% of African political borders follow either latitudinal or longitudinal lines.

I use historical information on ethnic dispersion to address the issue of omitted variable bias by conflicts, political effects, natural disasters or migration. The historical distribution of ethnicities mitigates the threat of reverse causality if migrants follow trade routes. Furthermore, the use of exogenous borders alleviates the threats posed by the endogenous formation of borders.

3.4 The Impact of Migration on Exports

The positive effects of migration on bilateral trade between African countries have been highlighted in the literature to a great extent. In this section, I provide evidence for a positive effect of migration between developing countries on bilateral trade using two approaches. First, I document the effect of migration on trade using the standard approach with past bilateral migration. Second, introducing baseline ethnic heterogeneity in exporting and importing countries highlights a substantial downward bias as subsequent migration flows are likely to be correlated with initial conditions that affect trade and migration.

Effects on Exports The main results are presented in Table 2 where I report the endogeneous ordinary least squares, the first stage and reduced form using the instrument, and the point estimate from instrumenting. I estimate the impact of the stock of migrants in 1990 on the value of bilateral exports in the period 1989–2014 for the full sample of countries with trade flows. The ordinary least squares point estimate suggests that a 1% higher migrant stock increases bilateral exports by 0.139% (s.e. 0.024). For the standard approach to obtain a causal estimate, I use the migrant stock in 1980 or 1960 which are valid instruments according to the F-Statistics and similar reduced form point estimates. Instrumenting the migrant stock in 1990 with its past values in the last row of Table 2 suggests a small downward bias of the OLS as the IV point estimates range from 0.166 (s.e. 0.027) to 0.199 (s.e. 0.031).

The exclusion restriction in this specification featuring exporting- and importing-country fixed effects and country-pair controls requires that no unobserved country-pair characteristic affects both migration and exports. In standard trade theory, however, large initial income differences reflect productivity differences which increase exports from rich to poor and migration from poor to rich countries as the marginal product of labor is equalized in both countries. In this setting, more migration is negatively correlated with exports and implies a downward bias on the OLS and the IV, when past migrant stocks are used as an instrument.

In column (3), I repeat the exercise for the sample of bordering countries and show that while the point estimates are insignificant due to the lack of power, qualitative conclusions carry over to this narrower defined sample. I introduce ethnic heterogeneity in column (4) and use the constructed initial ethnic match probability as an instrument for migration flows between two bordering countries. Both the reduced form (0.192, s.e. 0.080) and the first stage (0.260, s.e. 0.093) suggest a valid instrument and the resulting IV point estimate suggests that a 1% increase in the stock of migrants increases trade flows by about 0.739% (s.e. 0.417).

The point estimates from the standard approach (column 3) are about 50% smaller than the estimates using the approach that incorporates ethnic heterogeneity to the African context (column 4). The results suggest that migration likely increases trade and ethnic identification plays an important role, even when observing cross-country trade. Similar to well-identified studies on price differences in a narrow setting, ethnic identification also seems to influence the value of exports (Aker et al., 2014).

I conclude this set of results by introducing ethnic heterogeneity in the endogenous variable. By constructing the ethnic match probability using contemporaneous data on ethnic populations in countries, the migration variable now contains information about ethnic heterogeneity (column 5). Here, the OLS suggests that a 1% higher ethnic population in the exporting or importing country increases bilateral exports by 0.232% (s.e. 0.123). Instrumenting the contemporaneous match probabilities with the pre-colonial match probabilities corroborates the results in column (4) with a similar point estimate (0.665, s.e. 0.279).

While the F-Statistic is reasonably low (7.813 and 12.200), the point estimate is unlikely to be affected by a weak instrument problem as this would bias the estimate towards the OLS. The point estimate in columns (4) and (5) is unaffected by the above violation of the exclusion restriction as the constructed ethnic match probability is determined using the pre-colonial distribution of ethnicities and country borders that quasi-randomly displaced parts of an ethnic group in another country. However, as the degree of ethnic connectivity between two neighboring countries may impact bilateral conflicts and politics, other violations of the exclusion restrictions are possible. I investigate this possibility in the mechanisms section, but focus on reduced form estimates in the remainder of the paper.

The results from Table 2 highlight the potential bias when ethnic heterogeneity is not taken into account. Columns (4) and (5) indicate that the elasticity to migration is about 0.7, which is about 2.5 times larger as compared to the results from the conventional approach in columns (1)–(3).

Effects on sub-national development The uniqueness of the African context allows me to infer on the spatial distribution of the gains from trade. As bilateral trade flows, especially in Africa, are usually between capitals or major cities, I aim at identifying whether ethnic homelands benefit from increased exports. However, as bilateral trade is part of the national GDP and hence hard to disentangle, I focus on sub-national gains from trade using nighttime light emissions.

Light emitted in a region has been found to be a valid proxy for regional GDP (Michalopoulos and Papaioannou, 2013; Henderson et al., 2012) and has frequently been used to study sub-national development in Africa.[19] I replicate these findings at the country level in Table 3. Average nighttime light emitted per country is a strong predictor of per capita GDP, controlling for population, country characteristics, and conflicts. Including country fixed effects (column 4), a 1% higher nighttime luminosity is correlated with a 0.472% (s.e. 0.054) higher per capita GDP. Additionally, including year fixed effects confines the variation to the country-year level and decreases the point estimate and increases the noise (column 5).

The results established in the literature and Table 3 motivate to estimate the effects of exports on nighttime luminosity in Table 4. A 1% increase in exports is correlated with a 0.411% (s.e. 0.152) increase in nighttime luminosity of the ethnically connected regions of the exporter. To compare equally densely populated areas, I additionally control for population density and show the first-stage regressions of the ethnic match probabilities today and in the past in columns (2) and (3) for this sample. Both point estimates are less precisely estimated but suggest the same relationship as in Table 2. A larger ethnic match probability is positively correlated with larger exports and more nighttime lights in the ethnically connected region in the exporting country (columns 4 and 5). Instrumenting exports with the ethnic match probability (column 6) suggests that a one percent increase in exports increases sub-national GDP in the ethnic homeland by 1.6% (s.e. 0.635) over a mean of 0.213.[20] Following Henderson et al. (2012) and identifying the time dimension (column 7) yields the same conclusion with a lower point estimate of 0.618 (s.e. 0.200). Compared to the impact of total country exports on nighttime luminosity in the entire country (0.013, s.e. 0.006), the gains from exports

[19] In this study I use nighttime light data from NOAA available under `http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`. The recorded measure ranges from 0–63 and is available from 1992–2010.

[20] With a mean of 0.213 and a range of the nighttime light data of 0–63, there likely exists room for improvements in the border regions. The point estimate without controlling for population density and country-pair controls is 1.890 (s.e. 1.110) and 1.464 (s.e. 0.622), including country-pair controls without population density.

induced by a larger ethnically connected population are overwhelmingly centered inside the ethnic homeland.[21]

The results from Table 2 and 4 indicate the importance of including population heterogeneity when studying the economic effects of migration. The impacts of migration on exports are estimated to be about 2.5 times larger than with the conventional approach. Exports are correlated with increased nighttime luminosity, and the spatial distribution of the gains from trade are likely concentrated within the ethnic homelands.

3.5 Instrument Discussion

For the instrument to recuperate an unbiased estimate of the effect of migration, the exclusion restriction of the instrument needs to be fulfilled. However, as the exclusion restriction cannot be tested, I focus on the reduced form estimates for the remainder of the paper. In this section, I aim at showing that these reduced form estimates are not affected by specification choice, population figures, or the Murdock data itself. Having shown the stability of the instrument, I continue in the mechanisms section to test alternative hypotheses for the question of how ethnic connectivity affects exports.

In Table 5 I present the reduced form estimates in various specifications. The point estimate from the baseline specification (column 1) is robust to including country-pair controls (columns 3 and 4) and adding time varying conflict controls (column 5). To increase the precision of the estimate and put less weight on exporter-importer-pairs with less informational content, I weight the baseline specification with the number of times trade is observed between the exporting and importing country in columns (2), (4), and (5). As expected, the point estimate is unaffected and noise is reduced, resulting in smaller standard errors.

[21]The point estimate is from the regression: $\log \text{nighttime lights}_{i,t} = \beta \log \text{Total exports}_{i,t} + \delta_i + \delta_t + \varepsilon_i$ and is robust to including per capita GDP.

Since standard errors are likely to be correlated within the country pair, I cluster the standard errors at the country-pair level in the baseline specification. However, as shocks might instead be correlated across countries, I report standard errors clustered at the exporting and importing country separately in brackets and conclude that the original clustering is more conservative.

In the remainder of Table 5, I estimate the baseline model using the method suggested by Santos-Silva and Tenreyro (2006). As trade data is likely to be Poisson distributed, I re-estimate the baseline specification using the Pseudo-Poisson-Maximum-Likelihood estimator and show that the point estimates are not statistically different to the baseline.

To identify which country-pair characteristics are predicted by the ethnic match probability, I run the baseline specification using the characteristic as the dependent variable and report the point estimates in Table 6. Country pairs with a larger ethnic match probability share longer borders and the centroids of these countries are closer together. Moreover, these countries have boundaries that are less random as measured by the border fractionalization index of Alesina et al. (2011).[22] However, other relevant characteristics such as the number of ethnic connections, the judicial language, or colonial histories seems to confound the estimate. Estimating all country-pair controls jointly, the F-Statistic on joint significance is 1.572 and only the length of the shared border remains significant at the 10% level.

In the remainder of Table 6, I estimate whether the instrument predicts conflict incidence or severity. Contrary to the literature on ethnic fractionalization within a country, co-ethnic membership across countries does not increase conflict incidences or their severity.

[22]The fractionalization index is derived from a regression varying the size of boxes needed to cover the border: $\log(\text{square count}) = \alpha - \beta \log(squaresize)$. Intuitively, the number of squares needed to cover a straight line can be approximated by square size^{-1}. Any deviation will lead to a number $-1 - \delta$ and a higher fractionalization index.

In Table 7, I control for various measures of the not ethnically connected population to rule out the possibility that alternative measures of population affect the precision of the instrument. Including the log population and the non-ethnic match probability in column (4), the point estimate is virtually identical to the baseline reported in column (1). However, as the population data or the underlying Murdock-maps are potentially incorrectly drawn close to country borders, there may exist ethnicities that are incorrectly coded as split between countries. I drop ethnicities with a population share of less than 1% in columns (5)–(7) and show that while the point estimates increase, they are not statistically different from the baseline.

As a last robustness test, I draw upon the Geo-Referencing of Ethnic Groups (GREG) from the *Atlas Narodov Mira* created by Russian scientists and digitized by Weidmann et al. (2010). It shows the geographic dispersion of ethnicities around the world in 1960 and has been used to study ethnic inequality before. While the two ethnic match probabilities are strongly correlated, the Murdock map captures more ethnicities at an earlier point in time and, hence, more likely to fulfill the orthogonality assumption. However, as the GREG map is closer to independence, it is likely that it captures the ethnic composition more accurately. I repeat the specification tests as well as the IV estimations in Table 8. As the input is closer in time to the independence of countries and features fewer ethnic groups, the point estimates are larger and as robust as the original Murdock input. Furthermore, the IV estimations on exports (column 6) and night time luminosity (column 7) show the same magnitude as the baseline, supporting my initial conclusions.

3.6 Mechanisms

A causal link between migration and exports or economic outcomes is questionable if the initial ethnic match probability has other impacts on conflict, GDP, or government coalitions which, in turn, affect exports or nighttime luminosity. However, two identification decisions support such an interpretation. First, I restrict my analysis to exclusively split ethnicities and, hence, violations of the exclusion restriction must come from an intensive rather than an extensive margin. For example, it has been shown that split ethnicities are more likely to face conflicts and are worse of economically. Here, a violation of the identification assumption would require that larger split ethnicities are differently affected than smaller split ethnicities, and that this difference is correlated with exports. Second, as I include exporter and importer fixed effects in a cross-country regression, the identification relies solely on country-pair variation. For example, the history of conflicts within each country is captured by these fixed effects, leaving country-pair induced conflict variation as the only confounder.

In this section, I provide evidence on how co-ethnic matches can affect bilateral trade by testing four hypotheses that could explain the results. The first hypothesis considers the equator belt where many ethnicities were economically active before independence. In that case, the instrument would only confirm pre-existing trade patterns in sectors that have been trading long before independence. The second hypothesis concerns the effects of conflict in Africa. If cross-border ethnic links reduce conflict incidences, it might raise economic activity and explain the results. The third hypothesis is that ethnic match probabilities only capture similarities in languages which are larger between ethnically connected countries. Hence, it is possible that the main impact of migration on trade is via reducing language barriers. Then omitting linguistic similarity constitutes an omitted variable bias. The fourth hypothesis concerns the literature on the economic effects of institutions in Africa. It is possible that ethnicities with historically stronger institutions are more likely to participate in contemporaneous governments and instead of using their network, use government institutions to facilitate trade.

Pre-existing trade patterns I approach the first hypothesis of trade in preference goods that predates independence from three angles. First, I show that the effect of ethnic matches is stable across all sectors. Then, I document that these ethnic matches increase the amount of goods traded as well as the number of sectors in which they are traded. In the last step, I demonstrate that no country from the equator belt has a large impact on the estimation. Combined, these results indicate that co-ethnic connections do not only reflect trade in preference goods and likely increase the flow of information across countries. This hypothesis is supported by the evidence on the extensive margins of trade, as more sectors are actively trading.

The effect on increased trade might only capture pre-existing trade patterns for habitual goods or document a similarity of preferences for certain goods. Such preference goods were likely traded already before the independence of countries and, hence, the external validity of the impact of migration on exports would be limited. As it is likely that these goods are concentrated in the agricultural sector and not in other sectors, I evaluate the reduced form impact in all sectors. Since Figure 4 shows no differential impact of the instrument in the agricultural sector as compared to other sectors, a preference driven story is not supported by the data at hand. Also, if preference goods are more likely to be traded undetected, the point estimates presented here are lower bounds of the true impact of cross-border ethnicities on economic outcomes.

In contrast, the largest impacts are found in industrial sectors where information and trust are more important than preferences. If these co-ethnic connections do facilitate trade via decreasing information costs, exports should be more diverse in areas with larger ethnic match probabilities. In Table 9, I provide evidence in favor of decreasing information costs using data disaggregated into two-digit industries (SIC-2) for 1989–2014 and four-digit industries (SIC-4) for 2010–2014. Country pairs with larger ethnic match probabilities have more sectors actively trading. Compared to the relevant mean, we observe about 2-3% more sectors actively trading, suggesting that ethnic connections across countries in Africa increase trade at the extensive margin as well.

As the densely populated equator belt contains many ethnicities across many countries which were economically active before colonization, the instrument might reflect this initial economic activity. In Figure 3, I omit individual countries from the baseline specification to estimate their impact in the regression. The low point estimates for Angola and Zambia indicate that these countries have larger effects than the remaining countries. However, since the country borders for Angola are mostly straight lines and neither country belongs to the belt, Figure 3 provides no evidence in favor of a hypothesis based on pre-existing trade patterns.

Since the ethnic match probability affects all sectors and increases the number of sectors trading, a hypothesis based on pre-existing trade routes based on preferences is unlikely. This conclusion is supported by the result that no country from a densely populated area with a rich history of trade affects the point estimates significantly.

Conflict In the second hypothesis, I test whether an increased ethnic connectedness is associated with less conflict, which, in turn, increases the economic activity in Table 10. In addition to the null result on country wide conflict severity when testing for balance (Table 6), I use geocoded conflict data from the Uppsala Conflict Data Program and identify conflict intensity inside the homelands of cross-border ethnicities. Including conflict incidence or conflict severity in the exporter country, importer country, or both jointly, does not affect the point estimate throughout Table 10 and all interaction effects are insignificant, suggesting no heterogeneous effects. Additionally, I find no evidence that the ethnic match probability is predictive for any conflict measure (last row). In sum, the evidence from Table 10 suggests no direct channel of ethnicities affecting conflict and hence exports.

Linguistic similarity In the third hypothesis, I argue that two more ethnically connected countries are likely to have more similar languages and cultures. Then, an increased ethnic match probability might just reflect countries that are trading more because of similarity, and not because of migration or ethnicities. To obtain a measure beyond sharing a judicial language, I use data from Spolaore and Wacziarg (2015) that captures the linguistic similarity between a subset of country pairs in Africa as a measure between zero and one. In this subset, the point estimate is largely unchanged from the baseline with controls (Table 12, column 1), and a larger ethnic match probability is associated with a larger linguistic similarity (column 2). However, including the level effect and the interaction with the ethnic match probability does not affect the point estimate. Again, the interaction is insignificant, suggesting no heterogeneous effects of linguistic similarity.

Combined, the positive effects on the extensive margin of trade and the null-effect on conflict and linguistic similarity suggest that ethnic connections between countries are likely to work by providing information and increasing trust between business partners. However, as government institutions should alleviate problems of mistrust between business partners, the question remains whether these government institutions complement or substitute ethnic connections across country borders.

Government participation and institutions It is entirely possible that governments build on their ethnic connections to foster trade. However, it is equally likely that ethnicities trade with their counterparts in other countries, as they are actively discriminated against or excluded from power. I use data from the Ethnic Power Relations data set to identify the political status of cross-border ethnicities in Africa in Table 12 (Wimmer et al., 2009).[23] Unfortunately, the sample is severely reduced, but the main effect is robust to any definition of political status (columns 2–4). While the political status variable ranges from being discriminated to being the dominant ethnicity (column 2), I vary the definitions to increase power and show robustness (columns 3 and 4). In the raw data, the

[23] I exclusively focus on cross-border ethnicities, as I am interested in their political status, and not the status of an ethnicity that is not split, or irrelevant for this country pair.

impact of being discriminated or being a senior partner is lower than the effect of being a junior party in the omitted category. This suggests that at the one end, ethnicities at least need some economic freedom, and at the other end, ethnic groups use the tools of the government when they are at their disposal. The point estimates suggest that ethnic groups with considerable political power have a 60% lower impact on exports than their counterparts without power (column 4). The decreased point estimate could imply that ethnicities with political power use it to foster trade with other countries, or trade more within their own country. As the former is picked up by the country fixed effects, and the latter is unobservable, I cannot disentangle the two. However, it is clear that ethnicities that are not an influential part of government coalitions are likely to have a large effect on trade.

Whether ethnicities are part of governmental coalitions is likely influenced by their pre-colonial institutions. If an ethnic group had the organizational structure to manage cities and a political system, it is likely to take part in politics and coalitions. I use data from Michalopoulos and Papaioannou (2013) on pre-colonial ethnic institutions and show that while the point estimate is not affected when controlling for such institutions, the interaction effects on political centralization confirm the results from contemporaneous governmental coalitions (Table 13). Historically, more politically centralized ethnicities have a smaller effect on contemporaneous exports, likely because they are more likely to participate in governments.[24] In turn, if these ethnic groups then use government institutions to foster their economic development, this will explain the decreased impact on trade when they are part of government coalitions.

Country pairs with many cross-border ethnic groups are likely to have similar preferences and hence, their governments might work to form pref-

[24]Cross referencing the data from Wimmer et al. (2009) and Michalopoulos and Papaioannou (2013) only leads to 103 ethnicities in 15 countries. More historically centralized ethnicities are 50% more likely to have gained power in the years 1989–2010. The results are not shown.

erential trade agreements which are conducive to trade. Indeed, larger ethnic match probabilities are associated with more preferential trade agreements (Table 14, column 1) which, in turn, facilitate trade (column 2). However, including the endogeneous formation of preferential trade agreements (column 4), or interacting it with the instrument, does not affect the point estimate of ethnic match probabilities. The insignificant interaction suggests no heterogeneous effects of preferential trade agreements, which is consistent with the results on government coalitions, indicating that ethnic groups inside ruling coalitions are more likely to rely on institutions, rather than their connections, to facilitate trade.[25]

Consistent with a model of international trade, where ethnic migration decreases the fixed cost of exporting, a higher likelihood of ethnic matches across country pairs is associated with more trading along the intensive and extensive margin. Every sector of trade benefits, consistent with a hypothesis where ethnic networks across countries decrease information costs or increase the enforcement of cross country contracts. When these ethnicities are part of government coalitions, evidence suggests that institutions act as substitutes to such networks and become more important.

3.7 Conclusion

In this paper, I provide two pieces of evidence that add to our understanding of African economic development. First, I document that the standard approach to estimating the impact of migration on exports is biased when ignoring ethnic heterogeneity. Deriving a simple model of ethnic connectedness across neighboring countries, I document the positive impacts on exports and economic development using nighttime light data.

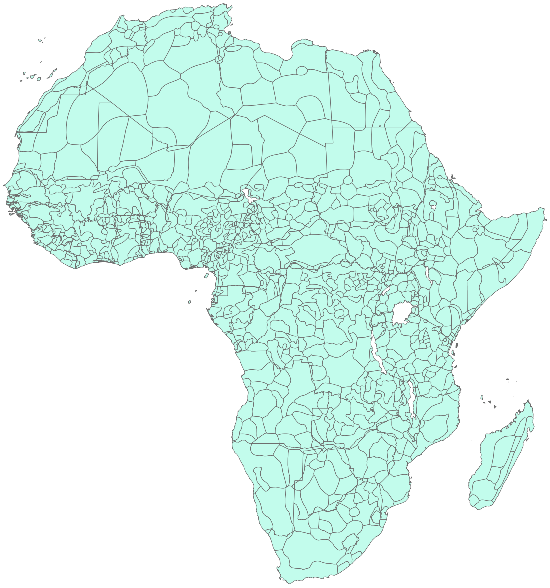
[25]The same conclusion is drawn when using the number of border crossings per country pair as an indicator of the 'willingness to trade'. While increasing trade, the point estimate of the instrument remains unchanged and I find no evidence of an heterogeneous effect.

The second result then concerns the reduced form effect of how ethnic connectedness between countries increases exports. I provide evidence against hypotheses based on preferences, conflict, and linguistic similarity, and show that the most likely mechanism is government exclusion. Ethnicities which are excluded from government participation show the strongest impacts on trade. I argue that ethnicities divert their economic activity to other countries when they are being discriminated against. In line with research on price dispersion (Aker et al., 2014), I argue in favor of information being transmitted and higher within-group trust that facilitates transactions.

In light of the vast amount of research done on the negative development outcomes of ethnicities (Alesina et al., 2016; Michalopoulos and Papaioannou, 2016), this paper provides evidence for a more nuanced view on ethnicities in Africa.

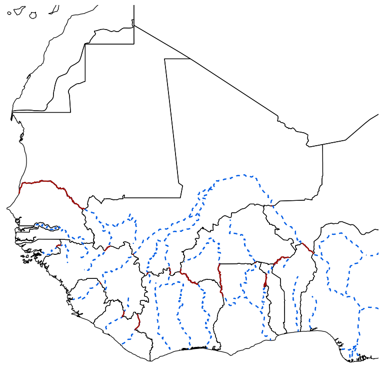
3.8 Tables and Figures

Figure 1: Input data



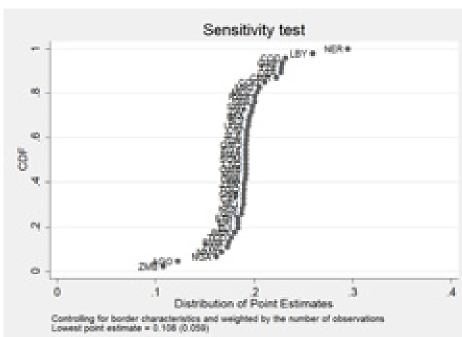
Distribution of ethnicities before colonization (Murdock, 1959).

Figure 2: Rivers as confounders



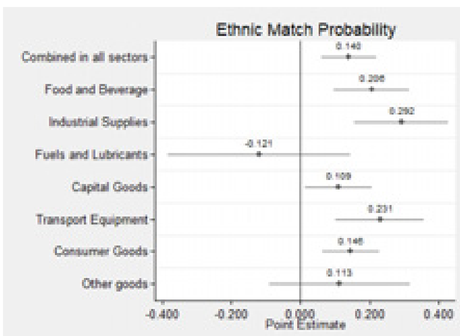
Country borders without rivers (black) and rivers that constitute country borders (red)

Figure 3: Sensitivity Analysis



Point estimates from leaving out individual countries from the baseline specification.

Figure 4: Heterogeneity across sectors



Identifying the impact of the ethnic match probabilities on bilateral trade in various sectors.

Table 1: Determinants of being divided: Historical characteristics of Ethnic groups in Murdock (1959)

	Tribe is divided between two or more countries							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log Population in 1960	0.041***	0.008						0.015
	(0.013)	(0.015)						(0.021)
	[0.011]	[0.011]						[0.017]
log Ethnic Area		0.109						0.138***
		(0.019)						(0.022)
		[0.013]						[0.016]
log Population Density			−0.031**				−0.050***	
			(0.015)				(0.021)	
			[0.011]				[0.014]	
Cities				−0.087			−0.084	−0.046
				(0.055)			(0.059)	(0.060)
				[0.050]			[0.051]	[0.049]
Mean Size of Local Communities					0.013		0.020*	0.004
					(0.012)		(0.011)	(0.011)
					[0.011]		[0.011]	[0.011]
Political Centralization						0.036	0.038	−0.072
						(0.055)	(0.053)	(0.050)
						[0.051]	[0.051]	[0.051]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	833	833	833	441	441	441	441	441
Adjusted R-squared	0.022	0.086	0.014	0.017	0.014	0.011	0.038	0.134

Every column shows the point estimate from a regression on the probability of an ethnicity being divided between two or more countries. Geographic Controls include latitude, longitude, and their product. log Population in 1960 taken from UNEP SIOUX grid cell data. log Ethnic Area is the total expansion area of an ethnicity as given by the Murdock map. Data in columns (4)–(8) taken from Michalopoulos and Papaioannou (2013) and coded as follows. 'Cities': If at least one ethnicity that crosses the border historically had permanent or complex settlements. 'Political Centralization' If at least one ethnicity that crosses the border historically had a jurisdictional level beyond the local level: Centralized Tribe≥2. 'Centralized Tribe' is the count variable of jurisdictional level beyond the local level (range: 0-3). Standard errors corrected for spatial correlation within 500km shown in parenthesis. Lower cutoffs decrease the standard errors to the robust standard errors level shown in brackets. Symbols reflect the significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: The effect of Migration on Bilateral Trade

	log Exports				
	Full sample of African countries		Sample of bordering African countries		
	(1)	(2)	(3)	(4)	(5)
Endogenous variable:	log Stock Migrants 1990				log Ethnic Match Probability, Today
Ordinary Least Squares	0.139***	0.139***	0.147	0.147	
	(0.024)	(0.024)	(0.147)	(0.147)	
					0.232*
					(0.123)
Instrument:	log Stock Migrants 1980	log Stock Migrants 1960	log Stock Migrants 1960	log Ethnic Match Probability	log Ethnic Match Probability
Reduced Form	0.153***	0.140***	0.142	0.192**	0.192**
	(0.025)	(0.023)	(0.095)	(0.080)	(0.080)
First Stage	0.922***	0.704***	0.544***	0.260***	0.289***
	(0.016)	(0.025)	(0.122)	(0.093)	(0.083)
IV Estimate	0.166***	0.199***	0.260	0.739*	0.665**
	(0.027)	(0.031)	(0.157)	(0.417)	(0.279)
Country-pair controls	Yes	Yes	Yes	Yes	Yes
Observations	1902	1902	168	168	168
F-Stat	3532.121	790.770	19.892	7.813	12.200

The samples consist of 45 countries that trade with every other country (columns (1)–(2)) or only their neighbor (columns (3)–(6)). Country-pair controls for the full sample are: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. In the border sample I additionally control for the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Nighttime lights as a predictor for per capita GDP

	log GDP, per capita				
	(1)	(2)	(3)	(4)	(5)
log Average Nighttime lights	0.380^{***}	0.557^{***}	0.553^{***}	0.472^{***}	0.162
	(0.057)	(0.051)	(0.053)	(0.054)	(0.162)
# Ethnicities		−0.003	−0.003		
		(0.008)	(0.008)		
Ruggedness		0.012	0.013		
		(0.055)	(0.055)		
log country area		0.575^{***}	0.574^{***}		
		(0.064)	(0.066)		
log population in 1960		-0.608^{***}	-0.608^{***}		
		(0.098)	(0.099)		
log # Conflicts		0.016	0.018	-0.037^{**}	−0.019
		(0.039)	(0.040)	(0.016)	(0.015)
log # Civilian Casualties		−0.026	−0.021	0.007	0.001
		(0.027)	(0.026)	(0.012)	(0.011)
Country Fixed Effects				Yes	Yes
Year Fixed Effects			Yes		Yes
Observations	893	874	874	893	893

Average Nighttime lights calculated in the period 1992–2010. # determined by the amount of tribes from the Murdock map in a country. Ruggedness taken from Nunn and Puga (2012). Conflicts and Casualties taken from the Uppsala Conflict Database. Standard errors clustered at the country level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The effect of bilateral trade on the economic activity of bordering ethnicities

	log Nighttime lights	log Exports		log Nighttime lights			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log Exports	0.411^{***}					1.664^{**}	0.618^{***}
	(0.152)					(0.635)	(0.200)
log Ethnic Match Probability, Today		0.640		0.812^{*}			
		(0.462)		(0.457)			
log Ethnic Match Probability			0.184^{**}		0.307^{***}		
			(0.079)		(0.110)		
Estimation	OLS	OLS	FS	OLS	RF	IV	IV
Population density in the exporter and importer country	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter×Year fixed effects							Yes
Exporter×Year fixed effects							Yes
Importer ×Year fixed effects							Yes
Observations	168	168	168	168	168	168	2,983
F-Stat						5.428	7.348

Nighttime lights calculated from the period 1992–2010 in the ethnically connected region in the exporting country. I control for log(0.01+population density in 1960) in all specifications following Michalopoulos and Papaioannou (2013). Mean luminosity in the sample is 0.214. Column (1) shows the point estimate from a regression of exports on nighttime lights and column (2) from the regression of the ethnic match probabilities today on nighttime lights. Column (3) is the first-stage estimate using the ethnic match probabilities from the Murdock map to instrument exports via the ethnic match probability today. Columns (4) and (5) then show the reduced forms of the ethnic match probabilities on nighttime lights and column (6) the instrumental variable estimate. In Column (7) I estimate the time varying version and include country×year fixed effects to account for all country-specific variables that might change over time. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: The effect of ethnic matches on bilateral exports: Sensitivity to specification

	log Exports					Exports	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log Ethnic Match Probability	0.135*	0.135***	0.192**	0.175***	0.201***	0.118*	0.127**
	(0.073)	(0.050)	(0.080)	(0.049)	(0.047)	(0.067)	(0.062)
	[0.061]	[0.046]	[0.060]	[0.042]	[0.040]		
Weighted by Number of Observations		Yes		Yes	Yes		Yes
Country-pair controls			Yes	Yes	Yes	Yes	Yes
Conflict Controls					Yes		
PPML estimation						PPML	PPML
Observations	168	3,287	168	3,287	3,287	168	3,287

Every column shows the point estimate from a regression of ethnic match probabilities on exports. In columns (1) and (2) I estimate the unrestricted, plain model with exporter- and importer-country fixed effects. In columns (3)–(7) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. In column (5) I include time varying conflict controls that include the logged amount of conflict, civilian casualties, total deaths, and unknown deaths. In columns (6) and (7) I estimate the Pseudo-Poisson-Maximum-Likelihood method as suggested in Santos-Silva and Tenreyro (2006). In columns (2), (4), (5), and (7), I weight every observation with the amount of positive trade observed in the time span 1989–2014 to put more weight on observations with more informational content. Standard errors clustered at the country-pair level shown in parenthesis and two-way clustered standard errors for the exporter and importer country shown in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Instrument Validity: Covariate Checks

	Mean	Std. Dev.	Observations	β	s.e.
Country-pair controls:					
log Border length	6.405	0.715	168	0.114***	(0.037)
log Distance Centroids	6.052	1.805	168	−0.148*	(0.088)
Border with River	0.601	0.491	168	0.022	(0.022)
# ethnic connections	3.548	1.999	168	0.114	(0.115)
Same Judicial Language	0.161	0.368	168	0.005	(0.029)
Shared Colonial History	0.411	0.493	168	−0.004	(0.018)
log Border Fractionalization	−0.004	0.026	168	0.002**	(0.001)
Border with mountain top ≥ 1000m	0.536	0.500	168	0.021	(0.029)
Border with mountain top ≥ 2000m	0.173	0.379	168	0.010	(0.017)
			F-Stat on joint significance:		1.572
Conflict controls:					
log # conflicts	6.039	1.760	168	0.043	(0.047)
lnCivillian	7.366	2.322	168	0.043	(0.065)
lnDeath	7.156	2.918	168	0.033	(0.083)
lnUnknown	6.731	2.672	168	0.030	(0.066)
		F-Stat on joint significance (incl. border country-pair controls):			1.765

I report β from the regression of the ethnic match probability on the variable in the first column. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Robustness to various population measures and cutoffs

	log Exports						
		Including non ethnic population			Excluding minorities in the country pair		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log Ethnic Match Probability	0.192^{**}	0.196^{**}	0.189^{**}	0.194^{**}	0.268^{*}	0.262^{*}	0.249^{*}
	(0.080)	(0.085)	(0.083)	(0.091)	(0.151)	(0.150)	(0.128)
log Non-Ethnic Population (Exporting Country)		0.161		0.163			
		(0.139)		(0.142)			
log Non-Ethnic Population (Importing Country)		−0.193		−0.192			
		(0.129)		(0.129)			
log Non-Ethnic Match Probability			−0.061	−0.041			
			(0.180)	(0.206)			
Country-pair controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cutoff in Exporter Country:					≥1%		≥1%
Cutoff in Importer Country:						≥1%	≥1%
Observations	168	168	168	168	168	168	168
Country pairs with ethnic connections:					152	154	125

Every column shows the point estimate from a regression of ethnic match probabilities on exports. In columns (1)–(7) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. In columns (2) and (4) I include the log population in the exporting and importing country that is not ethnically connected between the countries. In columns (3) and (4) I construct the Non-Ethnic Match Probability in the same way I construct the main explanatory variable. In columns (5)–(7) I exclude ethnicities who make up less than 1% of the population in the exporting or importing country. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Different input variable: Geo-Referencing of Ethnic Groups (1960)

	log Exports						log Nighttime lights
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log Ethnic Match Probability, 1960	0.353^{***}	0.297^{***}	0.301^{***}	0.246^{***}	0.389^{***}		
	(0.131)	(0.088)	(0.107)	(0.076)	(0.087)		
log Ethnic Match Probability, Today						0.831^{**}	
						(0.361)	
log Exports							0.346^{**}
							(0.133)
Weighted by Number of Observations		Yes		Yes	Yes		
Country-pair controls			Yes	Yes	Yes	Yes	Yes
Conflict Controls					Yes		
IV estimation						IV	IV
Controlling for population density							Yes
Observations	164	3,201	164	3,201	3,201	164	164
F-Stat						9.665	6.543

Every column shows the point estimate from a regression of ethnic match probabilities on exports. In columns (1) and (2) I estimate the unrestricted, plain model with exporter- and importer-country fixed effects. In columns (3)–(7) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. In column (5) I include time varying conflict controls that include the logged amount of conflict, civilian casualties, total deaths, and unknown deaths. In columns (2), (4), and (5), I weight every observation with the amount of positive trade observed in the time span 1989–2014 to put more weight on observations with more informational content. In column (7) I control for population density in the exporter and importer country separately as suggested by Henderson et al. (2012). Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Extensive Margin of Trade: More sectors active in Trading

	# Sectors trading			
	SIC-2, 1989-2014		SIC-4, 2000-2014	
	(1)	(2)	(3)	(4)
log Ethnic Match Probability	0.244***	0.027***	9.705**	0.147***
	(0.060)	(0.008)	(4.037)	(0.033)
# of Ethnic Connections	0.017	0.005	1.685	0.007
	(0.110)	(0.011)	(5.062)	(0.022)
Country-pair controls	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Poisson Estimation		Yes		Yes
Mean # Active Sectors	11.417	11.417	267.016	267.016
Observations	3,287	3,287	925	925

Every column shows the point estimate from a regression of ethnic match probabilities on exports. In columns (1)–(4) I include year fixed effects and the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. In columns (2) and (4) I account for the count structure of the data and use a poisson maximum likelihood estimation. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Heterogeneous Effects: Conflict and the effect of ethnic matches on bilateral exports

	log Exports								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log Ethnic Match Probability	0.149*	0.183*	0.158	0.204**	0.170	0.186	0.153**	0.206***	0.170**
	(0.085)	(0.092)	(0.096)	(0.102)	(0.104)	(0.115)	(0.075)	(0.077)	(0.072)
× Exporter Conflicts	0.019		0.023						
	(0.029)		(0.025)						
× Importer Conflicts		0.004	−0.008						
		(0.037)	(0.037)						
× Exporter Deaths				−0.004		−0.011			
				(0.022)		(0.019)			
× Importer Deaths					0.006	0.012			
					(0.026)	(0.027)			
× Exporter Civilian							0.018		0.020
							(0.018)		(0.018)
× Importer Civilian								−0.007	−0.011
								(0.023)	(0.019)
Level effect included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-pair controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	168	168	168	168	168	168	168	168	168
Exogeneity T-statistic:	−0.423	−0.213		−1.422	−1.213		0.330	0.436	

Every column shows the point estimate from a regression of ethnic match probabilities on exports. In columns (1)–(3) I include the level and interaction effect of conflict incidence inside the ethnic homeland in the exporting- and importing country as defined by the Murdock maps. In columns (4)–(6) I include the level and interaction effect of the amount of deaths inside the ethnic homeland in the exporting- and importing country as defined by the Murdock maps. In columns (7)–(9) I include the level and interaction effect of the amount of civilians dead inside the ethnic homeland in the exporting- and importing country as defined by the Murdock maps. In the last row I present the t-statistic from the regression of ethnic match probability on the relevant conflict variable in the exporting- and importing country. In columns (1)–(7) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Heterogeneous Effects: Linguistic Similarity and the effect of ethnic matches on bilateral exports

	log Exports	Linguistic Similarity	log Exports	
	(1)	(2)	(3)	(4)
log Ethnic Match Probability	0.226***	0.004*	0.222***	0.170**
	(0.075)	(0.002)	(0.075)	(0.067)
Linguistic Similarity			2.356	13.695
			(7.272)	(8.648)
Linguistic Similarity × log Ethnic Match Probability				1.009
				(0.609)
Country-pair controls	Yes		Yes	Yes
Observations	137	137	137	137

Every column shows the point estimate from a regression of ethnic match probabilities on exports. Linguistic Similarity is defined as 1-weighted distance of language as in Spolaore and Wacziarg (2015). Its mean is 0.155 with a standard deviation of 0.107. In columns (1), (3), and (4) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Heterogeneous Effects: Government Participation and the effect of ethnic matches on bilateral exports

	log Exports			
	(1)	(2)	(3)	(4)
log Ethnic Match Probability	0.454***	0.659***	0.676***	0.510***
	(0.112)	(0.178)	(0.177)	(0.142)
× Discriminated		−0.549***		
		(0.197)		
× Irrelevant		−0.120		
		(0.425)		
× Powerless		−0.355**		
		(0.175)		
× Regional Autonomy		−0.200		
		(0.417)		
× Senior Partner		−0.727***		
		(0.166)		
× Dominant Ethnicity		0.147		
		(0.425)		
× Discriminated			−0.430**	
			(0.206)	
× Minority or Powerless			−0.335**	
			(0.166)	
× Senior Partner, Regional Autonomy, or Dominant			−0.524***	
			(0.175)	
× Senior Partner, Regional Autonomy, or Dominant				−0.294***
				(0.108)
Country-pair controls	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,282	1,282	1,284	1,284

Every column shows the point estimate from a regression of ethnic match probabilities on exports. Here I match the Murdock ethnicities to the ones reported in the ethnic power relations data set (Wimmer et al., 2009) and only use county-pairs where I could match at least one ethnicity. In column (1) I estimate the baseline model in this sub-sample. Columns (3)–(4) are variations of Column (2) with broader definitions of government relations to increase power. The omitted category in columns (2) and (3) is 'Junior Partner' and in columns (4) 'Discriminated', 'Irrelevant', 'Powerless', and 'Junior Partner' In columns (1)–(4) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Heterogeneous Effects: Historical institutions and the effect of ethnic matches on bilateral exports

	log Exports				
	(1)	(2)	(3)	(4)	(5)
log Ethnic Match Probability	0.182*	0.183*	0.197**	0.249***	0.242**
	(0.093)	(0.092)	(0.088)	(0.092)	(0.099)
Ethnicities had Cities	0.262	0.243			0.297
	(0.690)	(1.166)			(1.145)
× Ethnicities had Cities		−0.003			0.026
		(0.189)			(0.194)
Political Centralization			−0.178	−1.882**	−1.884**
			(0.513)	(0.877)	(0.899)
× Political Centralization				−0.212**	−0.207**
				(0.093)	(0.099)
Country-pair controls	Yes	Yes	Yes	Yes	Yes
Observations	164	164	164	164	164

Every column shows the point estimate from a regression of ethnic match probabilities on exports. Data taken from Michalopoulos and Papaioannou (2013) and coded as follows. 'Ethnicities had Cities': If at least one ethnicity that crosses the border historically had permanent or complex settlements. 'Political Centralization' If at least one ethnicity that crosses the border historically had a jurisdictional level beyond the local level: Centralized Tribe≥2. 'Centralized Tribe' is the count variable of jurisdictional level beyond the local level (range: 0-3). In columns (1)–(5) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Heterogeneous Effects: Preferential Trade Agreements

	PTA	log Exports			
	(1)	(2)	(3)	(4)	(5)
log Ethnic Match Probability	0.018*		0.222***	0.218***	0.196***
	(0.009)		(0.045)	(0.044)	(0.038)
Preferential Trade Agreement		0.403*		0.262	1.014*
		(0.235)		(0.200)	(0.538)
×Preferential Trade Agreement					0.116
					(0.073)
Country-pair controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	3,012	3,012	3,012	3,012	3,012

Every column shows the point estimate from a regression of ethnic match probabilities on exports using the panel dimension of the data. In columns (1)–(5) I include the following country-pair controls: Whether the country pair shares a border, the same judicial language or a common colonial tie. Distance between the country pair and the amount of ethnic links between the country pair are included. Characteristics of the border feature the length of the border as well as whether the border contains a river or a mountain top above 1000 or 2000 meters. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.A Technical Appendix

In this section, I derive a model of international trade with firm and ethnic heterogeneity to provide a motivation for the main estimation equation (3.2). My framework draws on Chaney (2008) and nests the standard model while retaining tractability.

The economy consists of N countries which contain a subset $e \in E$ of predefined ethnicities. Not every ethnicity is present in every country. Furthermore, every economy produces a homogeneous composite good q_0, as well as horizontally differentiated goods $q(\omega)$. Any firm of ethnicity $e \in E$ producing a heterogeneous good $\omega \in \Omega$ from country $i \in N$, uses its ethnic counterpart $e' \in E$ in country $j \in N$ to maximize the expected profits from selling in market $j \in N$ according to:

$$\pi_{ij,ee'}(\omega) = p_{ij}(\omega)q_{ij}(\omega) - c_{ij,ee'}(\omega) \tag{3.4}$$

Where the price of a good $p_{ij}(\omega)$ is country specific, as is the demand for a good $q_{ij}(\omega)$.[26] $\tau_{ij} > 1$ represent variable trade costs, denoted as "iceberg trade costs". A firm needs to produce τ_{ij} goods in order to sell one unit in country j. The cost of producing a good $c_{ij,ee'}(\omega)$ is assumed to be ethnic dependent in home e and foreign e' and of the form:

$$c_{ij,ee'}(\omega) = \frac{\tau_{ij}}{\varphi} q_{ij}(\omega) + \left(\frac{L_{j,e'}}{L_j}\right)^{-\eta} f_{ij} \tag{3.5}$$

Here, φ denotes productivity which every firm draws from a Pareto distribution $G(\varphi) = 1 - \varphi^{-\gamma}$.[27] γ represents the degree of firm heterogeneity, with increasing values denoting decreasing firm heterogeneity. Firms learn

[26] Although Aker et al. (2014) show that ethnicities affect the prices between two countries, I assume that this is a result of a supply or demand shock. However, including a demand shock here would create a simple demand shift in the gravitiy equation. Alternatively, one could divide the product space into goods consumed by ethnicities which would yield a result similar to including different sectors.

[27] Following the literature standard I use the Pareto distribution as it mirrors the empirical distributions well (Axtell, 2001) and is notational convenient.

about their productivity when drawing from $G(\varphi)$ and, subsequently, decide to pay country pair specific fixed costs f_{ij} in order to serve market j.[28] These fixed costs are mitigated by the fraction of the population in country j that is of the same ethnicity $e' = e \in E$ as the owner of the firm.[29] I call the effect of the fraction $\left(\frac{L_{j,e'}}{L_j}\right)^{-\eta}$ the network effect of ethnic ties. This fraction lies within the unit interval and raised to the power of $\eta \in \left[0, \frac{\sigma-1}{\gamma}\right)$ that gives the importance of ethnic networks in decreasing the fixed costs of exporting. It can be interpreted as a decreased costs of acquiring information about the market structure in the destination country or market demand. Alternatively, its interpretation permits lower payments to government officials because of ethnic ties or it serves a proxy for the general trust-worthiness of a society. Empirical evidence by Grossman et al. (2006) suggests that factors like cultural distance and institutional development are particular relevant for the fixed cost of exporting. Ethnic networks should then be beneficial when firms try to circumvent bureaucratic hurdles. The larger the hurdles, the larger should be the impact of ethnic networks.

In every country, households maximize their utility according to:

$$U = q_0^{1-\mu} \left(\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}\mu} \tag{3.6}$$

That is, they consume a freely traded homogeneous good q_0 and consume every available variety of the heterogeneous good ω. The share of income spent on the heterogeneous good is given by μ and the elasticity

[28] The cost of producing a good are wages times $c_{ij,ee'}(\omega)$. Due to the production in the freely traded homogeneous good q_0 wages in both sectors are normalized to unity to simplify the expressions. Furthermore, since there are infinitely many possible firms of each ethnicity, I can characterize the costs of producing variety ω simply by the ethnicity and the productivity of the firm φ.

[29] A similar approach has been undertaken by Krautheim (2012) where the fraction is the number of domestic firms active in the destination market. In the following, I assume that every ethnicity has at least one member in every country. I can relax this assumption and assume that there is an additional fixed cost to pay when dealing with non co-ethnic members. The results are robust.

of substitution is given by $\sigma > 1$. Standard results lead to a pricing of $p_{ij}(\varphi) = \frac{\sigma}{\sigma-1}\frac{\tau_{ij}}{\varphi}$ and a demand:

$$q_{ij}(\varphi) = p_{ij}(\varphi)^{-\sigma} P_j^{\sigma-1} \mu \left(1 + \frac{\Pi}{L}\right) L_j. \tag{3.7}$$

Here, $\left(1 + \frac{\Pi}{L}\right) L_j$ denotes the fraction of world capital Π and labor L income that belongs to country j.[30] Hereof, a fraction μ is spend on heterogeneous goods. Combining the profit function, pricing and demand yield the ethnicity dependent productivity cutoff above which firms start to export due to non-negative profits $\pi_{ij,ee'} \geq 0$:

$$\varphi^*_{ij,ee'} = \left(\frac{\sigma}{\sigma-1}\right) \frac{\tau_{ij}}{P_j} \left[\frac{\mu}{\sigma}\left(1 + \frac{\Pi}{L}\right) L_j\right]^{\frac{1}{1-\sigma}} \left(\frac{L_{j,e'}}{L_j}\right)^{\frac{\eta}{1-\sigma}} f_{ij}^{\frac{1}{\sigma-1}} \tag{3.8}$$

The price index P_j can be solved explicitly by summing all prices from all exporting countries together, taking their productivity cutoffs into account.[31] Then, the productivity cutoff can be expressed in terms of primitives:

$$\varphi^*_{ij,ee'} = \left[\frac{\gamma}{\gamma - (\sigma-1)}\right]^{\frac{1}{\gamma}} \left[\frac{\mu}{\sigma}\left(1 + \frac{\Pi}{L}\right)\right]^{-\frac{1}{\gamma}} L_j^{\frac{\eta-1}{\gamma}} \frac{\tau_{ij}}{\theta_j} f_{ij}^{\frac{1}{\sigma-1}} (L_{j,e'})^{\frac{\eta}{1-\sigma}} \tag{3.9}$$

As in Chaney (2008), the total foreign population decreases the cutoff due to market size effects $L_j^{\frac{\eta-1}{\gamma}}$. This effect is dampened by $\frac{\eta}{\gamma}$ because the ethnic population has a stronger effect on the cutoff than the total pop-

[30] Due to the sector that produces the homogeneous goods, wages are driven down to unity.

[31] $P_j = \left(\sum_{k=1}^{N} L_k \sum_{e\in E} \int_{\varphi^*_{kj,ee'}}^{\infty} \left(\frac{\sigma}{\sigma-1}\frac{\tau_{kj}}{\varphi}\right)^{1-\sigma} dG(\varphi)\right)^{\frac{1}{1-\sigma}}$.

ulation.[32] θ denotes the multilateral resistance term that approximates how distant a market is in comparison to all other markets.[33] Equation (3.9) suggests that much of the ethnic network effect will work through the extensive margin of trade. If the fixed costs of exporting are higher due to corruption, the cutoff for ethnically connected and non-connected firms increases, but to a lesser extent for the former group.[34]

In order to obtain a testable equation, I aggregate individual demand[35] to an network extended gravity equation:

$$X_{ij} = \mu \left(1 + \frac{\Pi}{L}\right) L_j f_{ij}^{\frac{\sigma-1-\gamma}{\sigma-1}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{-\gamma} \sum_{e \in E_i \cap E_j} L_{i,e} (L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}} \qquad (3.10)$$

Total exports between any pair of countries increase in market size $\mu\left(1 + \frac{\Pi}{L}\right) L_j$ and multilateral resistance θ and decrease in variable trade cost τ_{ij} and fixed costs f_{ij}. The network term is increasing the total trade flows since $\nu \equiv \frac{\eta(\sigma-1-\gamma)}{1-\sigma} \in [0,1)$ in order to obtain interior solutions for

[32]The original cutoff in Chaney (2008) can be recovered by setting $\eta = 0$. The effect of the foreign ethnic population is greater since $\frac{\eta}{\gamma} < \frac{\eta}{\sigma-1}$ due to the assumption $\gamma > \sigma - 1$ that guarantees interior solutions.

[33]$\theta_j = \left[\sum_{k=1}^{N} f_{kj}^{\frac{\sigma-1-\gamma}{\sigma-1}} \tau_{kj}^{-\gamma} \sum_{e \in E} L_{k,e} (L_{j,e})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}\right]^{-\frac{1}{\gamma}}$. A popular example is the comparison between Portugal and Spain with New Zealand and Australia. Similar in terms of GDP, the latter trade relatively more with each other due to their distance to all other markets in the world.

[34]Putting it differently, in a world where all the fixed cost consist of corruption and trust, the ethnic networks are paramount to exporting. We should observe only ethnically connected firms. A similar exercise can be done by changing the cost function into a part which is ethnic dependent (trust and corruption) and a part that is non ethnic dependent. Then ethnic networks do not matter when there is no ethnic dependent fixed costs, but matter a lot when there is no non ethnic dependent fixed cost.

[35]$X_{ij} = L_i \sum_{e'=e \in E} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi)$, where $\frac{L_{i,e}}{L_i}$ is the ethnic fraction in country i. An alternative summation would be to include the non ethnic population in foreign and their cutoffs: $X_{ij} = L_i \left[\sum_{e \in E_i \cap E_j} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi) + \sum_{e' \neq E_i \cap E_j} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi)\right]$. The second term would be condensed to the part in Chaney (2008).

the system of equations.[36] If the number of ethnicities is greater then the number of countries, the system of equations is under-identified and individual parameters in ν cannot be identified. A way around is to assume specific values for ν and conduct sensitivity analyses. Specifically, if ν takes on the value one, the ethnic network variable leads to a search and matching interpretation and gives the likelihood that two randomly selected firms from both countries are of the same ethnicity, when controlling for population size.

The introduction of ethnic heterogeneity in the framework of Melitz (2003) and Chaney (2008) introduced a second source of heterogeneity that creates a particular feature regarding export decisions. Firms owned by an ethnic minority might first export to other markets and only later serve their home market. This feature is similar to capital-constraint firms that cannot export in Chaney (2016) and implies imperfect selection into exporting. Firms that export might have lower productivity than firms that do not and, thus, create welfare losses.

The empirical equivalent of this equation is given by:

$$\log(X_{ij,t}) = \beta \log\left(\sum_{e\in i\cap j}^{E} L_{i,e}(L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}\right) + B_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t}$$

Since the importer and exporter fixed effect also capture population in each country and $(L_j \times L_i)^{-1} = -\log L_j - \log L_i$ one can rewrite the equation as:

$$\log(X_{ij,t}) = \beta \log\left(\sum_{e\in i\cap j}^{E} \frac{L_{i,e}}{L_i} \times \frac{(L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}}{L_j}\right) + B_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t}$$

[36] I further require that $\gamma > (\sigma - 1)$ and $\eta < \frac{(\sigma-1)}{\gamma}$ to guarantee an interior solution.

which as $\frac{\eta(\sigma-1-\gamma)}{1-\sigma} \rightarrow 1$ approaches equation (3.2). This equation can be interpreted as a search and matching model, where the population in the importing country has to incur a penalty, thus needs a larger population to have the same effect on trade as the exporting population.

3.A.1 Inter-ethnic Trade

So far I assumed that connections can only exist within ethnicities and neglected the possibilities of inter-ethnic connections. Here, I relax this initial assumption and assume that every ethnicity has an implicit (weak) ranking of every other ethnicity. Then, for every ethnicity I can order the other ethnicities according to the cost they have to incur in order to conduct business with them. This cost is similar to the fixed costs discussed earlier, in the sense that it reflects learning costs between ethnicities. Therefore, I assume there exists a matrix $F_{E\times E}$ that reflects this ordering between every possible combination of ethnicities. The cost of producing and exporting are then given by:

$$c_{ij,ee'}(\varphi) = \frac{\tau_{ij}}{\varphi} q_{ij}(\varphi) + \left(\frac{L_{j,e'}}{L_j}\right)^{-\eta} f_{ij} f_{ij,ee'}$$

with $f_{ij,ee'}$ being an element from $F_{E\times E}$. Here bilateral fixed costs are disentangled from ethnic specific cost. Every firm has to incur bilateral fixed costs to set up the firm, but also have to invest in ethnic relations in order to mitigate the additional ethnic specific fixed costs.[37] The gravity equation is then given by:

$$X_{ij} = L_j \mu \left(1 + \frac{\Pi}{L}\right) f_{ij}^{1-\frac{\gamma}{\sigma-1}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{-\gamma} \sum_{e\in E\cap E'} L_{i,e} (L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}} f_{ij,ee'}^{1-\frac{\gamma}{\sigma-1}}$$

[37] The basic model is a special case of this case where the off diagonal elements of $F_{E\times E}$ are assumed to be so high that only within ethnicity connections can occur.

Now, the effect of ethnic match probabilities is not only measured within ethnicities, but also between ethnicities. If the fixed costs of creating ties between ethnicities are low enough, this specification should fit the data better. Combining the findings on the extensive margin formulation and the ethnic specific fixed costs, ethnicities have a two fold effect on trade flows. They increase the number of firms exporting in distrustful environments by affecting the extensive margin. However, trade volumes between two countries are negatively affected by the ethnic specific fixed costs. Then if these fixed costs represent trust or corruption issues, the above model puts a strong emphasis on reducing corruption and increase trust among ethnicities.

Table 15: Inter ethnic networks: Using the distance between ethnicities to proxy for the cost it takes to create trust

	log Exports							
	Border Sample				Entire Africa			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log Stock Migration 1990	0.096			0.436*	0.139***			0.692***
	(0.126)			(0.228)	(0.024)			(0.023)
log Distance weighted Match Probability		1.192**	2.734***			2.067***	2.986***	
		(0.593)	(0.617)			(0.071)	(0.079)	
Country-pair controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV estimation				IV				IV
Observations	168	168	168	168	1,902	1,902	1,902	1,902
F-Test				19.651				1421.488

Country-pair controls included: Distance between the country pair and the amount of ethnic links between the country pair are included. I additionally control for sharing a colonial history, same language, or a border. Standard errors clustered at the country-pair level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Zusammenfassung

Die vorliegende Dissertation besteht aus drei selbstständigen Aufsätzen die zentrale Fragen der Entwicklungsökonomie und der politischen ökonomie behandeln. Jeder Aufsatz behandelt ein separates Problem in heutigen Entwicklungsländern und sucht nach Beispielen in der Vergangenheit um dieses zu lösen. Wie müssen die Ressourcen verteilt sein damit auch zukünftige Generationen diese nutzen können, ohne dass heutige Generationen darunter leiden müssen? Wie reagieren Menschen auf gewalttätige Erfahrungen in Konflikten, und welchen Einfluss haben diese auf die Gesellschaft? Wie sehr beeinflussen Migration und ethnische Vielfalt die ökonomische Entwicklung Afrikas? In dieser Arbeit versuche ich unser Verständnis für diese Fragen anhand von neuen Daten, Methoden und ökonometrischer Schätzverfahren zu vergrößern um den politischen Diskurs mit Informationen zu fundieren.

Property Rights, Resources, and Wealth Im ersten Kapitel dieser Thesis beschäftige ich mich mit der Frage, wie Eigentumsrechte definiert sein sollten um Ressourcen bestmöglich zu verteilen und den Wohlstand anzuheben. Besonders im heutigen Afrika hat die voranschreitende Privatisierung zu einer Knappheit von Land geführt, welches vorher von der dortigen Landbevölkerung genutzt wurde. Traditionell werden Ressourcen wie Acker- oder Weideland oftmals durch Gewohnheitsrecht reguliert. In diesem Fall sind die Profitabilität und die langfristigen Landnutzungsrechte dadurch bestimmt, wie lange jemand eben jenes Land bereits bewirtschaftet.

Durch das Bevölkerungswachstum müssen sich immer mehr Landwirte das vorhandene Weideland teilen, welches dann durch die resultierende überweidung immer unproduktiver wird. Diese Situation ist allgemein als das 'Tragedy of the Commons' (Allmende Problematik) bekannt. Es existieren zwei Lösungen um eben jenes Problem zu lösen, die in der Wissenschaft diskutiert werden (Hardin, 1968). Der erste Vorschlag zur Lösung der überweidung ist, dass das Weideland vollständig privatisiert wird. Dieser Lösungsansatz beruht auf der Annahme, dass in einer Situation ohne Transaktionskosten in der jede Person Land kaufen kann, jeder das Land bekommt was seine Produktivität maximiert. Die daraus resultierende Verteilung des Landes ist theoretisch gesehen optimal. Das Gute an diesem Lösungsansatz ist, dass die ursprüngliche Verteilung des Landes keine Rolle spielt, weil jedes Stück Land von jedem Landwirt gekauft werden kann. Wenn aber nur wenige Personen alles Weideland besitzen, es Transaktionskosten gibt, oder sich manche keinen Kredit für ihr Land leisten können, ist diese Lösung nicht länger optimal. Deswegen beruht der zweite Lösungsansatz darauf, das Land in staatlicher Hand zu lassen. Hier verpachtet der Staat das Weideland zu festen Konditionen an Landwirte, welche dann im Rahmen ihres Vertrag das Land optimal nutzen.

Da in beiden Lösungsansätzen die Anreize für die Landwirte so gesetzt sind, dass sie das Land nicht ausnutzen, sondern optimal bewirtschaften, ist das Ergebnis, theoretisch gesehen, das gleiche. Unter der Annahme, dass jeder Landwirt einen Kredit bekommen kann und Weideland frei gehandelt wird, wird das Weideland unter beiden Lösungsansätzen optimal bewirtschaftet. Leider ist dies nur ein theoretisches Resultat und nicht empirisch fundiert, da das beste Weideland oft schon lange in privater Hand ist und niemand bereit ist das schlechtere Land zu kaufen. Der Vergleich von privatisiertem Land und staatlichem Besitz ist unter diesen Gegebenheiten problematisch, weil sich die unterliegende Qualität des Landes unterscheidet. Deswegen hat die empirische Forschung auch keine Antwort auf folgende wichtige Fragen: Sollte ein Staat sein Weideland privatisieren oder im staatlichen Besitz belassen und verpachten?

Welche Lösung ist besser für die Produktivität der Ressource und welche ist besser für den Landwirt?

Ich beantworte eben diese Fragen anhand einer großen Landreform in den Vereinigten Staaten. Im Jahr 1934 wurde eine Behörde gegründet, die mehr als 600.000 Quadratkilometer Land an Landwirte verpachtete. Die Landwirte nutzen das Land bereits vor der Reform um ihre Kühe dort zu weiden. Die Zugangs- bzw. Besitzrechte waren jedoch nicht geregelt, was zur überweisung des Weidelandes führte. Um dieser Entwicklung entgegenzuwirken beschloss die Zentralregierung in Washington D.C., Teile des Weidelandes an die Landwirte zu verpachten. Gleichzeitig wurde bis 1934 Land zu fixen Konditionen verkauft. Diese Konstellation erlaubt es mir erlaubt Weideland, welches 1934 privatisiert wurde, mit verpachtetem und frei nutzbarem Weideland in der Nähe zu vergleichen. Da diese Ländereien nur durch einen Zaun getrennt wurden, ist davon auszugehen, dass die unterliegende Bodenqualität dieselbe ist und alle heutigen Unterschiede durch die unterschiedlichen Besitzrechte hervorgerufen sind. Da die Anzahl der Kühe davon abhängt wie viel Gras eine Weide hat, nutze ich moderne Satellitenaufnahmen um die Vegetation auf beiden Seiten des Zaunes zu vergleichen.

Meine Resultate bestätigen die alten Theorien von Coase (1960) und Samuelson (1954). In Gegenden in die Polizei Gesetze auch durchgesetzt hat, gibt es keinen Unterschied zwischen Weideland welches privatisiert oder verpachtet wurde. Beide Formen haben heute mehr Vegetation als freies Weideland in unmittelbarer Nähe, welches also unter 'Tragedy of the Commons' (Allmende Problematik) litt. Jedoch haben Landwirte mit Zugang zu verpachten Weideland heute mehr Einkommen und Besitz, als Landwirte die sich Land kauften. Dieses Resultat kommt allerdings vor allem dadurch zustande, dass die ursprünglichen Pachtrechte weiterverkauft werden konnten.

Um den Kreis zu schließen komme ich zu meiner Ausgangsfrage zurück: wie sollten Eigentumsrechte definiert sein um Ressourcen bestmöglich zu verteilen und den Wohlstand zu vergrößern? Die Ergebnisse meiner Studie zeigen, dass Privatisierung dann zu den besseren Resultaten führt, wenn Zugangs- und Besitzrechte nur unzureichend umgesetzt werden können. Sofern jedoch, wie in vielen Entwicklungsländern, Landwirte keinen Zugang zu Krediten haben oder es hohe Transaktionskosten gibt, kann Verpachtung effizienter und besser sein. Für einen Landwirt, der freies Weideland benutzt, welches er mit anderen teilen muss, unterscheidet sich ein Pachtvertrag mit klaren Konditionen und Preisen nur geringfügig von Eigentum. Wenn dieses Eigentum von Banken und potentiellen Käufern anerkannt wird, erhöht das den Bürgschafts- oder Verkaufswert des Betriebs und der Landwirt hat mehr Handlungsspielraum das zu tun, was optimal für ihn ist. In solch einer Situation tritt das Coase-Theorem in Kraft und die Verteilung des Landes ist nach Abschluss aller Transaktionen optimal. Wenn es nun doch optimal sein sollte, dass ein einzelner alles Land besitzt, wird am Ende aller Transaktionen auch diese Aufteilung entstehen, unabhängig davon ob das Land privatisiert oder verpachtet wurde. Der einzige Unterschied ist das jetzt Landwirte ihr Land verkaufen und entschädigt werden, und nicht der Staat.

State Repression, Exit, and Voice Das zweite Kapitel behandelt eine grund- legende Frage im Zusammenhang mit gewalttätigen Konflikten. Wie reagieren Menschen auf Gräueltaten? Anstelle der quantitativen Frage ‘wie sehr sie reagieren’ kann man auch eine tiefer gehende Frage stellen: Wie erwarten wir, verändert sich eine Bevölkerung die einen Genozid erlebt hat? Es ist einerseits möglich, dass die Menschen sich stärker der eigenen Gruppe zuwenden, wenn ein externer Aggressor diese Gräueltaten vollzieht, da eine starke Einheit zukünftige Aggressoren abwehren könnte. Andererseits, wenn diese Gräueltaten von Mitgliedern der eigenen Gruppe begangen werden, könnte das das Vertrauen in die Gruppe zerstören und in ein starkes Mistrauen gegenüber Führungspersönlichkeiten aus der gleichen Gruppe hervorrufen. Ausgangspunkt der zweiten Analyse ist einer der schlimmsten Genozide der modernen Geschichte in dem fast 30% der Bevölkerung Kambodschas umgebracht wurden. Als die Khmer Rouge 1975 an die Macht kamen, wurden Menschen aufgrund ihrer landwirtschaftlichen Nützlichkeit eingeteilt.

Menschen in Städten sowie Intellektuelle wurden in Arbeitslager verschleppt um dort entweder durch Arbeit zu sterben oder hingerichtet zu werden.

Da die kambodschanische Gesellschaft auf einer starken Hierarchie aufgebaut war, welche sich durch ein starkes Vertrauen in die Führungspersönlichkeiten ausdrückte, gab es kaum Widerstand gegen diese Gräueltaten. Da diese Arbeitslager in den ländlichen Kommunen errichtet worden sind, ist unser Startpunkt dann: Wie haben sich die Anwohner, welche nicht im Arbeitslager waren, verändert? Hierzu entwickeln wir ein Model, welches jeder Person die Wahl zwischen, wählen gehen, die lokale Gemeinschaft zu meiden, oder einfach loyal zu ihren Anführern zu sein.

Unsere empirischen Resultate zeigen, dass sich Menschen, die sich diese Gräueltaten ansehen mussten, stark verändert haben. Anstelle ihren Anführern blind zu vertrauen, informieren sie sich selbst und formen ihre eigene Meinung zu den Vorgängen im Land. Dies führte zur Ablehnung von starken Führungspersönlichkeiten und der stärkeren Unterstützung von demokratischen Werten, vor allem bei Wahlen. Abgesehen davon zieht sich diese Gruppe von Menschen aus dem öffentlichen Leben zurück, wahrscheinlich um nicht aufzufallen, da dies während des Genozids die beste Strategie war um nicht aufzufallen. Ein Nebeneffekt ist ein ebenfalls gesunkenes Vertrauen in Nachbarn und Beamte. Dieses führt wiederrum zu einer verringerten Abholzung der Wälder. Da Rodung auch in Kambodscha illegal ist, müssen sich alle Beteiligten einigen sich nicht gegenseitig zu verraten und Beamte, sofern notwendig, gemeinschaftlich bestochen werden. Da der Genozid einen negativen Einfluss auf diese hierarchische Gesellschaft hatte, in welcher der Staatsapparat bestochen werden konnte, nahm die Abholzung der Wälder über die Jahre hin ab. Korruption ist in hierarchischen Gesellschaften einfacher, jedoch ist sie in einer Umgebung von niedrigerem Vertrauen und einer niedrigen Partizipation im täglichen Leben schwieriger umzusetzen.

Migration and Ethnic Heterogeneity Im dritten Kapitel stelle ich die alte, aber wieder aktuelle, Frage, ob Migration gut für die Gesellschaft ist und versuche dabei eine Lücke in der Forschung zu schließen. Während wir viele Beispiele für die positiven Effekte von Migration zwischen 'reichen' Ländern haben, wissen wir wenig über Migration zwischen Entwicklungsländern, obwohl sie die häufigste Art von Migration ist.

In diesem Kapitel liegt der Fokus auf Exporten zwischen afrikanischen Ländern, die einerseits den Großteil der Migration zwischen Entwicklungsländern stellen, andererseits zwei große Besonderheiten aufweisen, die die Analyse erschweren. Erstens sind sowohl das Exportvolumen also auch die Migrationsströme aufgrund unzureichender Datenerfassung nur annäherungsweise messbar. Zweitens ist die Identifikation mit der eigenen Ethnizität oft stärker als die Identifikation mit der Nationalität. Dadurch werden die Standardmodelle ungenau, da diese von einer nationalen Identifikation ausgehen.

In diesem Papier nutze ich bilaterale Exporte und Migration und zeige, dass die Anzahl der Migranten aus dem Jahr 1990 mit den Exporten in den darauffolgenden Jahren korreliert ist. Da jedoch Exporte im Jahr 1990 mit vorhergehenden Exporten korreliert sind und Migration im Jahr 1990 ebenfalls durch vorhergehende Exporten beeinflusst sein könnten, nutze ich die Anzahl der Migranten aus dem Jahr 1960 als Instrument um einen kausalen Effekt zu schätzen. Diese Vorgehensweise, welche Standard in der Literatur ist, basiert jedoch auf der Annahme, dass sich Migranten mit ihren jeweiligen Herkunftsländern identifizieren. Hinzukommt, dass selbst die Anzahl der Migranten aus dem Jahr 1960 von Einkommensdifferenzen zwischen den einzelnen Ländern bestimmt sein kann, welche gleichzeitig die Exporte beeinflussen. Reiche Länder ziehen relativ mehr Migranten an und exportieren mehr, was die Resultate negativ beeinflusst.

Um diesen negativen Einfluss zu umgehen, greife ich auf eine große Literatur bezüglich afrikanischer Ethnien zurück. Insbesondere nutze ich eine Karte, welche die geographische Verteilung von Ethnien in Afrika vor der Kolonialisierung wiederspiegelt, um ein Maß für die ethnischen Verbindungen zwischen den Ländern zu erhalten. Da die Grenzen zwischen den afrikanischen Ländern von Europäischen Mächten im späten 19.

Jahrhundert gezogen wurden, sind sie nicht durch lokalen Begebenheiten beeinflusst. Als die afrikanischen Staaten ihre Unabhängigkeit erhielten, wurden so bestimmte ethnische Gruppen in zwei Länder geteilt. Diese Gegebenheit stellt ein sogenanntes 'natürliches Experiment' dar, da diese Leute nicht aufgrund von ökonomischen Anreizen 'migriert' sind, sondern durch die Grenzziehung in den Ländern 'platziert' wurden.

Diese Besonderheit wird im Folgenden als Instrument für die Migration im Jahr 1990 genutzt. Der daraus resultierende Effekt von Migration aus dem Jahr 1990 auf Exporte ist etwa doppelt so groß ist wie der ursprünglich berechnete Effekt, welcher die ethische Vielfalt nicht berücksichtigt. Um nachzuweisen, dass eine größere ethnische Verbindung zwischen Ländern keine anderen Effekte hat, teste ich eine Reihe von Hypothesen, die die Validität des Instruments unterstreichen. Die Tests zeigen, dass das Instrument nicht mit einer größeren Sprachgleichheit, weniger Konflikten, oder länderübergreifenden Präferenzen korreliert ist.

Des Weiteren zeige ich in der Analyse, dass Ethnien, die an Regierungskoalitionen beteiligt sind, diese ethnischen Netzwerke nicht benötigen. Die Effekte sind am stärksten für Ethnien, welche innerhalb ihres Landes gezielt diskriminiert werden und sich andere Möglichkeiten, wie Migration, zur ökonomischen Entfaltung suchen müssen. Um diese Interpretation zu verifizieren nutze ich Daten über die politische Struktur der einzelnen Ethnien vor der Kolonialisierung. Ethnien, welche stärker zentralisiert waren und Städte verwalteten, sind häufiger in der Regierungen vertreten und sind weniger auf ihre Verbindungen in die Nachbarländer angewiesen.

Die Ergebnisse aus diesem Kapitel lassen zwei Schlüsse zu. Erstens zeigen die positiven Effekte von geteilten Ethnien auf den Anteil der Exporte und der wirtschaftlichen Entwicklung eines Landes, dass sich die bislang bestehende Annahme, laut welcher ethnische Vielfalt das Wirtschaftswachstums Afrikas behindert, nur teilweise bestätigen l´asst. Zweitens, da diese ethnischen Verbindungen wiederum andere ethnische Gruppen ausschließen, sind sie weniger effizient als gute staatliche Institutionen, welche den Handel fördern. Würde man diese Institutionen fördern, könnte dies Afrika einen weiteren Entwicklungsschub geben.

Sammanfattning

Denna avhandling består av tre fristående uppsatser om ämnen som är centrala för utvecklingsekonomi och politisk ekonomi. Varje uppsats behandlar ett separat problem i dagens utvecklingsländer och försöker finna svar baserat på tidigare erfarenheter. Hur ska vi allokera äganderätter för att bevara resurser för framtida generationer utan att skada dess användares levebröd? Hur reagerar människor när de erfar våldsamma grymheter och ger uttryck för sitt missnöje? Hur påverkar migration och etnisk heterogenitet i afrikanska länder ekonomisk aktivitet? I den här avhandlingen har jag för avsikt att ge insikter i ovanstående frågor genom att använda primärdata, där en hel del digitaliserats och används för första gången, och noggrann ekonometri för att ge information för framtida policydiskussioner.

Property Rights, Resources, and Wealth [Äganderätter, resurser och förmögenhet] Det första kapitlet beaktar den pågående debatten om hur man ska fördelar äganderätter till gemensamma resurser. I det moderna Afrika har en ökande privatisering medfört en minskning av de jordbruks- och betesmarker som använts av befolkningen på landsbygden under århundraden. Dessa gemensamma resurser reglerades av rättigheter enligt gammal hävd där användningens varaktigheten reglerade resursens säkerhet och lönsamhet. (Goldstein and Udry, 2008).

Då en befolkningsökning utgjorde ett tryck på de kvarvarande gemensamma resurserna, så minskade den överexploatering som denna ledde till användarnas inkomster. Denna situation myntades av Hardin (1968) som 'Tragedy of the Commons' och två lösningar diskuteras sedan dess. Den

första lösningen som baseras på privatisering kan ofta härledas tillbaka till en inflytelserik artikel av Coase (1960). I en situation utan transaktionskostnader kommer den totala privatiseringen av resursen att ge ett effektivt resultat, då de minst lönsamma bönderna kommer att sälja den egendom de nyligen införskaffat till högproduktiva bönder. Det är viktigt att den initiala fördelningen av ägande är irrelevant då man kan handla fritt med resursen. Om den initiala fördelningen emellertid är koncentrerad till en enda individ, individerna har kapitalbegränsningar, eller vi inför transaktionskostnader, är effektiviteten inte längre garanterad. Därmed den andra lösningen som är baserad på allmännyttans princip (public goods principle) från Samuelson (1954) där en institution äger alla resurser. Genom att äga alla resurser och hyra ut rätten att använda dessa resurser till det optimala priset för användaren kan det effektiva utfallet uppnås.

Det är viktigt att båda lösningarna, privatisering eller allmänt ägande, ger samma utfall under antagandet om perfekta marknader och inga transaktionskostnader. Resursen bevaras och dess ägare erhåller en stadig inkomstström. Då äganderätter formas endogent, har ekonomer emellertid begränsad empirisk bevisning i nyckelfrågor: Jämfört med ett system utan några som helst former av äganderätter, borde en beslutsfattare privatisera resursen eller hyra ut tillgången till denna resurs och förvalta den genom att använda en institution? Vilken lösning är bäst för resurshantering? Vilken lösning är bäst för bönderna?

I detta kapitel ger jag bevis i dessa frågor genom att använda en unik storskalig markreform i USA 1934. Här placerades mer än 140 miljoner tunnland under allmän förvaltning och bönder som tidigare använt marken tilldelades rättigheter till dessa betesmarker för betande boskap. Samtidigt privatiserades land fram till 1934 baserat på en strikt regel vad gäller hemman, som gör det möjligt för mig att jämföra effekterna av privatisering och offentlig förvaltning på produktivitet. Då mängden växtlighet är en god prediktor för den mängd boskap som en ängsmark kan föda, använder jag moderna satellitdata och jämför växtligheten på båda sidor om staketet.

Resultaten tyder på att om betesrätterna upprätthålls så är de initiala prediktionerna från Coase och Samuelson korrekta. Privatisering och offentlig förvaltning ökar växtligheten, mätt 50 år senare, med samma mängd, vilket tyder på att de är lika effektiva. Om vi emellertid jämför bönder med tillgång till dessa betesmarker med bönder utan, så observerar vi en högre inkomst och förmögenhet för bönder med tillträde mer än 50 år efter implementeringen av denna reform. Det är viktigt att dessa förmögenhetsresultat endast uppstår i områden med lägre transaktionskostnader, vilket tyder på att bönder handlade med dessa rättigheter och omallokerade till andra yrken.

När cirkeln sluts gjorde beslutsfattare i dagens utvecklingsländer rätt när de privatiserade resurser i områden med lågt utnyttjande. Resultaten i detta kapitel tyder emellertid på att om bönder är kreditbegränsade eller det finns en informationsasymmetri mellan banker och bönder, kan offentlig förvaltning potentiellt vara mer effektiv. För en bonde som tidigare använde en gemensam resurs utgör ett långsiktigt kontrakt för samma kvantitet och tid som verkställs och garanteras av staten en äganderätt. Detta ökar denna bondes bokförda värde och mildrar antingen kreditbegränsningarna eller gör det möjligt för denna bonde att sälja sin gård till ett högre pris än tidigare. Viktigt är att, än en gång åberopande Coaseteoremet, allokeringen är optimal efter att alla transaktioner har utförts. I den här situationen, om det effektiva utfallet är att få personer äger stora stråk av land, så kommer detta att hända i vilket fall. Den enda skillnaden är riktningen på transfereringarna som nu involverar bönder och inte regeringen.

State Repression, Exit, and Voice [Statsförtryck, utträde, och röst] Det andra kapitlet diskuterar en tyvärr nog allestädes närvarande fråga som uppstått efter konflikten. Hur reagerar folk på att iakttaga grymheter? Utöver de kvantitativa delarna så finns här en mer grundläggande fråga. Hur anser vi som forskare att folk bör reagera på att iakttaga grymheter? å ena sidan, om man iakttar en utländsk styrka som begår våldsamma handlingar, så kan man eventuellt anta att det optimala agerandet är att öka “inomgrupps”-kohesionen för att ha en bättre chans att förhindra ytterligare våldshandlingar. å andra sidan, om dessa våldshandlingar utförs av medlemmar i din egen grupp, kan det vara optimalt att inte lita på starka ledare som påkallar sådana handlingar

Vår utgångspunkt är ett av de värsta folkmorden i modern historia där ca 30% av befolkningen massakrerades av sitt eget folk baserat på huruvida de ansågs vara till nytta för staten. När Röda khmererna tog över makten i Kambodja 1975, så delades människorna in i grupper av människor som var till nytta för att skapa ett jordbruksemperium och de som inte var det. Människor som bodde i tätorter skickades till arbetsläger på landsbygden för att öka produktiviteten och slutligen dö av svält eller dödas.

Eftersom det kambodjanska samhället var byggt på starka beskyddare skyddsling förhållanden, där förtroendet för de starka ledarna var starkt, så fanns det inget motstånd mot dessa grymheter i samhällena på landsbygden. Vår utgångspunkt är då, hur reagerade människor som iakttog dessa grymheter? Framför allt så utvecklar vi ett enkelt ramverk där överlevaren har möjlighet att uttrycka sitt missnöje, lämna civilsamhället eller vara lojal mot sin ledare.

Våra resultat tyder på att iaktta våld leder till en dramatisk attitydförändring. I stället för att lita på sina lokala ledares övertygelser och information, så informerar överlevare sig och skapar sina egna individuella preferenser. Detta leder till att starka ledare förkastas och till mer demokratiska värden, framför allt vid valurnorna. Vidare så frigör sig människor från sitt lokala liv, då återverkningarna av att avslöjas som nonkonformist med det traditionella synsättet, har inpräntats genom deras erfarenheter under folkmordet. En minskning av avskogningen är en oavsiktlig, men välkommen, bieffekt av minskat lokalt förtroende och lokalt engagemang. Trots att avskogning är illegal i Kambodja så är den fortfarande vittspridd på landsbygden då korruptionen sprider sig som en löpeld. Då övertygelsen att grannar och lokala ledare inte kommer att se till att en muta leder till åtal utgör en nyckelkomponent för korruption, minskade folkmordet avskogningen genom att minska beroendet av beskyddare-skyddsling samhället. Speciellt i omgivningar med lågt förtroende-litet civildeltagande, är korruptionsjämvikter svåra att upprätthålla.

Migration and Ethnic Heterogeneity [Migration och etnisk heterogeneitet] I denna avhandlings tredje kapitel så studerar jag den gamla, men relevanta, frågan om huruvida migration är bra för samhället och försöker skaffa information on en lucka i litteraturen. Medan vi har gott om bevis för att migration från rika länder till rika länder ökar exporten och kulturella erfarenheter, så finns betydligt mindre kunskap om migration från fattiga till rika eller från fattiga till fattiga länder, trots att dessa utgör majoriteten av migrationsflödena.

Genom att fokusera på exporten mellan afrikanska länder så siktar jag på kärnan i migrationen mellan fattiga länder men står inför två avsevärda hinder för kvalificerad forskning. För det första mäts både migrations- och exportflöden med avsevärt störningar. För det andra är identifikationen med etnicitet mycket starkare än identifikationen med nationalitet och standardtillvägagångssättet kommer sannolikt att skapa snedvridna beräkningar.

I det här papperet använder jag bilateral export och migrationsdata för att visa att mängden av migranter 1990 är korrelerad med exporten under perioden 1989-2014. För att övervinna eventuella endogenitetsfrågor så använder jag standardtillvägagångssättet i litteraturen och instrumenterar mängden migranter med deras motsvarighet 1960. Emellertid, eftersom detta standardtillvägagångssätt baseras på antagandet om nationell identitet och migrationen 1960 kan drivas av initiala inkomstskillnader mellan länder som formar exporten, hävdar jag att detta tillvägagångssätt inte ger beräkningar av den sanna effekten som inte är snedvridna.

För att övervinna denna möjliga snedvridning, gör Afrika det möjligt för mig att använda en rik historia av forskning om etnisk identifikation. Framför allt använder jag en karta som visar den förkoloniala fördelningen av etniska grupper och korsar dessa med nuvarande landsgräns för att erhålla ett mått på etnisk bindning mellan två länder som gränsar till varandra i Afrika. Det är viktigt att eftersom dessa gränser drogs upp av imperialmakter på sent 1800-tal så speglar de inte gruppernas preferenser i Afrika. Vidare, då fördelningen av dessa etniciteter fastställdes innan självständigheten, så är den inte skapad av initiala inkomstskillnader och tillhandahåller sålunda ett rimligt instrument för migration.

Genom att inkludera etnisk heterogeneitet genom att använda den förkoloniala fördelningen av etniciteter, visar jag att de beräknade effekterna av migration på export är ungefär två gånger så stora. Att dra slutsatsen att denna beräkning inte är snedvriden av en utesluten variabel, så avvisar jag hypoteser som är baserade på språkliga likheter, liknande preferenser som sammanbinder länder och konflikter.

Genom att använda data om regeringskoalitioner finner jag stöd för hypotesen att etniciteter ersätter etniska kontakter med gynnsamma institutioner. Effekterna är större för etniciteter som utgör en minoritet eller är utsatta för aktiv diskriminering och inte utgör en del av regeringskoalitionerna. Jag bekräftar detta resultat genom att använda historisk information om den politiska centraliseringen av etniciteter. Grupper som historiskt sett var mer centraliserade var mer sannolika att erövra en större del av regeringsapparaten och sålunda vara mindre beroende av etniska band.

Resultaten i detta kapitel tyder på två slutsatser. För det första, de positiva effekterna av splittrade etniciteter på export tyder på att den bild av afrikansk underutveckling som baseras på etnisk fraktionalisering är inkomplett då etniciteter använder existerande band för att övervinna handelshinder och hinder för ekonomisk utveckling. För det andra, då användningen av deras etniska nätverk länder mellan sannolikt är sämre än bra institutioner, skulle en inkludering av dessa etniciteter i regeringskoalitioner och förbättrade institutionerna ytterligare kunna öka den afrikanska ekonomiska utvecklingen.

Bibliography

Acemoglu, D., T. A. Hassan, and J. A. Robinson (2011). Social structure and development: A legacy of the Holocaust in Russia. *The Quarterly Journal of Economics 126*(2), 895–946.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Ades, A. and R. Di Tella (1999). Rents, competition, and corruption. *American Economic Review 89*(4), 982–993.

Aker, J. C., M. W. Klein, S. A. O'Connell, and M. Yang (2014). Borders, ethnicity and trade. *Journal of Development Economics 107*(1), 1–16.

Alesina, A., W. Easterly, and J. Matuszeski (2011). Artificial states. *Journal of the European Economic Association 25*, 246–277.

Alesina, A. and N. Fuchs-Schündeln (2007). Goodbye lenin (or not?): The effect of communism on people. *American Economic Review 97*(4), 1507–1528.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy*.

Alston, L. J., G. D. Libecap, and B. Mueller (2000). Land reform policies, the sources of violent conflict, and implications for deforestation in the Brazilian Amazon. *Journal of Environmental Economics and Management 39.*

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review 69*(1), 106–116.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Aragón, F. M. (2015). Do better property rights improve local income?: Evidence from First Nations' treaties. *Journal of Development Economics 116*(1), 43–56.

Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (2017). Data retrieved from `http://www.chikyu.ac.jp/precip/english/`.

Athey, S. and G. Imbens (2017). Chapter 3 - the econometrics of randomized experimentsa. In A. V. Banerjee and E. Duflo (Eds.), *Handbook of Field Experiments*, Volume 1 of *Handbook of Economic Field Experiments*, pp. 73 – 140. North-Holland.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science 293*(5536), 1818–1820.

Baaz, M. and M. Lilja (2014). Understanding hybrid democracy in Cambodia: The nexus between liberal democracy, the state, civil society, and a "politics of presence". *Asian Politics & Policy 6*(1), 5–24.

Bandyopadhyay, S., C. c. Coughlin, and H. J. Wall (2008). Ethnic networks and US exports. *Review of International Economics 16*, 199–213.

Bates, R. H. (2008). *When Things Fell Apart: State Failure in Late-Century Africa.* Cambridge University Press.

Battisti, M., G. Peri, and A. Romiti (2016). Dynamic effects of co-ethnic networks on immigrants' economic success. *NBER Working Paper* (22389).

Bauer, M., C. Blattman, J. Chytilová, J. Henrich, E. Miguel, and T. Mitts (2016). Can war foster cooperation? *Journal of Economic Perspectives 30*(3), 249–274.

Bellows, J. and E. Miguel (2009). War and local collective action in Sierra Leone. *Journal of Public Economics 42*(11-12), 1144–1157.

Bennett, C. (Ed.) (2015). *To Live Amongst the Dead: An Ethnographic Exploration of Mass Graves in Cambodia*. University of Kent.

Besley, T. (1995). Property rights and investment incentives: Theory and evidence from Ghana. *Journal of Political Economy 103*(5), 903–937.

Besley, T., K. B. Burchardi, and M. Ghatak (2012). Incentives and the de Soto effect. *The Quarterly Journal of Economics 127*(1), 237–282.

Besley, T. and T. Persson (2011). The logic of political violence. *The Quarterly Journal of Economics 129*(3), 1411–1445.

Besley, T., T. Persson, and D. M. Sturm (2010). Political competition, policy and growth: Theory and evidence from the US. *Review of Economic Studies 77*(10), 1329–1352.

Bidner, C., P. Francois, and F. Trebbi (2015). A theory of minimalist democracy. *Working Paper*.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics 114*(2), 577–599.

Blattman, C. (2009). From violence to voting: War and political participation in Uganda. *American Political Science Review 103*(2), 231–247.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Bleakley, H. and J. Ferrie (2016). Shocking behaviour: Random wealth in antebellum Georgia and human capital across generations. *The Quarterly Journal of Economics 2016*(4), 1455–1495.

Bromley, D. W. and M. M. Cernea (1989). The management of common property natural resources. *World Bank Discussion Papers 57.*

Brownlee, J. (Ed.) (2007). *Authoritarianism in an Age of Democratization.* New York: Cambridge University Press.

Burchardi, K., T. Chaney, and T. Hassan (2016). Migrants, ancestors, and investments. *Working Paper*.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Pardo i Miguel (2015). The value of democracy: Evidence from road building in Kenya. *American Economic Review 105*(6), 1817–1851.

Burke, M. B., E. Miguel, S. Satyanath, J. A. Dykema, and D. B. Lobell (2009). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences 106*(49), 20670–20674.

Calef, W. (Ed.) (1960). *Private Grazing and Public Lands.* Chicago: The University of Chicago Press.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association 110*(512), 1753–1769.

Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017). Curriculum and ideology. *Journal of Political Economy 125*(2), 338–392.

Cantoni, E. (2016). A precinct too far: Turnout and voting costs. *Job Market Paper*.

Chandler, D. (Ed.) (1988). *The Tragedy of Cambodian History. Politics, War, and Revolution since 1945.* New Haven, CT: Yale University Press.

Chandler, D. (Ed.) (2008). *A history of Cambodia.* Boulder, CO: Westview Press.

Chandler, D., B. Kiernan, and C. Boua (Eds.) (1988). *Pol Pot Plans the Future: Confidential Leadership Documents from Democratic Kampuchea, 1976-1977.* New Haven, CT: Yale Southeast Asia Studies Monograph Series.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–1721.

Chaney, T. (2016). Liquidity constrained exporters. *Journal of Economic Dynamics and Control 72*, 141–154.

Chen, C., R. Santaeulalia-Llopis, and D. Resuccia (2017). The effects of land markets on resource allocation and agricultural productivity. *NBER Working Paper* (24034).

Chernina, E., P. C. Dower, and A. Markevich (2015). Property rights, land liquidity, and internal migration. *Journal of Development Economics 116*(1), 43–56.

Christensen, D., A. Hartman, and C. Samii (2017). Legible institutions and land demand: The effect of property rights systems on investments in Liberia. *Working Paper*.

Ciccone, A. (2013). Estimating the effect of transitory economic shocks on civil conflict. *Review of Economics and Institutions 4*(2), 1–14.

Clark, W. R., M. Golder, and S. N. Golder (2017). An exit, voice, and loyalty model of politics. *British Journal of Political Science 47*(4), 719–748.

Clochard, G. and G. Hollard (2018). Africa's growth tragedy, 20 years on. *Working Paper*.

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics 3*, 1–44.

Coggins, G. C. and M. Lindeberg-Johnson (1982). The law of public rangeland management II: The commons and the Taylor Act. *Environmental Law 13*(1), 1–102.

Cotula, L., S. Vermeulen, R. Leonard, and J. Keeley (Eds.) (2009). *Land grab or development opportunity? Agricultural investments and international land deals in Africa.* London/Rome: IIED/FAO/IFAD.

Davenport, C. and M. Inman (2012). The state of state repression research since the 1990s. *Terrorism and Political Violence 24*(4), 619–634.

de la Sierra, R. S. and J.-P. Z. Mutakumura (2014). On the impact of states: Contract enforceability and ethnic institutions in East Congo. *Working Paper*.

De Soto, H. (Ed.) (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West but Fails Everywhere Else.* New York: Basik Books, Bantam Press/Random House.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica 78*(6), 1863–1903.

Dell, M. (2012). Path dependence in development: Evidence from the Mexican revolution. *Working Paper*.

Dercon, S. and R. Gutierrez-Romero (2012). Triggers and characteristics of the 2007 Kenyan electoral violence. *World Development 40*(4), 731–744.

Desbarats, J. (Ed.) (1995). *Prolific survivors: Population change in Cambodia 1975-1993.* Arizona State University.

Dunlevy, J. A. (2006). The influence of corruption and language on the protrade effect of immigrants: Evidence from the American states. *Review of Economics and Statistics 88*(1), 182–186.

Dunlevy, J. A. and W. K. Hutchinson (1999). The impact of immigration on American import trade in the late nineteenth and ealry twentieth centuries. *Journal of Economic History 59*, 1043–1062.

Eifert, B., E. Miguel, and D. N. Posner (2010). Political competition and ethnic identification in Africa. *American Journal of Political Science 54*(2), 494–510.

Everett, S. and T. Meisburger (2014). Democracy in Cambodia. Technical report, Asia Foundation. `https://asiafoundation.org/resources/pdfs/DemocracyinCambodia.pdf`.

Fafchamps, M. (2003). Ethnicity and networks in African trade. *Contributions to Economic Analysis and Policy 2*(1).

Fafchamps, M. and F. Gubert (2007). Risk sharing and network formation. *American Economic Review 97*(2), 75–79.

Federal Deposit Insurance Corporation (2001). Federal deposit insurance corporation data on banks in the United States, 1920-1936. http://doi.org/109.3886/ICPSR0007.v1.

Fehr, E. and A. Leibbrandt (2011). A field study on cooperativeness and impatience in the tragedy of the commons. *Journal of Public Economics 95*(9-10), 1144–1155.

Felbermayr, G. J., B. Jung, and F. Toubal (2010). Ethnic networks, information, and international trade: Revisiting the evidence. *Annals of Economics and Statistics 98*(4), 41–70.

Felbermayr, G. J. and F. Toubal (2012). Revisiting the trade-migration nexus: Evidence from new OECD data. *World Development 40*(5), 928–937.

Field, E. (2005). Property rights and investments in urban slums. *Journal of the European Economic Association Papers and Proceedings 3*(2-3), 279–290.

Field, E. (2007). Entitled to work: Urban tenure security and labor supply in Peru. *The Quarterly Journal of Economics 122*(4), 1561–1602.

Foley, D. K. (1970). Lindahl's solution and the core of an economy with public goods. *Econometrica 38*(1), 66–72.

Foss, P. O. (Ed.) (1960). *Politics and Grass.* Seattle: University of Washington Press.

Galiani, S. and E. Schargrodsky (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics 94*(9-10), 700–729.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2014). Competition and ideological diversity: Historical evidence from US newspapers. *American Economic Review 104*(10), 3073–3114.

Giry, S. (2015). Autopsy of a Cambodian election: How Hun Sen rules. *Foreign Affairs 94*(5), 144–159.

Giuliano, P. and A. Spilimbergo (2014). Growing up in a recession. *Review of Economic Studies 81*(2), 787–817.

Glaser, C., C. Romaniello, and K. Moskowitz (2015). Cost and consequences. the real price of livestock grazing on America's public lands. *Report for the Center for Biological Diversity*.

Goldstein, M. and C. Udry (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy 116*(6), 981–1022.

Gould, D. M. (1994). Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows. *Review of Economics and Statistics 76*, 302–316.

Grossman, G. M., E. Helpman, and A. Szeidl (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics 70*(1), 216–238.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). High-resolution global maps of 21st-century forest cover change. *Science 342*(6160), 850–853. Data retrieved from: `http://www.glad.umd.edu/dataset`.

Hansen, Z. K. and G. D. Libecap (2004). Small farms, externalities, and the dust bowl of the 1930s. *Journal of Political Economy 112*(3), 665–694.

Harari, M. and E. L. Ferrara (2018). Conflict, climate and cells: A disaggregated analysis. *The Review of Economics and Statistics, forthcoming*.

Hardin, G. (1968). The tragedy of the commons. *Science 162*(3859), 1243–1248.

Heder, S. and B. D. Tittemore (2001). Even candidates for prosecution: Accountability for the crimes of the Khmer Rouge. *Washington, War Crimes Research Office, Washington College of Law and Coalition for International Justice*.

Henderson, J., A. Storeygard, and D. N. Weil (2012). Measuring economic growth from outer space. *American Economic Review 102*(2), 994–1028.

Herander, M. G. and L. A. Saavedra (2005). Exports and the structure of immigrant-based networks: The role of geographic proximity. *Review of Economics and Statistics 87*, 323–335.

Heuveline, P. (1998). 'Between one and three million': Towards the demographic reconstruction of a decade of Cambodian history (1970-1979). *Population studies 52*(1), 49–64.

Hiebert, M. S. (2017). Genocide, revolution, and starvation under the Khmer Rouge. *Genocide Studies International 11*(1), 68–86.

Hirschman, A. O. (Ed.) (1970). *Exit, Voice, and Loyalty. Responses to Decline in Firms, Organizations, and States.* Cambridge, MA: Harvard University Press.

Hirschman, A. O. (1978). Exit, voice, and the state. *World Politics 31*(1), 90–107.

Hjort, J. (2014). Ethnic divisions and production in firms. *Quarterly Journal of Economics 129*(4), 1899–1946.

Holmes, T. J. (1998). The effect of state policies on the locatiuon of manufacturing: Evidence from state borders. *Journal of Political Economy 106*(4), 667–705.

Hornbeck, R. (2010). Barbed wire: Property rights and agricultural development. *The Quarterly Journal of Economics 125*(2), 767–810.

Hornbeck, R. (2012). The enduring impact of the American dust bowl: Short- and long-run adjustments to environmental catastrophe. *American Economic Review 102*(4), 1477–1507.

Hsiang, S. M., M. Burke, and E. Miguel (2013). Quantifying the influence of climate on human conflict. *Science 341*(6151).

Johnson, S., J. McMilland, and C. Woodruff (2002). Property rights and finance. *American Economic Review 92*(5), 1335–1356.

Keele, L. J. and R. Titiunik (2015). Geographic boundaries as regression discontinuities. *Political Analysis 23*(1), 127–155.

Kiernan, B. (Ed.) (2008). *The Pol Pot Regime.* Yale University Press.

Klyza, C. M. (1994). Ideas, institutions, and policy patters: Hardrock mining, forestry, and grazing policy on United States public lands, 1870–1985. *Studies in American Political Development 8*(2), 341–374.

Knight, R. S. (Ed.) (2010). *Statutory recognition of customary land rights in Africa.* Rome: Development Law Service, FAO Legal Office.

Krautheim, S. (2012). Heterogenous firms, exporter networks and the effect of distance on international trade. *Journal of International Economics 87*(1), 27–35.

Kudamatsu, M., T. Persson, and D. Strömberg (2016). Weather and infant mortality in africa. *Working Paper*.

Ledgerwood, J. and J. Vijghen (2002). *Decision making in Khmer Villages.* DeKalb, IL: Northern Illinois University, Center for Southeast Asian Studies.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Libecap, G. D. (Ed.) (1981). *Locking up the Range.* Pacific Institute for Public Policy Research.

Libecap, G. D. and D. Lueck (2011). The demarcation of land and the role of coordinating property institutions. *Journal of Political Economy 119*(3), 426–467.

Lindahl, E. (1919). Die Gerechtigkeit der Besteuerung. In R. A. Musgrave and A. T. Peacock (Eds.), *Classics in the theory of public finance*, Chapter 4. London: Macmillan.

Madestam, A. and D. Yanagizawa-Drott (2011). Shaping the nation: The effect of fourth of July on political preferences and behavior in the United States. *Working Paper*.

Magaloni, B. (Ed.) (2006). *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico.* New York: Cambridge University Press.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics 126*(1), 373–416.

McKenzie, D. and H. Rapoport (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics 84*, 1–24.

Meisburger, T. (2003). Democracy in Cambodia. Technical report, Asia Foundation. `https://asiafoundation.org/resources/pdfs/DemocracyinCambodia.pdf`.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica 81*(1), 113–152.

Michalopoulos, S. and E. Papaioannou (2014). National institutions and subnational development in Africa. *Quarterly Journal of Economics 129*(1), 151–213.

Michalopoulos, S. and E. Papaioannou (2016). The long-run effects of the scramble for Africa. *American Economic Review 106*(7), 1802–1848.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy 112*(4), 725–753.

Ministry of Education, Youth and Sport (2002). Census on Cambodian Schools. Data retrieved from `http://www.moeys.gov.kh/en/emis.html`.

Minnesota Population Center (2016). National historical geographic information system: Version 11.0 [database]. http://doi.org/10.18128/D050.V11.0. Obtained via https://www.nhgis.org/.

Montero, E. (2017). Cooperative property rights and agricultural development: Evidence from land reform in El Salvador. *Job Market Paper*.

Moulton, B. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics 32*(3), 385–397.

Muhn, J. A. (1987). The Mizpah-Pumpkin creek grazing district: Its history and influence on the enactment of a public lands grazing policy. *Master's Thesis, Montana State University, Bozeman*.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the U.S. labor market. *Quarterly Journal of Economics 118*(2), 549–599.

Murdock, G. (1959). *Africa: Its Peoples and Their Culture History*. New York: McGraw-Hill.

National Institute of Statistics (2014). Cambodia Socio-Economic Survey, 1996–2014. Data retrieved from `https://www.nis.gov.kh/index.php/en/about/general-information/14-cses/12-cambodia-socio-economic-survey-reports`.

National Oceanic and Atmospheric Administration (2018). Data retrieved from `ftp://ftp.cpc.ncep.noaa.gov/fews/S.Asia/data/`.

Nesbitt, H. J. (Ed.) (1997). *Rice production in Cambodia*. IRRI-Cambodia: International Rice Research Institute.

Norén-Nilsson, A. (Ed.) (2016a). *Cambodia's Second Kingdom. Nation, Imagination, and Democracy*. Ithaca, NY: Cornell University Press.

Norén-Nilsson, A. (2016b). Good gifts, bad gifts, and rights: Cambodian popular perceptions and the 2013 elections. *Pacific Affairs 89*(4), 795–815.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *Quarterly Journals of Economics 123*(1), 139–176.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics 94*(1), 20–36.

Nunn, N. and L. Wantchekon (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review 101*(7), 3221–3252.

Olken, B. and P. Barron (2009). The simple economics of extortion: Evidence from trucking in Aceh. *Journal of Political Economy 117*(3), 417–452.

Open Development Cambodia (2018). Data retrieved at various points in time from `https://opendevelopmentcambodia.net/`.

Ornaghi, A. (2016). Civil service reforms: Evidence from U.S. police departments. *Job Market Paper*.

Ostrom, E. (Ed.) (1990). *Governing the Commons.* Cambridge: Cambridge University Press.

Owen, T. and B. Kiernan (2006). Bombs over Cambodia. *The Walrus*.

Pande, R. (2011). Can informed voters enforce better governance? Experiments in low-income democracies. *Annual Review of Economics 3*(1), 215–237.

Partridge, J. and H. Furtan (2008). Immigration wave effects on Canada's trade flows. *Canadian Public Policy 34*(2), 193–214.

Peri, G. and F. Requena-Silvente (2010). The trade creation effect of immigrants: Evidence from the remarkable case of Spain. *Canadian Journal of Economics 43*(4), 1433–1459.

Pfeffer, E. L. (Ed.) (1951). *The Closing of the Public Domain: Disposal and Reservation Policies, 1900–1950.* Stanford: Stanford University Press.

Powell, J. W. (1878). Report on the lands of the arid region. *Washington: Government Printing Office*.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics 48*, 7–35.

Rauch, J. E. and V. Trindade (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics 84*(1), 116–130.

Restuccia, D. (2016). Resource allocation and productivity in agriculture. *Working Paper*.

Rice, S. and J. Tyner (2017). The rice cities of the Khmer Rouge: An urban political ecology of rural mass violence. *Transactions of the Institute of British Geographers 42*(4), 559–571.

Rogowski, J. C. (2016). Presidential influence in an era of congressional dominance. *American Political Science Review 110*(2), 325–341.

Rundle, S. (2004). The once and future federal grazing lands. *William and Mary Law Review 45*(1), 1803–1838.

Samuelson, P. A. (1954). The pure theory of public expenditure. *Review of Economics and Statistics 36*(4), 387–389.

Santos-Silva, J. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

Scheftic, W., X. Zeng, P. Broxton, and M. Brunke (2014). Intercomparison of seven NDVI products over the United States and Mexico. *Remote Sensing 6*(1), 1057–1084.

Scott, J. C. (Ed.) (2009). *The Art of Not Being Governed. An Anarchist History of Upland Southeast Asia.* New Haven, CT: Yale University Press.

Skaggs, R., Z. Edwards, B. Bestelmeyer, and P. Smith (2010). Vegetation maps at the passage of the Taylor Grazing Act (1934): A baseline to evaluate rangeland change after a regime shift. *Rangelands 33*(1), 13–20.

Spolaore, E. and R. Wacziarg (2015). Ancestry, language and culture. *Working Paper*.

Strangio, S. (Ed.) (2014). *Hun Sen's Cambodia.* New Haven, CT: Yale University Press.

Strömberg, D. (2004). Radio's impact on public spending. *The Quarterly Journal of Economics 119*(1), 189–221.

Svensson, J. (1998). Investment, property rights and political instability: Theory and evidence. *European Economic Review 42*(7), 1317–1341.

Svensson, J. (2003). Who must pay bribes and how much? Evidence from a cross section of firms. *Quarterly Journal of Economics 118*(1), 207–230.

The National Election Committee (2017). Data retrieved from `https://www.necelect.org.kh/english/`.

Thion, S. (Ed.) (1993). *Watching Cambodia: Ten paths to enter the Cambodian tangle.* Bangkok: White Lotus.

Twining, C. (1988). *The Economy.* Princeton, NJ: Princeton University Press.

Tyner, J. (Ed.) (2017a). *From Rice Fields to Killing Fields. Nature, Life, and Labor Under the Khmer Rouge.* Syracuse University Press.

Tyner, J. (Ed.) (2017b). *Landscape, memory, and post-violence in Cambodia.* New York: Rowman & Littlefield.

Un, K. (2015). *The Cambodian People Have Spoken: Has the Cambodian People's Party Heard?*, pp. 102–116. ISEAS-Yusof Ishak Institute.

Vickery, M. (Ed.) (1999). *Cambodia 1975-1982.* Chiang Mai: Silkworm Books.

Vinez, M. (2017). History and access to land at the frontier: Lessons from the colonial legacy in the Democratic Republic of Congo. *Job Market Paper*.

Voigtländer, N. and H.-J. Voth (2012). Persecution perpetuated: The medieval origins of anti-semitic violence in Nazi Germany. *The Quarterly Journal of Economics 127*(3), 1339–1392.

Voors, M. J., E. E. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. Van Soest (2012). Violent conflict and behaviour: A field experiment in Burundi. *American Economic Review 102*(2), 941–964.

Walder, A. G. (2014). Rebellion and repression in China, 1966-1971. *Social Science History 38*(3-4), 513–539.

Walker, J. M. and R. Gardner (1992). Probabilistic destruction of common-pool resources: Experimental evidence. *The Economic Journal 102*(414), 1149–1161.

Weidmann, N. B., J. K. Rød, and L.-E. Cedermann (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research 47*(4).

White, R. (2007). Immigrant-trade links, transplanted home bias and network effects. *Applied Economics 39*, 839–852.

Wimmer, A., L.-E. Cederman, and B. Min (2009). Ethnic politics and armed conflict. A configurational analysis of a new global dataset. *American Sociological Review 74*(2), 316–337.

Yale University (2014). The Cambodian Genocide Program. Data retrieved from Ben Kiernan and available under: `www.yale.edu/cgp`.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics 129*(4), 1947–1997.

Yimsut, R. (Ed.) (2011). *Facing the Khmer Rouge: A Cambodian Journey*. New Brunswick, NJ: Rutgers University Press.

MONOGRAPH SERIES

1. Michaely, Michael, *The Theory of Commercial Policy: Trade and Protection*, 1973

2. Söderström, Hans Tson, *Studies in the Microdynamics of Production and Productivity Change*, 1974

3. Hamilton, Carl B., *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost*, 1974

4. Nyberg, Lars and Staffan Viotti, *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975

5. Myhrman, Johan, *Monetary Policy in Open Economies*, 1975

6. Krauss, Melvyn, *International Trade and Economic Welfare*, 1975

7. Wihlborg, Clas, *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes*, 1976

8. Svensson, Lars E.O., *On Competitive Markets and Intertemporal Resources Allocation*, 1976

9. Yeats, Alexander J., *Trade Barriers Facing Developing Countries*, 1978

10. Calmfors, Lars, *Prices, Wages and Employment in the Open Economy*, 1978

11. Kornai, János, *Economics of Shortage, Vols I and II*, 1979

12. Flam, Harry, *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981

13. Persson, Torsten, *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach*, 1982

14. Erzan, Refik, *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983

15. Horn af Rantzien, Henrik, *Imperfect Competition in Models of Wage Formation and International Trade*, 1983

16. Nandakumar, Parameswar, *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies*, 1985

17. Sellin, Peter, *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990

18. Werner, Ingrid, *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990

19. Svedberg, Peter, *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991

20. Nordström, Håkan, *Studies in Trade Policy and Economic Growth*, 1992

21. Hassler, John, Lundvik, Petter, Persson, Torsten and Söderlind, Paul, *The Swedish Business Cycle: Stylized facts over 130 years*, 1992

22. Lundvik, Petter, *Business Cycles and Growth*, 1992

23. Söderlind, Paul, *Essays in Exchange Rates, Business Cycles and Growth*, 1993

24. Hassler, John A.A., *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994

25. Daltung, Sonja, *Risk, Efficiency, and Regulation of Banks*, 1994

26. Lindberg, Hans, *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994

27. Stennek, Johan, *Essays on Information-Processing and Competition*, 1994

28. Jonsson, Gunnar, *Institutions and Incentives in Monetary and Fiscal Policy*, 1995

29. Dahlquist, Magnus, *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995

30. Svensson, Jakob, *Political Economy and Macroeconomics: On Foreign Aid and Development*, 1996

31. Blix, Mårten, *Rational Expectations and Regime Shifts in Macroeconometrics*, 1997

32. Lagerlöf, Nils-Petter, *Intergenerational Transfers and Altruism*, 1997

33. Klein, Paul, *Papers on the Macroeconomics of Fiscal Policy*, 1997

34. Jonsson, Magnus, *Studies in Business Cycles*, 1997

35. Persson, Lars, *Asset Ownership in Imperfectly Competitive Markets*, 1998

36. Persson, Joakim, *Essays on Economic Growth*, 1998

37. Domeij, David, *Essays on Optimal Taxation and Indeterminacy*, 1998

38. Flodén, Martin, *Essays on Dynamic Macroeconomics*, 1999

39. Tangerås, Thomas, *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000

40. Lidbom, Per Pettersson, *Elections, Party Politics and Economic Policy*, 2000

41. Vestin, David, *Essays on Monetary Policy*, 2001

42. Olofsgård, Anders, *Essays on Interregional and International Political Economics*, 2001

43. Johansson, Åsa, *Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity*, 2002

44. Groth, Charlotta, *Topics on Monetary Policy*, 2002

45. Gancia, Gino A., *Essays on Growth, Trade and Inequality*, 2003

46. Harstad, Bård, *Organizing Cooperation: Bargaining, Voting and Control*, 2003

47. Kohlscheen, Emanuel, *Essays on Debts and Constitutions*, 2004

48. Olovsson, Conny, *Essays on Dynamic Macroeconomics*, 2004

49. Stavlöt, Ulrika, *Essays on Culture and Trade*, 2005

50. Herzing, Mathias, *Essays on Uncertainty and Escape in Trade Agreements*, 2005

51. Bonfiglioli, Alessandra, *Essays on Financial Markets and Macroeconomics*, 2005

52. Pienaar, Natalie, *Economic Applications of Product Quality Regulation in WTO Trade Agreements*, 2005

53. Song, Zheng, *Essays on Dynamic Political Economy*, 2005

54. Eisensee, Thomas, *Essays on Public Finance: Retirement Behavior and Disaster Relief*, 2005

55. Favara, Giovanni, *Credit and Finance in the Macroeconomy*, 2006

56. Björkman, Martina, *Essays on Empirical Development Economics: Education, Health and Gender*, 2006

57. Larsson, Anna, *Real Effects of Monetary Regimes*, 2007

58. Prado, Jr., Jose Mauricio, *Essays on Public Macroeconomic Policy*, 2007

59. Tonin, Mirco, *Essays on Labor Market Structures and Policies*, 2007

60. Queijo von Heideken, Virginia, *Essays on Monetary Policy and Asset Markets*, 2007

61. Finocchiaro, Daria, *Essays on Macroeconomics*, 2007

62. Waisman, Gisela, *Essays on Discrimination and Corruption*, 2008

63. Holte, Martin Bech, *Essays on Incentives and Leadership*, 2008

64. Damsgaard, Erika, *Färnstrand Essays on Technological Choice and Spillovers*, 2008

65. Fredriksson, Anders, *Bureaucracy, Informality and Taxation: Essays in Development Economics and Public Finance*, 2009

66. Folke, Olle, *Parties, Power and Patronage*, 2010

67. Yanagizawa-Drott, David, *Information, Markets and Conflict*, 2010

68. Meyersson, Erik, *Religion, Politics, and Development*, 2010

69. Klingelhofer, Jan, *Models of Electoral Competition*, 2010

70. Perrotta, Maria Carmela, *Aid, Education and Development*, 2010

71. Caldara, Dario, *Essays on Empirical Macroeconomics*, 2011

72. Mueller, Andreas, *Business Cycles, Unemployment and Job Search*, 2011

73. von Below, David, *Essays in Climate and Labour Economics*, 2011
74. Gars, Johan, *Essays on the Microeconomics of Climate Change*, 2012
75. Spiro, Daniel, *Some Aspects of Resource and Behavioral Economics*, 2012
76. Ge, Jinfeng, *Essays on Macroeconomics and Political Economy*, 2012
77. Li, Yinan, *Institutions, Political Cycles and Corruption*, 2012
78. Håkanson, Christina, *Changes in Workplaces and Careers*, 2013
79. Qin, Bei, *Essays on Empirical Development and Political Economics*, 2013
80. Jia, Ruixue, *Essays on the Political Economics of China's Development*, 2013
81. Campa, Pamela, *Media Influence on Pollution and Gender Equality*, 2013
82. Seim, David, *Essays on Public, Political and Labor Economics*, 2013
83. Shifa, Abdulaziz B., *Essays on Growth, Political Economy and Development*, 2013
84. Panetti, Ettore, *Essays on the Economics of Banks and Markets*, 2013
85. Schmitt, Alex, *Beyond Pigou: Climate Change Mitigation, Policy Making and Decisions*, 2014
86. Rogall, Thorsten, *The Economics of Genocide and War*, 2015
87. Baltrunaite, Audinga, *Political Economics of Special Interests and Gender*, 2016

88. Harbo Hansen, Niels-Jakob, *Jobs, Unemployment, and Macroeconomic Transmission*, 2016
89. Stryjan, Miri, *Essays on Development Policy and the Political Economy of Conflict*, 2016
90. Karadja, Mounir, *On the Economics and Politics of Mobility*, 2016
91. Kitamura, Shuhei, *Land, Power and Technology, Essays on Political Economy and Historical Development*, 2016
92. Malmberg, Hannes, *Human Capital in Development Accounting and Other Essays in Economics*, 2017
93. Öberg, Erik, *On Money and Consumption*, 2017
94. Lane, Nathaniel, *States of Development: Essays on the Political Economy of Development in Asia*, 2017
95. Prawitz, Erik, *On the Move: Essays on the Economic and Political Development of Sweden*, 2017