LUND UNIVERSITY
School of Economics and Management

Master's Programme in Innovation and Global Sustainable Development

Agricultural Investments and Ecological Disruptions

The Effects of Transnational Agricultural Large-Scale Land Acquisitions on Local Biodiversity in Latin America and The Caribbean

by

Attila Mihálka
attila.mihalka.6882@student.lu.se

Abstract There is a notable research gap in understanding the differential impacts of transnational versus domestic agricultural large-scale land acquisitions (LSLAs) on local biodiversity. This thesis addresses this gap within the context of Latin America and the Caribbean. Utilizing spatial panel datasets of the Biodiversity Habitat Index (BHI) from 2000 to 2020, LSLA area approximations, and a diverse set of control variables, the research employs a DiD methodology with various regression models. The findings reveal that BHI levels are initially lower in areas affected by transnational LSLAs. The subsequent impact on BHI levels is highly context-dependent, with transnational investments leading to negative outcomes in certain countries while resulting in positive outcomes in others. This highlights the complexity of biodiversity responses to transnational agricultural investments and underscores the need for country-specific analyses in future research.

EKHS35
Master's Thesis (15 credits ECTS)
2024 May
Supervisor: Alexandra Lopez Cermeno
Examiner: Erik Green
Word Count: 14 625

Disclaimer

An AI writing assistant has been used as a means to aid the writing process of this paper, to enhance language and reading comprehension, and improve on R and LaTeX codes.

Acknowledgements

I would like to express my deepest gratitude to several individuals who have supported me throughout the journey of completing this thesis.

First and foremost, I extend my heartfelt thanks to my supervisor, Alexandra, for her unwavering support and patience. Her guidance and encouragement have been invaluable throughout this process.

A special thanks goes to Jean for his critical insights into my data analysis. His keen observations helped me identify and correct significant errors, ultimately steering my research in the right direction.

I am also immensely grateful to Adrian, Axel, Fanny, Fanny, Felix, and Molly. Their companionship and the joyous moments filled with music and fun provided much-needed relief from the stress of thesis writing.

To Elsa, Pontus, Roman, and Thadeu, thank you for keeping me company whenever I wrote my thesis in Alfa 1 basement. Your presence was a constant source of motivation.

Finally, I would like to thank Janthe, Maca, and Pablo for their unwavering support during the last month of my thesis preparations. Their encouragement and positivity helped me stay sane and focused.

Thank you all for your invaluable contributions and support.

Contents

1 Introduction **9**

1.1 Research Problem 10

1.2 Aim and Scope 10

1.3 Outline of the Thesis 11

2 Theoretical Framework **12**

2.1 Context 12

2.2 Literature Review 14

2.3 Data and Methodological Considerations 23

2.3.1 Importance of Data Selection 23

2.3.2 Methodological Approaches 24

3 Data **26**

3.1 Source Material 27

3.1.1 Biodiversity Habitat Index 27

3.1.2 Large-Scale Land Acquisitions 27

3.1.3 Additional Control Variables 29

3.2 Data Processing 34

3.2.1 Unit Approximation 34

3.2.2 Zonal Statistics 35

3.2.3 Additional Calculations 38

3.3 Final Dataset 38

3.3.1 Summary of the Final Dataset 38

3.3.2 Descriptive Statistics 38

4 Empirical Analysis **47**

4.1 Methodology 47

4.1.1 Integration of Theory and Data 47

4.1.2 Difference in Differences 48

4.1.3 The General Formula 49

4.1.4 OLS Models 49

4.1.5 Fixed Effects Models 50

4.1.6 Matched Fixed Effects Models 50

4.1.7 Robustness Checks 51

4.2 Results 52

4.2.1 OLS Results 53

4.2.2 FE Results 55

4.2.3 Matched FE Results 56

4.2.4 Summary of Different Models 56

4.2.5 Robustness checks 58

4.3 Discussion 59

5 Conclusion **62**

5.1 Summary of Data and Methodologies 62

5.2 Summary of Key Findings 63

5.3 Theoretical and Practical Implications 63

5.4 Future Research Directions 64

5.5 Policy Recommendations 64

5.6 Limitations 65

5.7 Conclusion 65

6 Bibliography **66**

A Figures and Tables **71**

B General Regression Formula **75**

List of Figures

2.1 Living Planet Index in Latin America and the Caribbean 12
2.2 Total cropland area in South America 13
2.3 Globalization, shown in the Trade Openness Index 14

3.1 A comparison of BHI in 2000 and 2020 28
3.2 LSLA locations 29
3.3 Comparison between the exact area and area approximation 34
3.4 Zonal statistics through the example of elevation in Uruguay 36
3.5 Box plot of BHI per year 40
3.6 Box plot of BHI per LSLA type 40
3.7 Box plot of BHI per year and LSLA type 41
3.8 Density plot of BHI, by year and LSLA type 41
3.9 Scatter plots and linear approximations between BHI and average location size 42
3.10 Scatter plots and linear approximations between BHI and other geography controls 43
3.11 Scatter plots and linear approximations between slope and ruggedness 43
3.12 Scatter plots and linear approximations between BHI and climate controls 44
3.13 Scatter plots and linear approximations between minimum and maximum temperature 44
3.14 Scatter plots and linear approximations between BHI and human activity controls 45
3.15 Scatter plot and linear approximation between NTL and GDP 46

4.1 Weight brackets of the control group 52

List of Tables

3.1 Summary statistics table of the final dataset 39

4.1 OLS results 53
4.2 OLS results with different subsets 54
4.3 FE Results 55
4.4 Matched FE Results 57

A.1 Summary Table of Different Models 71
A.2 Country Interaction Coefficients 72
A.3 Random Effects Models 73
A.4 Matched Random Effects Models 73
A.5 LM and Hausman tests 74

List of Abbreviations

BHI .... Biodiversity Habitat Index
BII .... Biodiversity Intactness Index
DiD .... Difference-in-differences
FAO .... Food and Agriculture Organization of the United Nations
FDI .... Foreign Direct Investment
FE .... Fixed Effects
GDP .... Gross Domestic Product
GIS .... Geographic Information System
IPRI .... International Property Rights Index
IUCN .... International Union for Conservation of Nature
LM .... Lagrange multiplier
LPI .... Living Planet Index
LSLA .... Large Scale Land Acquisitions
NIE .... New Institutional Economics
NTL .... Nighttime Lights
OLS .... Ordinary Least Squares
PCL .... Potentially Cultivable Land
TALSLA .... Transnational Agricultural Large Scale Land Acquisitions

1

Introduction

Biodiversity loss has emerged as one of the most pressing environmental issues of our time. The rapid decline in species and habitats across the globe poses significant challenges not only to ecosystems but also to human well-being. Latin America and the Caribbean, regions rich in biodiversity, have experienced particularly severe impacts. Large-scale land acquisitions (LSLAs), often for agricultural development, are increasingly scrutinized for their role in accelerating these changes.

Understanding the effects of LSLAs on biodiversity is crucial for devising strategies to mitigate their negative impacts. This thesis investigates the differential effects of transnational versus domestic agricultural investments on biodiversity in Latin America and the Caribbean. By examining a range of econometric models, this research seeks to uncover the nuanced relationships between land acquisitions and biodiversity outcomes.

The study employs a comprehensive methodological approach, utilizing Ordinary Least Squares (OLS) models, Fixed Effects (FE) models, and full matching. These models are designed to provide robust insights into how different types of LSLAs impact biodiversity, accounting for various control variables and potential biases.

The subsequent sections will outline the research problem, define the aim and scope of the study, and provide an overview of the thesis structure. These sections will lay the groundwork for a detailed examination of the theoretical framework, data sources, and empirical findings, ultimately contributing to a deeper understanding of the complex dynamics at play between agricultural investments and biodiversity in these critical regions.

1.1 Research Problem

Research exploring the complex relationship between biodiversity and economic development has been varied, encompassing both qualitative and quantitative dimensions. Scholars have delved into various facets of this dependency - notably, as a precursor of this research, I aimed to contribute to this discourse by shedding light on the significance of agricultural land size in determining a country's biodiversity levels across regions such as Latin America and the Caribbean, as well as Europe (Mihálka, 2023).

Central to understanding this dynamic is the role of land acquisitions, particularly on a large scale, which, naturally, significantly influence a country's agricultural land size. While research, such as that conducted by Davis et al. (2023), provides compelling evidence of the detrimental impact of transnational land acquisitions on biodiversity, an important gap remains unaddressed: whether these effects differ from those of domestic investments. Additionally, despite existing literature supporting the notion that foreign investments can negatively affect sustainability metrics (for more detail, see section 2.2), these researches' specific implications for biodiversity still remain unclear.

Addressing these gaps is crucial for advancing our understanding of the relationship between economic development and wildlife conservation. This research seeks to bridge these gaps by exploring how transnational and domestic land acquisitions impact local biodiversity in the Global South.

1.2 Aim and Scope

This paper seeks to fill this critical gap in the existing research landscape by providing a nuanced exploration of the impact of transnational agricultural LSLAs on local biodiversity, compared to their domestic counterparts. Drawing from the ideas of the likes of Dasgupta (2021) and Davies et. al. (2023), and partly building upon the foundations laid down in Mihálka (2023), this study endeavors to offer a comprehensive understanding of this complex relationship.

Thus the research question of this thesis is the following:

How do foreign investments in agriculture affect local biodiversity in Latin America and the Caribbean?

As the question states, the scope of this research is concentrated on Latin America and the Caribbean, regions of cardinal importance for biodiversity conservation. By focusing on these areas, the study aims to uncover region-specific dynamics and contribute to a deeper understanding of the interaction between land acquisitions and biodiversity conservation efforts. Furthermore, the research will use the presence of agricultural LSLAs as a proxy for agricultural investments.

Through the synthesis of existing literature and careful empirical analysis, this study aims to contribute valuable insights to both academic discourse and policy formulation for sustainable development. Additionally, the findings of this research may serve as a foundation for future exploration of the topic, providing a methodological framework that can be applied to other regions, or more specific case studies. By considering regional specifics and adapting the methodology accordingly, future research endeavors may further enhance our understanding of both the local, and global implications of large-scale land acquisitions on biodiversity.

1.3 Outline of the Thesis

The thesis begins with a discussion of the Theoretical Framework (chapter 2), which introduces the context, reviews past literature, and outlines the considerations relevant to the research's data and methodology. Following this, the Data chapter (chapter 3) presents the source material, details the data processing procedures undertaken, and provides a summary and descriptive statistical analysis of the final dataset. Subsequently, in the Empirical Analysis (chapter 4), the methodology employed in the research is explained, and the results are presented and discussed in detail. The thesis wraps up with a Conclusion (chapter 5) that encapsulates the findings and implications of the study.

2

Theoretical Framework

2.1 Context

The world is undeniably experiencing a major biodiversity crisis. This issue has garnered significant attention from not just the scientific community and governing bodies (e.g., OECD, 2018A & 2019B; Dasgupta, 2021; Convention on Biological Diversity, 2022A & 2022B, cited in Mihálka, 2023), but also larger news sources and the general public (e.g., Greenfield, Rainis, 2022; Pelley, 2023, cited in Mihálka, 2023). However, despite its critical importance, this global issue often remains overshadowed by other global challenges.

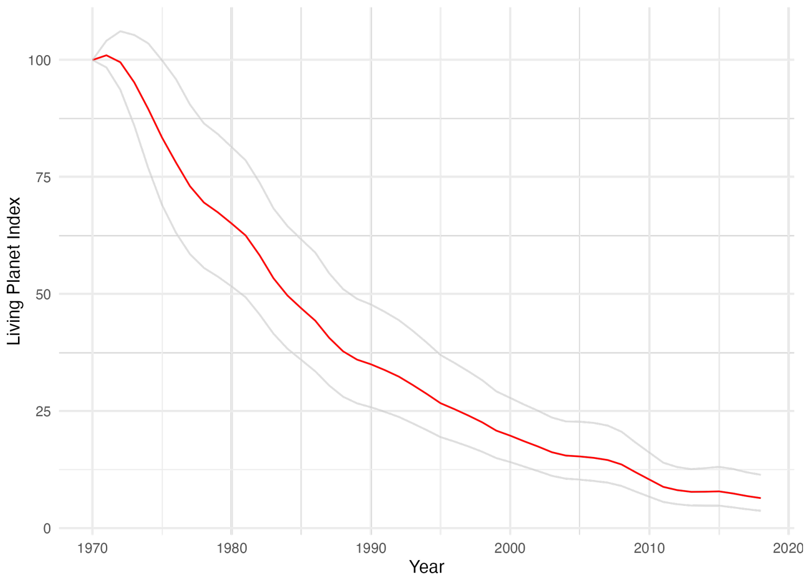


Figure 2.1: Living Planet Index in Latin America and the Caribbean
Source: Our World in Data (Ritchie, Spooner, Roser, 2024)

Latin America and the Caribbean hold particular significance in this context, as the region is home to vast areas of tropical forests and numerous biodiversity hotspots. Nonetheless, the region's wildlife has been in severe decline over the past half-century. As illustrated in figure 2.1, the region's Living Planet Index (LPI) — "[...] a measure of the state of the world's biological diversity based on population trends of vertebrate species from terrestrial, freshwater and marine habitats" (Zoological Society of London, WWF, 2023) — has plummeted from its baseline of 100% in 1970 to a deeply concerning 6.4% in 2020. This drastic decline poses severe consequences not only for the inhabitants of the region but also for the global ecosystem. Therefore, it is crucial to enhance our understanding of this issue and investigate the impact of human activities on the species we need to protect and foster.

Agriculture stands as one of the most significant impactors of biodiversity. The growing global population necessitates increased agricultural production, whether by enhancing yield per cropland through advanced technology and process optimization or by expanding the agricultural land area. For example, as shown on figure 2.2, the total cropland area of South America has more than doubled between 1970 and 2020. Without the implementation of more sustainable farming practices, this expansion could have dire consequences on the future of local biodiversity.

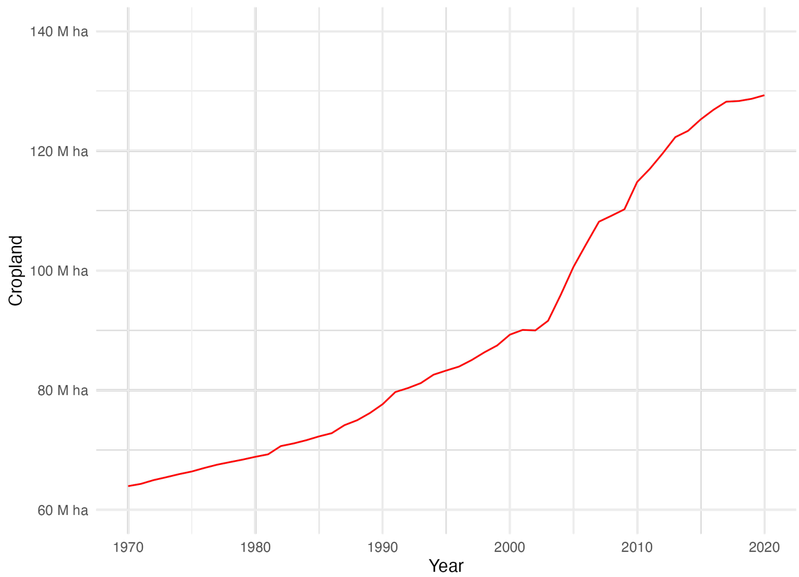


Figure 2.2: Total cropland area in South America
Source: Our World in Data (Ritchie, Roser, 2024)

Additionally, the increasingly globalized economy has profoundly impacted biodiversity. As depicted in figure 2.3, globalization has shown an increasing trend over the past two decades, despite occasional fluctuations. The rise of transnational corporations and their increasing power have exerted considerable and continuous pressure on ecosystems (as discussed by e.g. Österblom et. al., 2022).

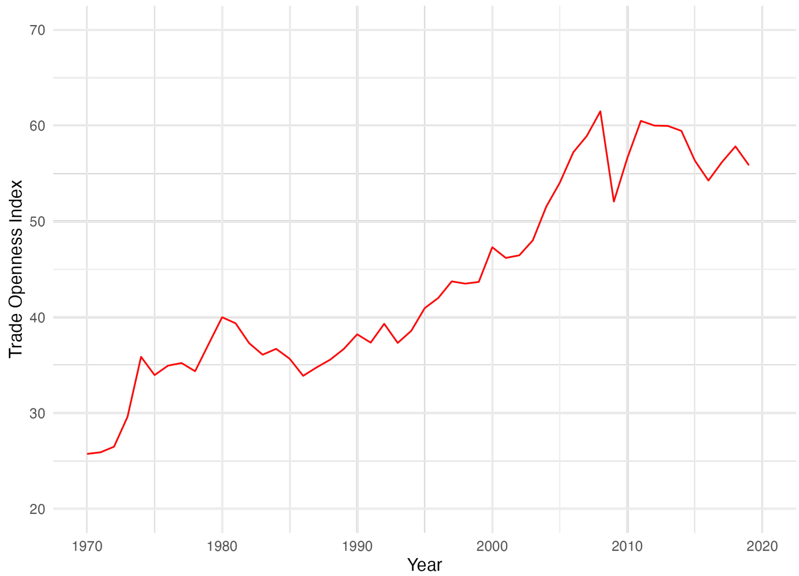


Figure 2.3: Globalization, shown in the Trade Openness Index
Source: Our World in Data (Ortiz-Ospina, Beltekian, Roser, 2024)

This research aims to interlink these three topics — biodiversity, agriculture, and the emergence of transnational entities — and uncover their effects on biodiversity in the crucial region of Latin America and the Caribbean. To deepen our understanding of these topics, the next section will review key pieces of literature that explore them in more detail.

2.2 Literature Review

The literature review section provides an in-depth examination of existing research and theoretical frameworks that inform the current study. This review aims to further contextualize the research within the broader academic discourse on transnational agricultural large-scale land acquisitions and their impacts on local biodiversity. By critically analyzing essential studies such as Dasgupta (2021) and Davis et. al. (2023), this section will highlight significant findings, identify gaps in the literature, and establish a foundation for the methodological approaches employed in this thesis.

The Dasgupta Review (2021) provides a crucial framework for understanding the extensive impact of modern agricultural practices on biodiversity. These practices, primarily aimed at maximizing farm yields, have inadvertently led to a significant reduction in biodiversity. This reduction has compromised ecosystems' ability to deliver vital services—such as regulation, maintenance, and cultural services—that are essential for the sustainable provision of provisioning services like food and building materials. This fundamental trade-off, detailed in Chapter 2, underpins the central argument of this review: productivity and resilience are increasingly dependent on biodiversity, which has been decreasing over the past 70 years due to the rising need for provisioning services.

The review also notes how the current state of global biodiversity is alarming, as highlighted by several key indicators. The global LPI shows a stark 68% decline in the abundance of an enormous amount of vertebrate populations between 1970 and 2016. The International Union for Conservation of Nature (IUCN) Red List reveals that over more than one fourth of assessed species are threatened with extinction, including substantial proportions of mammals, amphibians, conifers, reef-building corals, and birds. Additionally, the Biodiversity Intactness Index (BII) reports that only 79% of naturally present biodiversity remained in terrestrial ecosystems by 2015, with most biomes falling below the proposed safe limit of 90%. Chapter 2 also offers an extensive analysis of these indicators, painting a a complex, but grim picture of the global biodiversity crisis.

One of the significant challenges addressed in The Dasgupta Review is the measurement of biodiversity and natural capital. Chapter 4 outlines four primary challenges—diversity, reliability, adaptability, and scalability—that complicate the valuation of biodiversity, given that most natural capital is not priced in markets. The review proposes a framework for measuring biodiversity through indicators of structure, function, composition, and resilience, applicable at multiple levels. Developing robust measurement tools and frameworks is crucial for capturing the true value of biodiversity and natural capital. This underscores the choice of using a comprehensive index to represent biodiversity in this thesis, ensuring a holistic and accurate assessment of the impacts of transnational agricultural land acquisitions.

The Review underscores the importance of incorporating both temporal and spatial dimensions in biodiversity research. It explains that traditionally, research has focused on temporal trends, often neglecting spatial changes. To fully understand spatial changes, it is necessary to examine the specific conditions that support particular landscapes or ecosystems. For example, Chapter 3 discusses how biosphere tipping points can have widespread spatial implications, affecting regional water cycles and leading to mismatches between ecosystem service demands and sustainable supply capabilities. The review advocates thus for a holistic approach that integrates both temporal and spatial aspects to effectively address biodiversity loss.

The Review also addresses the significant impact of transnational investments on biodiversity, particularly in developing countries. Chapter 8 highlights how extractive activities such as mining, logging, and agricultural businesses can lead to severe ecosystem degradation and biodiversity loss. These negative impacts can disproportionately affect local communities in biodiversity-rich regions, worsening poverty and socio-economic challenges. The review argues for stringent environmental and social performance standards for transnational companies and stronger government regulations and tax policies to account for the externalities of environmental degradation. This chapter emphasizes the need for comprehensive policies that recognize the value of natural capital and mitigate the adverse effects of transnational investments.

According to Dasgupta, institutions play a pivotal role in managing biodiversity. The Review highlights the challenges posed by the open-access nature of global commons like the atmosphere and oceans. Effective institutions are essential for enforcing laws, establishing protected areas, incentivizing conservation, and supporting community-based natural resource management. Institutional change can drive behavioral shifts, reducing unsustainable consumption patterns and addressing the drivers of biodiversity loss. Chapter 7 elaborates on the significant role of institutions in biodiversity conservation, arguing that creating robust institutions is vital for sustainable natural capital management.

The Dasgupta Review concludes that protecting biodiversity through sustainable development is essential for ensuring the long-term delivery of ecosystem services. Sustainable practices not only enhance ecosystem resilience and reduce poverty but also create new economic opportunities. Chapters 3 and 5 discuss how sustainable development practices that protect biodiversity can offer significant benefits, including mitigating environmental disturbances such as climate change, natural disasters, and emerging diseases. By integrating the insights from The Dasgupta Review, this literature review highlights the critical importance of sustainable development in preserving biodiversity and promoting human well-being.

In summary, The Dasgupta Review (2021) provides a comprehensive framework for understanding the relationship between biodiversity, ecosystem services, and human well-being. It emphasizes the need for robust measurement tools, effective institutions, and sustainable development practices to address the ongoing biodiversity crisis. These insights are crucial for informing policies and practices aimed at protecting biodiversity and ensuring a sustainable future.

Bucheli's (2008) article delves into the historical dynamics between the United Fruit Company (later United Brands) and Central American and Caribbean nations during the 19th and 20th centuries. The multinational corporation wielded significant influence in the region, contributing to economic development through infrastructure projects while simultaneously facing scrutiny for its dealings with local dictators which in many cases resulted in negative effects on social and environmental factors.

Despite United Fruit's role in bolstering economic activity through initiatives such as establishing plantations and infrastructure, its reputation in the region remains tarnished due to its collaborations with authoritarian regimes. Bucheli challenges the notion of a natural alliance between multinational corporations and dictatorial regimes, asserting that cooperation was contingent upon economic benefits and stability provided by the corporations, alongside the strength of local labor movements.

The study situates itself within the nexus of business history and the political economy of foreign direct investment, drawing upon New Institutional Economics (NIE) to underscore the impact of political regimes on corporate operations. Bucheli underscores the significance of research on foreign direct investment in the primary sector, highlighting its potential influence on local politics and susceptibility to political violence.

The article sheds light on United Fruit's diminishing influence in the face of rising nationalism and shifting economic paradigms, ultimately leading to divestment and local ownership of plantations. It underscores the complex interplay between multinational corporations and political regimes, emphasizing the imperative of integrating theoretical debates in political economy into analyses of corporate operations in developing nations.

In conclusion, Bucheli's work offers valuable insights into the historical dynamics between multinational corporations and Central American nations, underscoring the multifaceted impacts of transnational investments on regional development and political stability.

The study conducted by Davis et al. (2023) investigates the impact of transnational agricultural large-scale land acquisitions (TALSLAs) on forest covers and biodiversity in the Global South. Utilizing data from 178 locations across 40 countries, the authors aim to assess the effects of TALSLAs on deforestation rates and vertebrate biodiversity in various regions.

The findings reveal significant evidence linking TALSLAs to deforestation in Asia and Africa, whereas no discernible difference is observed in Europe or Latin America. Regarding vertebrate biodiversity, the study indicates that most TALSLAs result in significant losses in relative species richness, although outcomes for relative species abundance are more varied. Notably, nearly 40% of the reviewed locations are situated, either wholly or partially, within biodiversity hotspots.

Davis et al. (2023) contextualize the research by defining large-scale land acquisitions as transactions involving a minimum of 200 hectares. They underscore the growing prevalence of such acquisitions, particularly transnational ones for agriculture, which account for a substantial portion of the total land area acquired. The authors emphasize the unprecedented scale of these acquisitions and their implications, rooted in historical colonial and imperial legacies. Additionally, they discuss the misconception surrounding land acquisitions in "marginal lands," often revealed to intersect with biodiversity hotspots.

Furthermore, the study highlights the potential ecosystem services provided by some land acquisitions, juxtaposed with the negative environmental impacts associated with agricultural conversions. The authors stress the significance of assessing the impacts of TALSLAs on forest covers and biodiversity through a quantitative lens, thus complementing existing research in the field.

Methodologically, Davis et al. (2023) obtain TALSLA deal centroids from the Land Matrix (2024a & b) and create approximations based on deal sizes, acknowledging the limitations of this approach in accurately representing actual deal areas. They also create additional buffer zones to examine potential spillover effects. Regarding biodiversity assessment, the study reveals severe losses in relative species richness across most locations, with mixed effects on relative species abundance. Notably, many TALSLAs are found to intersect with biodiversity hotspot areas, indicating negative biodiversity impacts even in buffer zones.

Despite its invaluable contributions, the study has areas of improvement, including the low observation count in South America relative to other regions, potentially impacting the generalizability of results. Moreover, the focus on forest loss rather than general overview of habitats warrants further investigation, particularly in regions with significant biodiversity importance such as Latin America.

Dogan (2022) provides an example of research investigating the impact of FDI in agriculture on a sustainability measure. Although the study does not directly address biodiversity, it offers valuable insights and methodologies that can be adapted for our purposes. Dogan's research, utilizing an unbalanced panel dataset from 56 developing countries between 2005 and 2020, applies the terminology of the Food and Agriculture Organization of the United Nations (FAO), defining agriculture as "agriculture, forestry, or fishing" (FAO, cited in Dogan, 2022, p. 55). This empirical study employs quantitative methodologies, specifically fixed effects regression models, to uncover a significant negative correlation between agricultural FDI flows and food security. These finding aligns with several papers reviewed in their literature review, suggesting that agricultural FDI can adversely affect sustainability measures. Furthermore, the study reveals that robust land governance systems can mitigate these negative impacts. Dogan's research underscores the importance of quantitative methods in investigating the effects of FDI in agriculture on sustainability measures and highlights the importance of considering the institutional context in such analyses.

Santangelo (2018) reinforces the findings of Dogan (2022) through similar quantitative methods. This research not only confirms the negative impacts of agricultural FDI on sustainability but also emphasizes the difficulties in obtaining detailed data on FDI in agriculture, as noted by Cotula et al. (2009, cited in Santangelo, 2018, p. 79). This challenge of data acquisition is crucial to consider for our research, as comprehensive and accurate data are necessary for robust analysis. The convergence of findings between Dogan and Santangelo strengthens the argument for the use of quantitative methodologies in studying the impacts of agricultural FDI and underscores the significance of data accessibility and quality in such research.

Ferrier and Guisan (2006) argue that modeling biodiversity and its changes using large multi-species datasets at the community level is often more effective for general goals than species-level modeling. This perspective significantly influences the approach to biodiversity research within this study. Given the relatively short scope and time-span of this thesis, resources are insufficient to analyze individual species-level data comprehensively. Ferrier and Guisan's work suggests that employing a more complex, larger biodiversity dataset will simplify the methodology while still yielding adequate results. This approach aligns well with the aim of producing robust findings within the constraints of the research timeframe.

Mann and Smaller (2010) discuss the recent trend of foreign investment in farmland for agriculture across various parts of the world, particularly focusing on crops used for food, feed, and energy. These investments, often in the form of purchases or long-term leases of large tracts of arable land, are primarily driven by concerns over food, water, and energy security. Besides highlighting the potential dangers to food and water security and certain social aspects, Mann and Smaller emphasize that the lack of regulations concerning pesticides, herbicides, water protection, and biodiversity in some host states poses significant risks to other water users, soil management, and the long-term sustainability of agricultural projects.

The role of institutional arrangements is crucial in regulating the terms and conditions of foreign investment in farmland. The presence of strong regulatory frameworks is essential for establishing clear rules for investors and safeguarding public interest. Although Mann and Smaller's research mainly reviews situations in African countries such as Ethiopia, Kenya, and Mozambique, their findings are considered applicable to Latin America to a considerable extent. However, regional differences should be acknowledged and accounted for in the analysis.

Sändig (2021) also discusses LSLAs and their effects in countries of the Global South. Their research adopts an opposite perspective compared to the previously discussed studies, exploring how local communities influence LSLAs through various avenues. Using a sample of 28 LSLAs, primarily focused on agriculture, Sändig identifies three main ways in which communities can manifest resistance: everyday resistance, contentious politics, and legal mobilization. It is found that in less constrained environments, communities can adopt a more organized approach in resisting these investments.

While legal mobilization is not discussed in detail in Sändig's examples, the overall findings suggest that the institutional setting once again plays a significant role in the outcomes of LSLAs. The ability of local communities to organize and resist these investments depends largely on the strength and structure of local institutions. This insight aligns with the broader understanding that robust institutional frameworks are crucial in managing the impacts of transnational agricultural investments and ensuring that local interests are protected.

The review of existing literature reveals several critical insights pertinent to understanding the impacts of LSLAs on local biodiversity in Latin America and the Caribbean.

The works of e.g. Mann and Smaller (2010), Dasgupta (2021) and Davis et al. (2023) confirm the severity of the biodiversity crisis, providing numerous examples through their extensive reviews. Both sources highlight how modern agricultural practices and FDI in agriculture tend to negatively affect sustainability measures, encompassing both environmental and social dimensions. This consistent finding across multiple studies underscores the significant ecological trade-offs associated with agricultural expansion and foreign investment.

Additionally, data availability issues are a recurring theme in prior research (e.g. Santangelo, 2018, Davis et. al., 2023). Assumptions and approximations, coupled with clear explanations of their potential pitfalls, are often necessary to achieve meaningful results. The reviewed literature often approximates FDI in agriculture through the lens of TALSLAs. This method, utilized in multiple studies, including Davis et al. (2023), offers a practical approach for acquiring adequate data on agricultural investments. Consequently, this approximation will be adopted in the present research as the appropriate measure of FDI in agriculture.

Davis et al. (2023) specifically highlight that LSLAs negatively impact biodiversity through various mechanisms. Their research, in line with the Dasgupta Review's Chapter 2 and 3 (2021) emphasizes the importance of considering temporal and spatial dimensions in understanding these effects, as well as the critical role of institutional settings. However, a notable gap in prior research is the differentiation between domestic and transnational LSLAs, which could significantly influence sustainable development outcomes.

The region of Latin America and the Caribbean is often overlooked in these studies. Although Mann and Smaller (2010) provide a comprehensive qualitative review of LSLAs, and Davis et al. (2023) employ a robust methodology yielding significant results, their focus remains predominantly global, evidenced by for example a low observation count in this specific region. This gap underscores the need for more focused research on Latin America and the Caribbean, regions of key importance for biodiversity conservation.

Given the critical importance of biodiversity protection in these regions and their under-representation in existing research, this study will measure foreign investments in agriculture using LSLAs, following the methodological framework of Davis et al. (2023).

To address the gap in prior research, and answer the research question presented in chapter 1, the following hypotheses are formulated based on the literature review:

H1: *All agricultural LSLAs have a significant and negative effect on local biodiversity in the region.*

H2: *Transnational LSLAs will have a higher long-term negative effect on local biodiversity in the region than domestic LSLAs.*

These hypotheses aim to investigate the multi-faceted impacts of agricultural investments on biodiversity, with a particular focus on the comparative effects of domestic versus transnational LSLAs. This research will contribute to a deeper understanding of how foreign agricultural investments influence local ecosystems in Latin America and the Caribbean, providing valuable insights for sustainable development policies.

2.3 Data and Methodological Considerations

2.3.1 Importance of Data Selection

Data selection is critical to addressing the research questions effectively. The data must align with the theoretical considerations discussed earlier in the literature review. The selection criteria for datasets include relevance to the research questions, reliability, and the ability to capture both spatial and temporal variations.

To investigate the impacts of LSLAs on biodiversity, diverse datasets capturing relevant variables will be employed. These datasets will encompass information on biodiversity, land use, climate conditions, and human activities. Each dataset will be selected based on its ability to provide high-resolution, accurate, and reliable data. The datasets also need to offer sufficient temporal coverage to enable analysis over time.

Biodiversity data will provide insights into the state of local ecosystems and the various species within them through a comprehensive index, as for example suggested by Dasgupta (2021). Data on land use will help in understanding the extent and nature of land acquisitions, while geography and climate data will control for environmental factors that could influence biodiversity independently of LSLAs. Additional human activity data will be considered to account for other influences on biodiversity, ensuring a comprehensive analysis.

Potential challenges related to data quality and availability will be addressed. Issues such as missing data or measurement errors will be mitigated through setting approximations or assumptions, which will be discussed in detail.

2.3.2 Methodological Approaches

The methodological approaches in this research will involve quantitative techniques to analyze the collected data. Given the need to address both spatial and temporal variations, spatial analysis and panel data analysis methods will be prominently featured.

Spatial analysis techniques will be employed to examine the relationships between LSLAs and biodiversity. These techniques will involve mapping and analyzing spatial patterns, providing visual and quantitative insights into the spatial dynamics of biodiversity changes. Geographic Information System (GIS) tools will be utilized to handle and process spatial data effectively.

Panel data analysis will be used to capture both spatial and temporal variations in the data. This method is advantageous as it controls for unobserved heterogeneity and captures dynamic changes over time. Spatial panel datasets will be crucial in allowing a more nuanced analysis of the interactions between LSLAs and biodiversity.

To isolate the specific effects of LSLAs on biodiversity, prior discussed various control variables such as geographic features, climate conditions and human activities will be included in the analysis. These controls are essential to account for other factors that could influence biodiversity, ensuring that the results specifically reflect the impacts of LSLAs.

Appropriate statistical models, including Ordinary Least Squares (OLS) and Fixed Effects (FE) models, and more advanced methodologies such as optimal full matching will be employed to analyze the data. The rationale for choosing these models will be explained to ensure that the methodology aligns with the research objectives.

Methodological constraints, such as the limitations of certain statistical techniques or potential biases, will also be discussed. Approaches to address these constraints, including robustness checks and alternative modeling strategies, will be proposed.

In summary, a robust spatial panel dataset has to be constructed, incorporating local biodiversity data, characteristics of LSLAs, geographic, climate, and other human activity factors. The chosen methodology will align with the research questions and aims, previous research, and the dataset's structure and characteristics. This approach ensures a comprehensive and careful analysis of the effects of transnational agricultural large-scale land acquisitions on local biodiversity in Latin America and the Caribbean.

3

Data

In the previous section, we underscored several important factors in choosing appropriate data, such as the significance of spatial and temporal patterns in studying biodiversity and its relationship with LSLAs. Thus, our research prioritized the utilization of spatial panel datasets. These datasets offer a dynamic foundation that allows for the exploration of spatial and temporal dynamics in the forthcoming analysis, aligning closely with the research objectives outlined in our study.

Moreover, the value of databases utilized in previous research was recognized, such as the Land Matrix employed in the study conducted by Davis et al. (2023), in investigating the impacts of LSLAs on biodiversity. While these databases provide a foundational understanding of the subject matter, this research required a comprehensive exploration of additional databases to meet the specific requirements of our study. By considering various data sources, we aimed to enrich our analysis and provide a more robust assessment of the relationship between LSLAs and biodiversity.

In this chapter, the data acquisition process is explored in detail, describing the approach to sourcing datasets across different categories. The arguments behind selecting each dataset and data processing, and how these choices align with the objectives of our research are examined. Additionally, descriptive statistics are used to further explore the final dataset. Through this overview, the research aims to establish a solid foundation for the subsequent analyses and interpretations presented in this study.

3.1 Source Material

3.1.1 Biodiversity Habitat Index

In the pursuit of identifying a reliable spatial biodiversity database, multiple datasets have been explored using the UN Biodiversity Lab (2024). This platform offers a comprehensive map function, allowing the examination of various layers of spatial data. Among the array of available datasets, the Biodiversity Habitat Index (BHI) database developed by Harwood et al. (2022b) emerged as the optimal choice for several reasons.

Foremost, the BHI database constructs a comprehensive biodiversity index that encapsulates the overall impact on biodiversity. While previous research, such as that conducted by Davis et al. (2023), utilized indicators like forest loss, relative species richness and relative species abundance, our focus on assessing the quality of habitats enables a holistic examination of the biodiversity landscape.

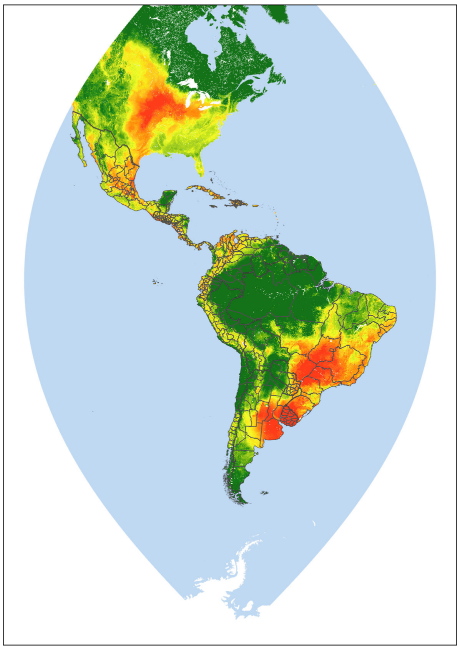
Moreover, the database employs high resolution (30 arc-seconds) spatial data, spanning the years 2000, 2005, 2010, 2015, and 2020. This extensive temporal coverage provides a large number of data points, allowing robust analysis and interpretation.

Additionally, the accompanying paper by Harwood et al. (2022a) underscores the trustworthiness and robustness of the dataset, further strengthening its suitability for the research aims. It provides detailed guidance on some necessary calculations, ensuring accuracy and reliability in data interpretation.

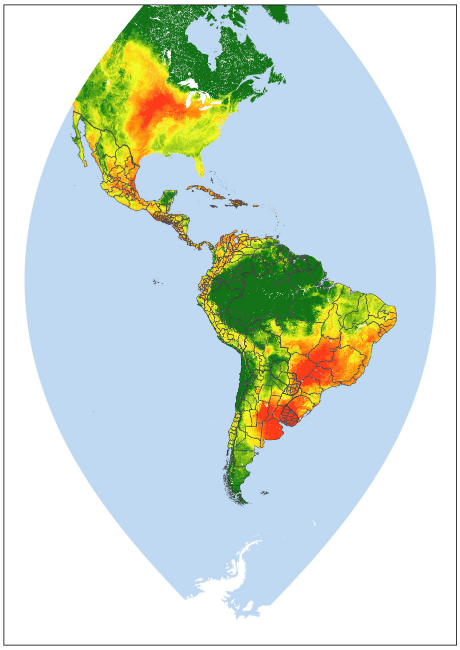
Considering these factors collectively, the BHI database can be established as a solid foundation for the research objectives. Its comprehensive nature, high resolution, and robust methodology align closely with the aims of this study, laying the groundwork for the subsequent analysis.

3.1.2 Large-Scale Land Acquisitions

As mentioned in the preceding chapter, the database by the Land Matrix (2024a and 2024b), also utilized by Davis et al. (2023), emerges as an invaluable resource for investigating the impacts of LSLAs on biodiversity. This repository is the product of an independent initiative aimed at enhancing transparency surrounding such



(a) 2000



(b) 2020

Figure 3.1: A comparison of BHI in 2000 and 2020
Red areas indicate a lower, while green areas indicate a higher BHI

investments within the Global South.

Of great importance to this study is the comprehensive nature of the database, which captures critical data points essential for analyzing LSLAs. Notably, the database carefully records the precise locations and estimated sizes of land acquisitions, encompassing both domestic and transnational transactions. Moreover, it offers a a large amount of supplementary information about the individual deals and their respective locations. This includes insights such as the purpose behind each land acquisition, the purchase price, the identities of acquiring entities, and their countries of origin, among other potentially relevant details. While these data points were not part of the subsequent analysis of the thesis, future research could delve into these details for more insights on the matter.

In essence, the Land Matrix database serves as a rich repository of information, providing researchers with a comprehensive toolkit to investigate LSLAs and their impacts on local factors, such as, in the case of this study, biodiversity.

To ensure alignment with the specific aims of this research, the original dataset has been filtered twofold. Firstly, the data was refined to encompass the geographic scope of this study, namely Latin America and the Caribbean. Additionally, the intention of the investments was restricted to agriculture or forestry, As suggested by the FAO definition cited in Dogan (2022), further refining the dataset to align with the research's scope.

Upon applying these filters, the dataset encompasses a total of 866 LSLA locations related to agriculture, comprising 370 transnational investments and 496 domestic investments. The distinction whether the individual locations are domestic or transnational serves as a crucial dummy variable, serving as the primary independent variable for addressing the research question at hand.

It is important to highlight that while the database provides coordinates for LSLA locations, it only includes exact areas for a very limited number of land acquisitions in the region. However, this limitation is addressed through the creation of area approximations, which will be explained in detail in section 3.2. This strategic approach ensures the robustness and reliability of the data for upcoming analysis and interpretation.

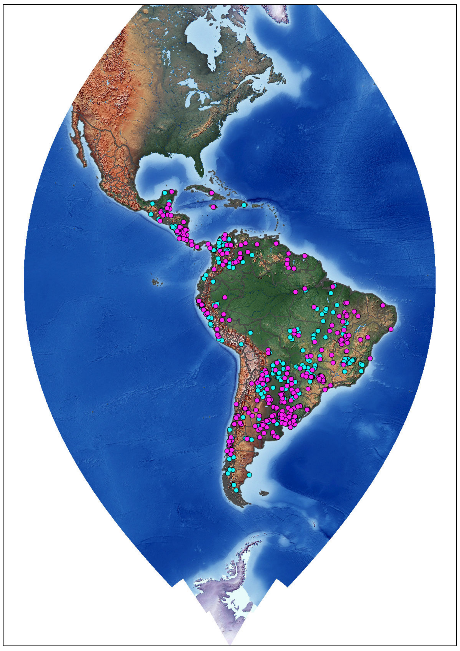


Figure 3.2: LSLA locations

Pink points represent transnational, while blue points represent domestic acquisitions

3.1.3 Additional Control Variables

In addition to the main independent variable of LSLA locations, this research acknowledges the necessity of incorporating additional control variables. It is understood that the state of biodiversity is influenced by several other factors beyond the specifics of LSLAs alone. While these variables may not serve as the primary focus of this study, they account for crucial components for constructing robust analytical models.

The additional control variables under consideration in this research can be broadly categorized into three distinct categories: geography, climate, and human activity. Each category covers a range of variables that may exert varying degrees of influence on biodiversity dynamics within the study region.

Geography

Given that both the dependent variable (BHI) and the main independent variable (LSLA locations) are spatial in nature, the inclusion of key geography controls takes on high importance. These controls serve to contextualize the spatial relationships under investigation and mitigate potential negating effects arising from geographic heterogeneity within the study region.

It is important to note that these geography controls can be considered constant for the purposes of this research, as their timescale of change is sufficiently large to be deemed negligible within the study timeframe of 2000-2020. As such, there is no need to account for temporal variability in these controls, as our analysis is confined to a specific timespan imposed by the availability of data from the BHI database.

Most evidently among these geography controls are latitude, longitude, and area data, which serve as fundamental variables of spatial location and extent. Since these variables are already included in the point dataset of the Land Matrix database, there is no need for acquiring an additional dataset.

In addition to latitude, longitude, and area data, it is vital to include additional geography controls such as elevation, slope, and terrain ruggedness in the analysis. These fundamental geographic factors have the potential to exert significant effects on both biodiversity patterns and the spatial distribution of land acquisitions. To incorporate these variables, raster datasets have been sourced from EarthEnv (2024) at a 1km resolution, using GMTED2010 as a source.

Moreover, proximity to bodies of water arises as another useful geographic parameter, given its potential impact on baseline biodiversity levels. Data on the proximity to water bodies has been obtained from Natural Earth (2024) in the form of shapefiles, compromised by a line dataset for rivers and polygon datasets for lakes. The method for extracting the actual distance from land acquisition locations to these water bodies will be elaborated upon in the upcoming section 3.2.

Additionally, an essential factor to consider is the agricultural suitability of the land, which can greatly influence both biodiversity and the spatial distribution of land acquisitions. To incorporate this variable into our analysis, data has been collected from the database of Schneider et. al. (2022), who have developed a comprehensive global inventory of potentially cultivable land (PCL) across various scenarios and time periods.

For the purposes of this research, the historical database spanning from 1980 to 2009 has been selected from Schneider et al.'s inventory. This dataset accounts for irrigation patterns, providing insights into the historical suitability of land for agricultural purposes. Specifically, raster data with a 30 arc-second resolution has been acquired to align with the resolution of other variables within our analysis.

Climate

Controlling for climate effects is also crucial in researching biodiversity, as climate variables can influence ecosystems in numerous ways. While a wide range of climate-related variables could be considered, this research focuses on two primary factors: temperature and precipitation. These variables are chosen for their significant impact on environmental conditions and their straightforward application in the upcoming empirical analysis.

Unlike geographic controls, climate variables are dynamic and their temporal changes must be accounted for in the analysis. Therefore, the datasets selected for this study are spatial panel datasets that span the observed years. The datasource that best meets these requirements is WorldClim (2022), which provides historical monthly raster datasets for minimum temperature, maximum temperature, and precipitation.

To create the necessary annual datasets, GIS tools were employed to calculate cellwise means for each raster dataset, by year of observation. By incorporating the measures for these three variables, the research ensures comprehensive control over the climatic changes that can occur within a region. This approach allows for a relatively simple but thorough examination of the impact of temperature and precipitation on biodiversity, contributing to the robustness of the analysis.

Human Activity

Once again, controlling for human activity assumes critical importance within the scope of this research. Human activities can exert significant pressures on biodiversity, necessitating their inclusion as control variables in the analysis.

Human activity variables are also dynamic in nature and cannot be considered constant over the study period. Ideally, spatial databases would be sought for these variables to capture their spatial heterogeneity and temporal dynamics comprehensively. However, due to the complexities associated with some of these variables, acquiring spatial databases may not always be feasible.

In Mihálka (2023), it has been demonstrated that population density has a significant negative effect on biodiversity. To represent population density, and by extension human activity, a spatial approximation based on nighttime lights activity can be employed. The use of nighttime light data for this purpose is well-established in scientific literature, with numerous research papers employing this approximation (e.g. Sutton, 1997, Liu et. al., 2011).

For this study, raster data has been procured from Li et al.'s (2023) Harmonized Global Night Time Lights dataset, once again offering a spatial resolution of 30 arc seconds. This dataset provides valuable insights into the spatial distribution and intensity of human activity, as indicated by nighttime lights (NTL) emissions (Li et. al., 2020). By incorporating this proxy for population density into our analytical models, we aim to capture the spatial heterogeneity of human activities and their potential impacts on biodiversity dynamics within the study region.

Additionally, once again reviewed in Mihálka (2023), Gross Domestic Product (GDP) emerges as an important variable in our analysis. The database compiled by Chen et al. (2022) stands out as the most suitable resource for our research needs, offering a gridded real GDP database. Notably, this database has been constructed using calibrated nighttime light data, which enhances its accuracy. However, it is important to acknowledge that this GDP dataset is likely to exhibit a high degree of correlation with the nighttime light dataset utilized in this study. Therefore, this correlation must be carefully considered and thoroughly tested before constructing the empirical analysis. This is further discussed in subsection 3.3.2.

Furthermore, it is essential to address the temporal limitations of the GDP dataset, which only extends until 2019. To mitigate this limitation, this paper will operate under the assumption that the GDP data for 2020 are sufficiently similar to those of 2019. Thus the 2019 GDP data is utilized in combination with 2020 data from other datasets.

Another crucial aspect considered in this research is institutional quality. While some previous quantitative studies, such as Welsch, 2004 and Tan et. al, 2022 (cited in Mihálka, 2023), have primarily focused on corruption, this research aims to take a broader approach by addressing institutional quality, which inherently encompasses corruption among other factors. This decision has also been influenced by the likes of Bucheli (2008), Dasgupta (2021), Sändig (2021) and Dogan (2022), which shows that a country's institutional environment is crucial when assessing biodiversity or transnational agricultural acquisitions in the region.

Acquiring a spatial database for institutional quality would not only be challenging but also unnecessary; country-level aggregation should suffice for the purposes of this research. To this end, the International Property Rights Index (IPRI) provided by the Property Rights Alliance (2024) has been selected. This index, ranging from 0 to 10, comprises components such as legal and political factors, judicial independence, and the rule of law, among others.

However, it is important to acknowledge the limitations of this dataset. The first available year for the IPRI is 2007, posing a temporal constraint on our analysis. To address this obstacle, this research adopts the assumption that IPRI values remained relatively stable between 2000 and 2005 and were comparable to the 2007 levels. As a result, the IPRI values for 2007 will be used as proxies for both 2000 and 2005 in our analysis.

3.2 Data Processing

3.2.1 Unit Approximation

One of the fundamental challenges encountered in this paper was determining the appropriate spatial unit for analysis, considering the dataset and research objectives. Drawing on insights from Davis et al. (2023), it became evident that creating area approximations around LSLA locations would be the most effective approach. These area approximations allow for a more in-depth understanding of the spatial extent and impact of LSLAs on biodiversity in Latin America and the Caribbean.

To generate these area approximations, the size of individual LSLA locations was calculated using deal sizes provided by the Land Matrix database, along with the number of locations assigned to each deal. Essentially, an average size was computed for each location based on these parameters. Subsequently, circular buffers were created around each LSLA location, with the radius of each buffer corresponding to the calculated average size. This method ensures that the area approximations adequately represent the spatial footprint of LSLAs and facilitate subsequent spatial analysis.

Through this approach, we aim to overcome the challenge of spatial representation inherent in LSLA data and provide a more comprehensive understanding of the spatial dynamics of large-scale land acquisitions and their implications for biodiversity in the study region.

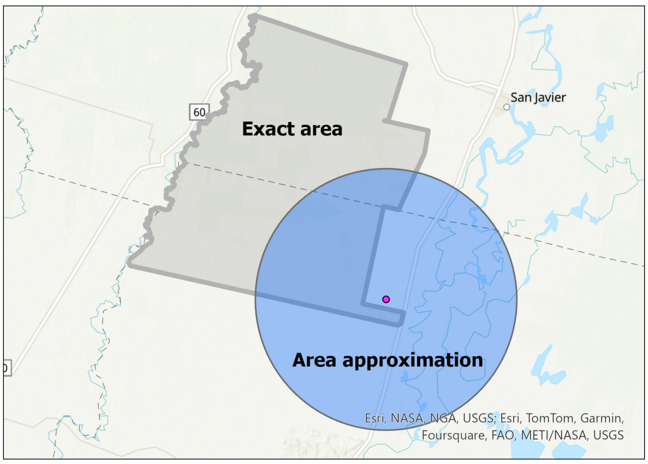


Figure 3.3: Comparison between the exact area and area approximation

Naturally, it is essential to consider that the actual spatial configurations of LSLA locations may differ significantly from the circular approximations generated.

For instance, a long rectangular land parcel could have the same geometric center as a circular approximation. Or as shown by figure 3.3, the location of the points and exact polygons can be slightly misaligned, resulting in a somewhat spatially shifted approximation. However, due to limitations in the availability of polygon data provided by the Land Matrix (2024b), which contains the actual shape and size of these locations, this research deems the use of circular approximations as an acceptable compromise. Nonetheless, it is crucial to bear in mind this limitation when interpreting the results, as the spatial accuracy of the approximations may vary.

By utilizing these area approximations as polygons, we can effectively extract the necessary data from the raster datasets reviewed earlier. This approach enables us to analyze the spatial relationships between LSLA locations and various geographic and environmental factors, providing insights into the potential impacts of large-scale land acquisitions on biodiversity in Latin America and the Caribbean.

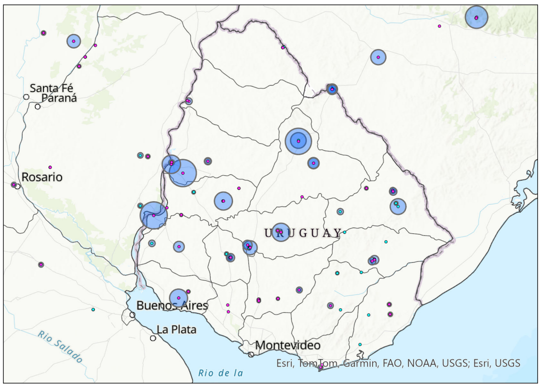
3.2.2 Zonal Statistics

Zonal statistics serves as a crucial tool for calculating statistical indicators from raster datasets, utilizing the spatial extent of another geographic layer, such as another raster or a polygon, as reference. Given the diverse array of raster-based datasets utilized in this research to align with the BHI dataset, these calculations are indispensable for acquiring accurate data for our selected spatial unit.

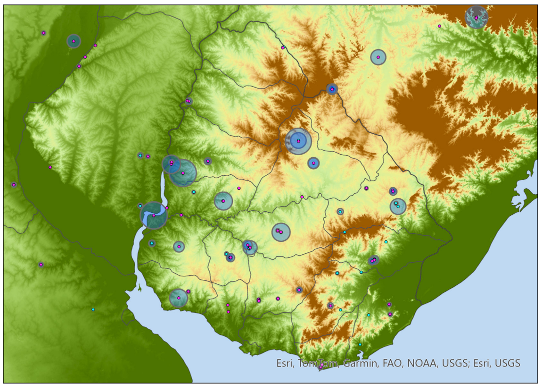
Calculating the weighted geometric means for BHI

Calculating the BHI value for each location in the LSLA dataset was informed by the supplementary paper provided by Harwood et. al. (2022a), which outlines the correct methodology for calculating the mean per territory. This supplement provides crucial insights into an additional raster dataset containing weights that should be utilized for the calculations. Simply put, these weights essentially provide adjustments for the spatial effects on biodiversity.

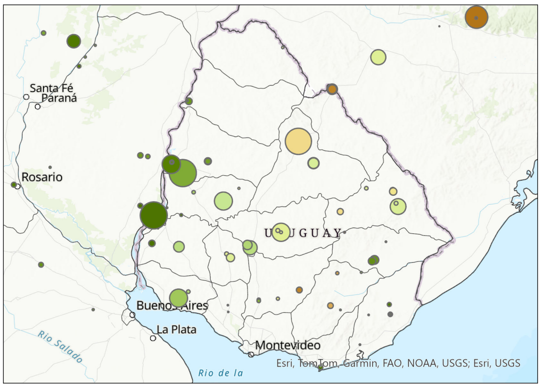
Moreover, the paper highlights that using an arithmetic mean would yield inaccurate results. Instead, it recommends the following formula for calculating the



(a) LSLA area approximations



(b) Elevation raster



(c) Calculated elevation geographic means

Figure 3.4: Zonal statistics through the example of elevation in Uruguay
Color scale adjusted for better visual presentation

weighted geometric means (Harwood et. al. 2022a)[1]:

$$X_r = e^{\frac{\sum_{i=1}^{i=n}[\frac{ln(x_i)}{A_i}]}{\sum_{i=1}^{i=n}[\frac{1}{A_i}]}}$$

Where:

- X : weighted geometric mean of BHI
- x : BHI raster
- A : weight raster
- r : region
- i : individual cell
- n : number of raster cells in the region

[1]The formula presented in this paper deviates slightly from the one in the original source due to a typo concerning the terms for which natural logarithms are required. Employing the original formula may lead to inaccurate outcomes.

In practice, the calculation of weighted geometric means for individual locations has been executed through multiple steps using GIS software. Initially, new raster layers were generated for the denominator and the numerator in each year. These rasters were computed using the following formulas:

$$wgm\_BHI\_num = \frac{ln(x)}{A}$$

Where wgm_BHI_num is the numerator for the weighted geometric mean;

$$wgm\_BHI\_denom = \frac{1}{A}$$

Where wgm_BHI_denom is the denominator for the weighted geometric mean.

Subsequently, zonal statistics tools were employed to determine the sum of cell values from both the denominator and numerator raster sets using the area approximation polygon dataset ($wgm\_BHI\_num\_sum_r$ and $wgm\_BHI\_denom\_sum_r$). These sums were then assigned to each unique location ID.

Finally, the weighted geometric means were computed for each year and location ID utilizing the following formula:

$$wgm\_BHI_r = e^{\frac{wgm\_BHI\_num\_sum_r}{wgm\_BHI\_denom\_sum_r}}$$

In summary, utilizing GIS software, the process involves multiple steps, including the generation of raster layers for both the denominator and numerator, computation of cell value sums using zonal statistics tools, and calculation of weighted geometric means based on a specific formula. This approach ensures the accurate assessment of the BHI across LSLA locations in Latin America and the Caribbean.

Calculating the geometric means of other raster datasets

For other datasets utilized in this research, calculating the geometric means did not require weighting, significantly simplifying the calculation process. Essentially, the geometric means were computed using GIS tools applied to the polygons representing the area approximations, and subsequently assigned to the individual location IDs.

In the case of variables that remain static and are based on raster datasets, such as geography controls, these calculations were conducted once and applied uniformly across all time points of the research. Conversely, for variables characterized by dynamism, such as climate controls and human activity indicators, these calculations were performed for each year of observation, ensuring temporal accuracy and relevance.

3.2.3 Additional Calculations

In addition to zonal statistics, supplementary calculations have been undertaken to supplement the analytical framework. One such calculation focused on determining the distance from bodies of water, which emerged as potentially important control variables for geographical effects. Using GIS software once again, this process required calculating the distance of each point in the location dataset to the nearest line or polygon, depending on whether it meant to rivers or lakes.

3.3 Final Dataset

3.3.1 Summary of the Final Dataset

Following the completion of data cleaning procedures, the final dataset comprises a total of 4043 observations spanning the five observed years. Within this dataset, 2357 observations represent domestic locations, while 1686 are transnational locations, encompassing 22 variables.

To establish a suitable methodology for analysis, the following subsection of the paper will conduct descriptive statistics. This approach aims to uncover the fundamental relationships between the variables, establishing a deeper understanding of the dataset's characteristics and informing subsequent analytical approaches.

3.3.2 Descriptive Statistics

Below, table 3.1 summarizes the source, units and additional characteristics of these variables. Spatial means and distance calculations were performed according to subsection 3.2.2.

To visualize one of the key trends, figure 3.5 presents the yearly state of BHI

Table 3.1: Summary statistics table of the final dataset
Categorical variables - country, location ID and deal ID - are excluded

Statistic	N	Mean	St. Dev.	Min	Max
BHI (0-1)	4,043	0.46	0.13	0.24	0.90
Year of obs.	4,043	2010	7.072	2000	2020
Transnational Dummy	4,043	0.42	0.49	0	1
Avg. Loc. Size (ha)	4,043	17,359.38	46,996.50	10	545,000
Year of impl.	4,043	2010	5.16	2000	2020
Longitude (°)	4,043	−63.40	11.95	−92.72	−37.80
Latitude (°)	4,043	−16.93	15.69	−51.75	21.15
Elevation (m)	4,043	370.12	493.10	2.74	3,951.52
Slope	4,043	2.03	3.29	0.08	22.02
Ruggedness (0-100)	4,043	7.57	11.59	0.48	75.22
Distance - Lakes (°)	4,043	3.63	3.05	0.00	14.22
Distance - Rivers (°)	4,043	0.56	0.60	0.00	5.87
Minimum Temp. (°C)	4,043	15.98	4.90	−2.46	24.48
Maximum Temp. (°C)	4,043	27.66	4.52	8.33	35.12
Precipitation (mm)	4,043	105.36	62.84	0.15	378.55
PCL (0-1)	4,043	0.68	0.21	0.00	0.86
NTL (0-63)	4,043	14.53	18.52	0.00	63.00
GDP (mUSD, 2017)	4,043	3.76	6.63	0.00	40.80
IPRI (0-10)	4,043	4.61	0.84	2.64	6.97

over the observed years, in the format of a boxplot. We see a clear decline over the 5 years in multiple measures. Whether it is mean, or upper/lower quantiles being measured, biodiversity is clearly in a worse state in these locations, based on the BHI measure. This, once again, supports the claims of the prior research reviewed related to the current state of global biodiversity.

Naturally, we also need to determine whether the state of biodiversity in domestic and transnational LSLA locations is different. For this, figure 3.6, a similar boxplot to the prior, shows a much lower level of BHI in transnational locations. This supports the findings of Davis et. al. (2023) saying transnational agricultural LSLAs are in fact a danger to local biodiversity. However, this plot provides a comparison to their domestic counterparts.

Combining the previous two plots, figure 3.7 shows the decline of the BHI over the 5 observed years, categorized into domestic and transnational LSLAs. However, something interesting can be seen here: the BHI in transnational locations appears to decline in a slower pace compared to domestic locations, even though BHI levels are generally lower in this category. While this essentially proves **H1**, it somewhat

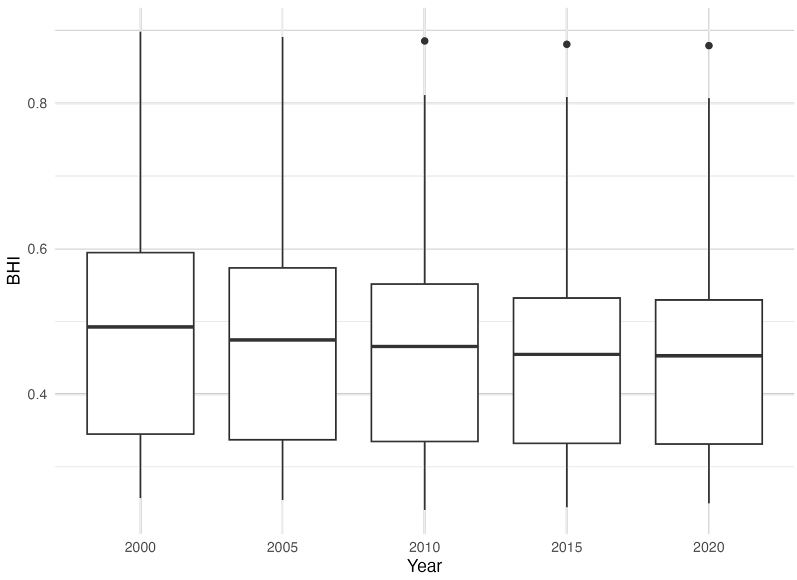


Figure 3.5: Box plot of BHI per year

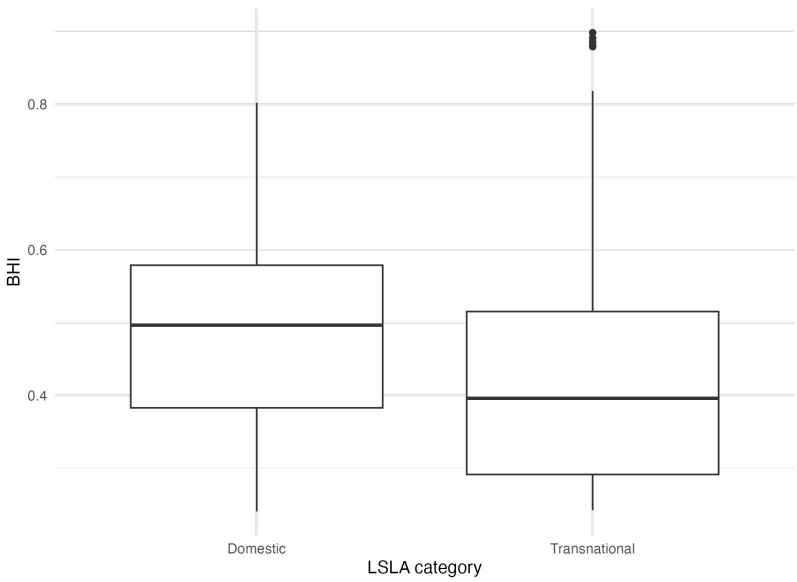


Figure 3.6: Box plot of BHI per LSLA type

counters **H2** of the research, and is thus investigated closely in the empirical research.

To gain more insights on this relationship, figure 3.8 shows a combination of 4 density plots, divided by the first and last years of observations, and the category of LSLAs. As also advised by the box plots of figure 3.7, these scenarios show a negative shift in time, or when the location is transnational. However, while the two plots showing domestic locations somewhat resemble a normal density curve, the other two show two distinct peaks in the density of the BHI, indicating the existence of additional sub-groups that divide this category. This finding is also something considered in the upcoming empirical analysis.

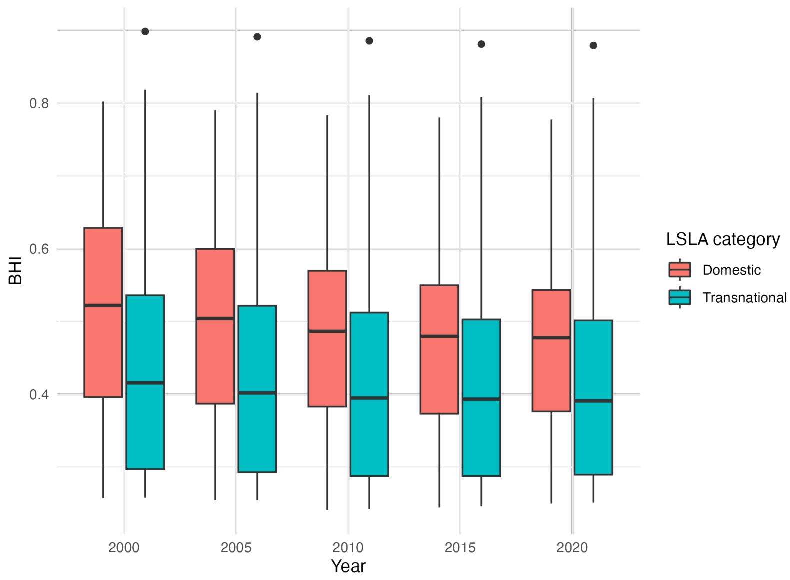


Figure 3.7: Box plot of BHI per year and LSLA type

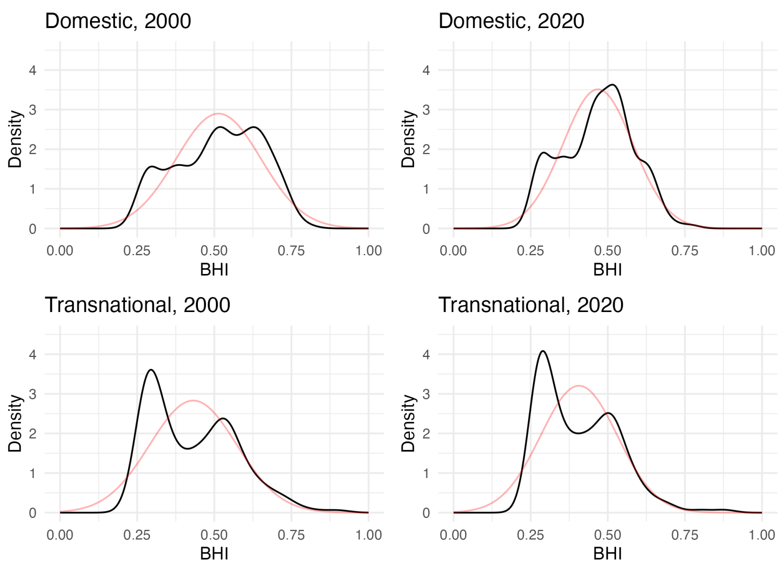


Figure 3.8: Density plot of BHI, by year and LSLA type

Shifting our attention from the relationships of the dependent to the main independent variable of research, the scatter plots of figure 3.9 and figure 3.10 show the relationship between the BHI and geography controls. It is important to note that all of these controls except the PCL index had to be log-adjusted to better observe these relationships - something to also consider during the empirical analysis.

Potentially the most important geography control, the location size, is presented in relation to biodiversity in figure 3.9. While the fitted line shows slight upward trend in the BHI when the location size increases, the scatter plots paints a much more noisy picture. The insight we can draw from this is there is no clear relationship between the two variables - however, including this control is still important for the upcoming analysis, as it can change in relation to other factors.

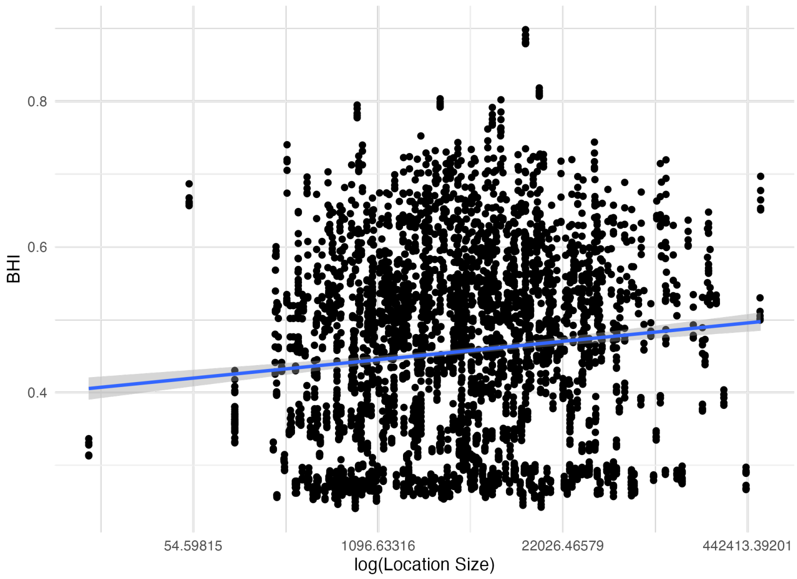


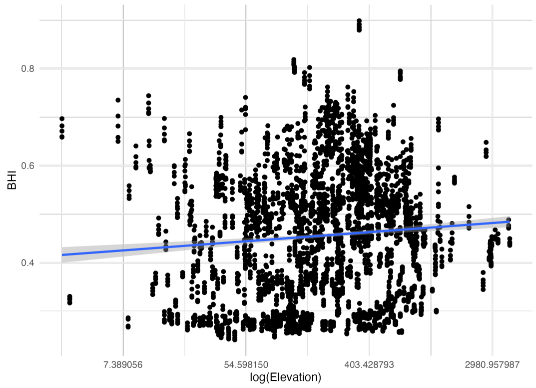
Figure 3.9: Scatter plots and linear approximations between BHI and average location size

For the rest of the geography controls presented in figure 3.10, we see a similar scenario: while the fitted lines do show some upwards or downwards trends, this is likely due to outliers in the dataset - most of it is noise. However, it is important to note that the relationship of BHI to the average slope and average ruggedness looks very similar, indicating correlation between the two variables.

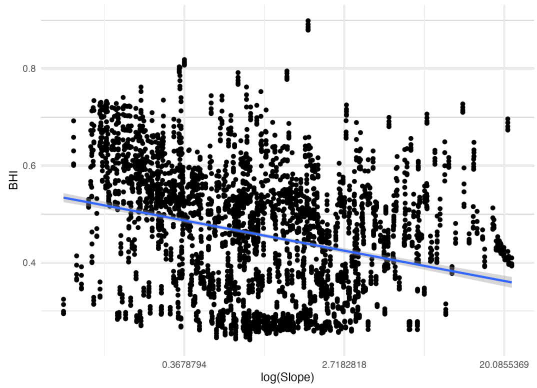
The scatter plot of figure 3.11 confirms a high correlation between the two variables: the observations almost perfectly sit on the fitted line of the graph. Thus moving forward, one of these variables, namely ruggedness, will be excluded from the dataset, to avoid auto-correlation in the upcoming analysis.

Similar to figure 3.10, figure 3.12 shows the scatter plots of the BHI and the climate variables. Once again, the relation of these controls to the dependent variable is unclear, but should be included in the analysis due to the spatial and temporal nature of the research.

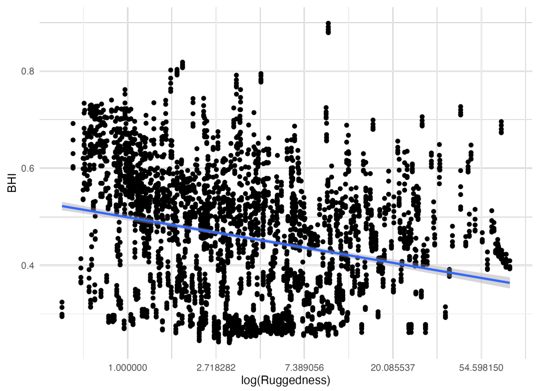
However, potential correlation is once again apparent between two explanatory variables. Naturally, the change of average minimum and maximum temperatures will go somewhat hand in hand. The scatter plot of figure 3.13 confirms this: the two independent variables are highly correlated in the case of this dataset too. Again to avoid auto-correlation, average minimum temperature will be removed from the



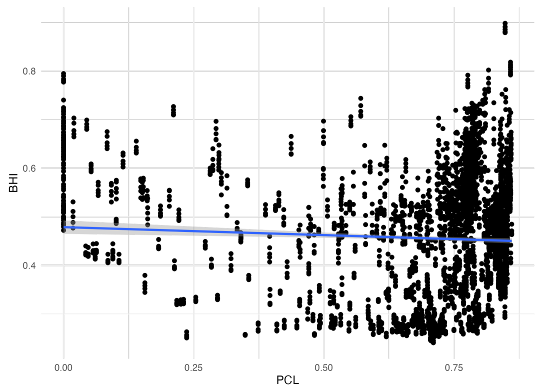
(a) log(Elevation)



(b) log(Slope)



(c) log(Ruggedness)



(d) PCL

Figure 3.10: Scatter plots and linear approximations between BHI and other geography controls

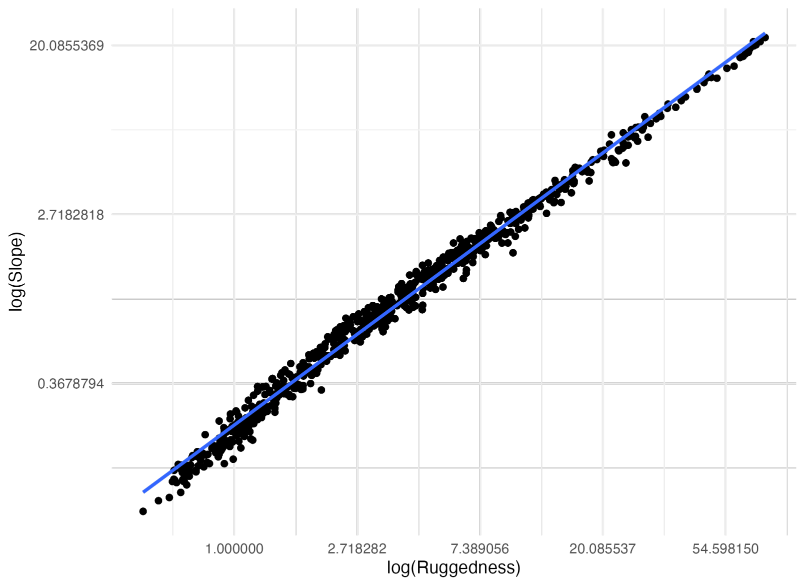
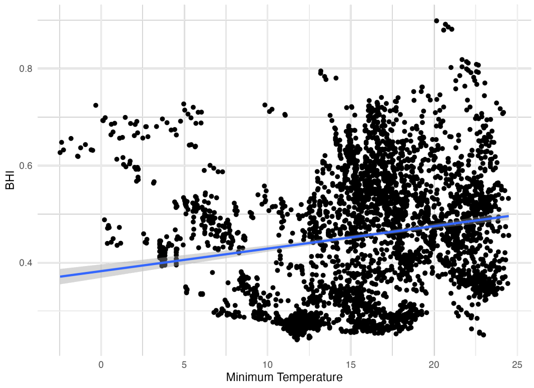
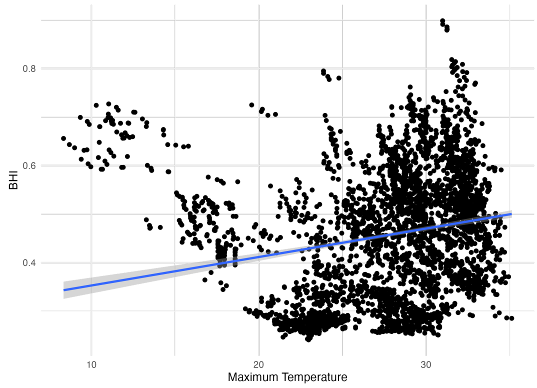


Figure 3.11: Scatter plots and linear approximations between slope and ruggedness

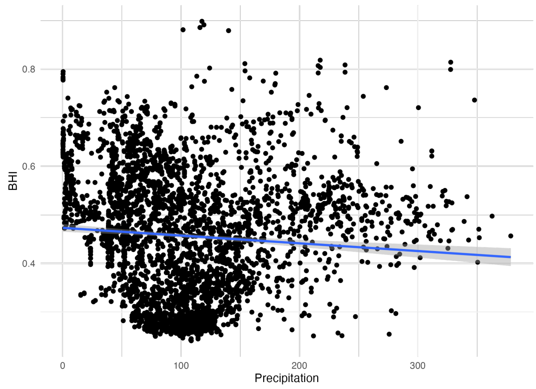
upcoming analysis. Average maximum temperature will be kept instead, as the rising high temperatures have a bigger effect on biodiversity change, this will likely hold more explanatory power in our analysis.



(a) Minimum Temperature



(b) Maximum Temperature



(c) Precipitation

Figure 3.12: Scatter plots and linear approximations between BHI and climate controls

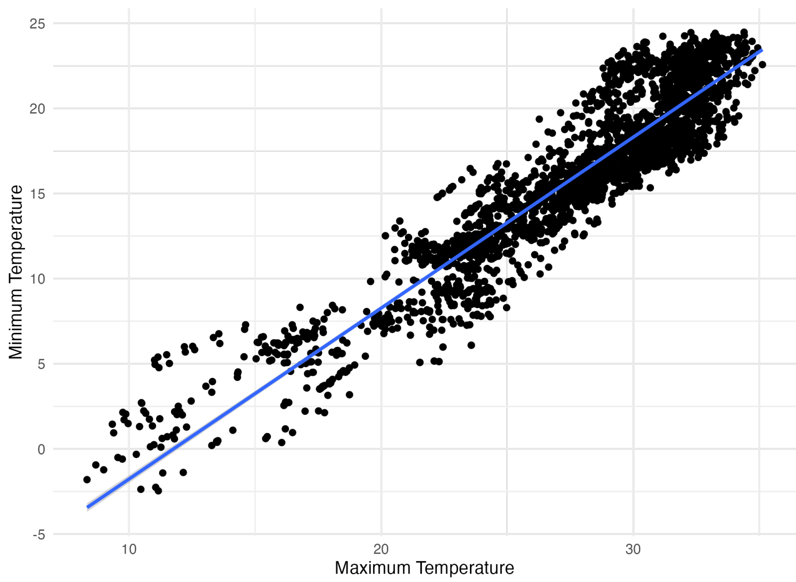
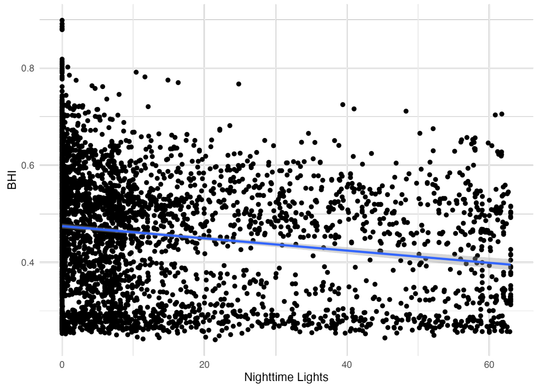


Figure 3.13: Scatter plots and linear approximations between minimum and maximum temperature

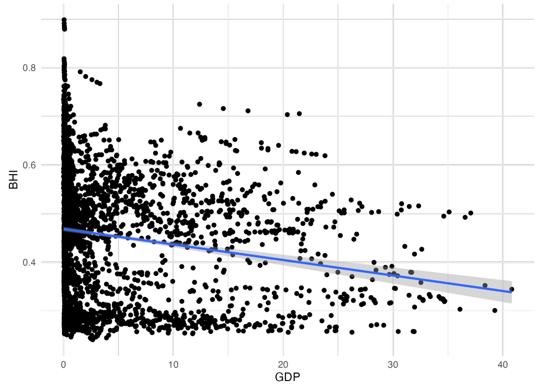
For the last group of controls, for human activities, figure 3.14 once again shows scatter plots in relation to the BHI. Here the trends seem somewhat clear for NTL and GDP data, both the scatters and the fitted lines show a somewhat negative relationship to the dependent variable.

However, this shows another case of possible autocorrelation, as the distribution of observations looks similar in both cases. To confirm this, figure 3.15 shows another scatterplot between the NTL and the natural logarithm of the GDP variables. While not as strong as in the two previous cases, this also shows correlation between the two variables. This of course makes sense, as it was prevously discussed that Chen et. al.'s (2022) spatial GDP database is in fact adjusted by NTL data. As the correlation between these two independent variables is not as precise as the previous examples, and given the high importance of both factors for this research, these will be both included in the subsequent analysis - although in separate models, to once again avoid auto-correlation.

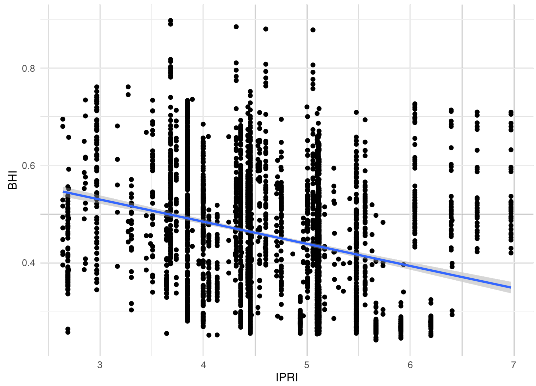
Regarding the IPRI, figure 3.14c shows even though the fitted line shows otherwise, there is once again no clear correlation between the BHI and IPRI - but this could again change when not viewed in isolation.



(a) NTL



(b) GDP



(c) IPRI

Figure 3.14: Scatter plots and linear approximations between BHI and human activity controls

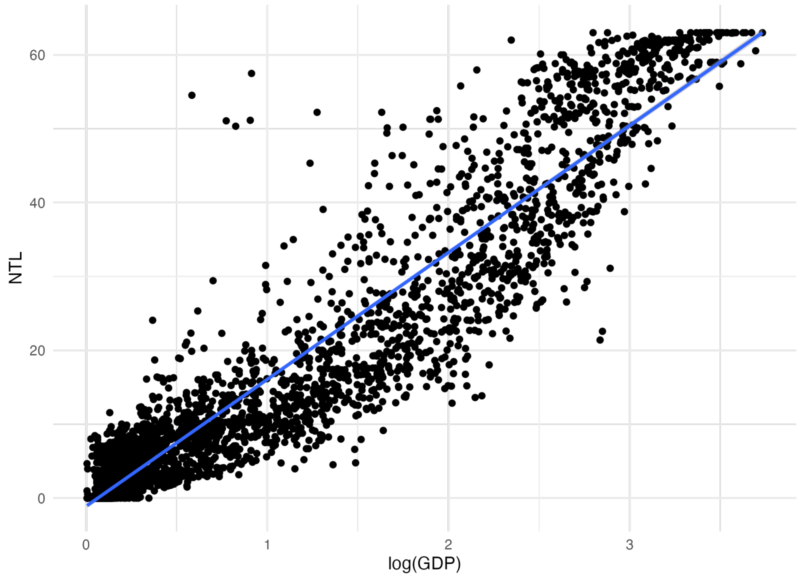


Figure 3.15: Scatter plot and linear approximation between NTL and GDP

Having discussed the basic relationships and descriptive statistics of the key variables, the focus now shifts to the empirical methodologies employed in this research.

4

Empirical Analysis

4.1 Methodology

4.1.1 Integration of Theory and Data

To reiterate, quantitative methodologies are employed in this research, with the state of biodiversity (BHI) serving as the dependent variable and the primary independent variable being whether the LSLA in question is a transnational or domestic investment, represented by the Transnational Dummy variable. To address the research question and test the two hypotheses, a difference-in-differences (DiD) method is deemed most suitable. The specifics of this method are further discussed in subsection 4.1.2. Additionally, the general formulad use throughout this analysis is explained in subsection 4.1.3.

Prior quantitative research reviewed in chapter 2 suggests the use of multiple regression models to adequately explore the effects in question. Employing a variety of models is crucial as each model has its shortcomings, and utilizing multiple approaches will assist in selecting the most appropriate one. It is common practice to commence such research with OLS models. Although the bias associated with OLS is well-known, it remains a valuable tool for assessing the basic relationship between the variables. This is elaborated upon in subsection 4.1.4.

In addition to OLS, fixed effects models will be utilized to control for time-variant unobserved heterogeneity and to mitigate endogeneity by accounting for individual-specific effects. Detailed discussion on this can be found in subsection 4.1.5. To further refine the methods, full matching will be employed, which assigns weights to the main independent variable, transnational dummy. This methodology is also employed due to the discrepancies observed between domestic and transnational locations, as illustrated in figure 3.7 and figure 3.8. Utilizing a matched dataset is expected to help bring the two groups of LSLAs closer together, addressing these discrepancies. The details of this approach are provided in subsection 4.1.6. To ensure thoroughness, the results will be subjected to robustness checks, which will be discussed further in subsection 4.1.7.

Based on insights from Davis et al. (2023), the average location size in the sample will be set above 200 hectares, as this criterion provides an adequate definition of large-scale acquisitions. Additionally, the regression models will employ the natural logarithm of the BHI to aid with a more straightforward interpretation of the results.

4.1.2 Difference in Differences

To address the research question and hypotheses outlined in the previous section, an appropriate methodological framework must be established. The empirical analysis will utilize the difference-in-differences (DiD) methodology. This approach is selected due to its suitability in assessing the impact of TALSLAs on local biodiversity, taking into account the temporal and spatial dimensions of the data structure.

The difference-in-differences methodology is a quasi-experimental design used to estimate causal relationships by comparing the changes in outcomes over time between a treatment group and a control group. The key advantage of this approach is its ability to control for unobserved confounding factors that might affect the treatment and control groups differently over time. This method has been extensively discussed in the literature (e.g. Angrist & Pischke, 2008 p. 169-182) and will be an effective tool in uncovering the impacts of these investments.

In this research, the "treatment" is defined as the presence of transnational LSLAs. The control group comprises domestic LSLAs. The primary outcome of interest is the change in the BHI before and after the implementation of these LSLAs. The pre-treatment period is set at the year 2000, representing the initial BHI levels before the large-scale implementation of LSLAs. The post-treatment period is set at 2020, the latest year for which BHI data is available.

The LSLA data has been subsetted to align with these temporal boundaries. Although this results in a reduced number of total observations, it simplifies the analysis while still yielding robust and meaningful results. The reduction in observations is considered an acceptable trade-off for the clarity and precision this methodology provides.

An essential aspect to consider is the year of implementation for each LSLA. Variations in the time elapsed between the LSLA implementation and the post-treatment period will exist across different locations. This variation could potentially influence the observed impact on biodiversity. The timing of LSLA implementation will be further discussed and controlled for in the subsequent sections to ensure the validity and reliability of the results.

4.1.3 The General Formula

Using the DiD methodology, the analysis will incrementally introduce control variables while constructing the comprehensive regression models. The initial step involves testing the effects of the three crucial dummy variables: transnational, time, and DiD. Subsequently, geography controls will be introduced, followed by climate controls, and finally, human activity controls.

Given the high correlation of the explanatory variables, as discussed in relation to Figure figure 3.15, human activity controls will first employ data from the GDP database. This will then be compared with control data from the NTL database. As the final step, a country interaction with the DiD variable will be incorporated into the models to uncover whether the effects of these locations are context-dependent. The general formula for the regressions is presented in Appendix B.

4.1.4 OLS Models

OLS models are a commonplace tool in econometric research due to their simplicity and accessibility. They provide quick results and an initial overview for assessing relationships between variables, serving as a valuable preliminary step before employing more refined methodologies. In this analysis, OLS models are used to establish the foundational relationships between the variables under consideration.

The control variables are gradually introduced into the OLS models, as discussed in subsection 4.1.3.

Given the delayed effects of LSLAs on biodiversity, as discussed earlier, the OLS models also include subsets of data based on the implementation year of the LSLAs. This approach helps in understanding the temporal dynamics of LSLA impacts on biodiversity.

Practically, these models have been constructed using the `lm()` function in R (RDocumentation, 2024). Despite the utility of OLS models, it is important to acknowledge their limitations. As Baltagi (2021b, p. 48) notes, OLS estimates can be biased and inconsistent under certain conditions. Therefore, the analysis also employs more refined methods to ensure robustness and reliability of the findings.

4.1.5 Fixed Effects Models

The next step of the analysis employs two-way fixed effects (FE) models to control for unobservable individual and time effects in the regressions, as suggested by Baltagi (2021b, pp. 47-50). Implementing these models is expected to result in a more robust and reliable analysis, thereby strengthening the findings.

Once again, additional controls were gradually added to the FE models, in accordance with subsection 4.1.3.

In practice, these regressions were conducted in R, using the `plm()` function from the `plm` package (Croissant & Millo, 2023), specifically employing the `"within"` model. The formula indexes were set to the individual location ID and the year variables, representing the time-invariant, individual-specific effects in these models.

Although these FE models are more refined than ordinary least squares (OLS) models, there remains room for further refinement to enhance the robustness of the analysis.

4.1.6 Matched Fixed Effects Models

To further enhance the robustness of the analysis, full matching has been employed to assign weights to the observations. This method allows for the creation of a weighted sample that accounts for potential confounding variables, thereby improving the precision of the estimates and reducing bias.

Hansen (2004), and later Hansen & Klopfer (2006) describe this method as the most effective approach for matching control and treatment groups in an experimental study, which they promptly demonstrate through the examples of coached and uncoached SAT scores, and women's and men's working conditions, respectively. In a nutshell, this methodology effectively reduces the disparity between the standard deviations of the two groups to a minimal level. While this will put more tension on the model, it will also show us a more true effect that the independent variables exert on the BHI.

This analysis' full matching relies on the assumption that the transnational dummy is influenced by several contextual factors of the location. Specifically, the institutional context (represented by the natural logarithm of the IPRI variable), the economic context (represented by the natural logarithm of the GDP variable), and the quality of land (represented by the PCL variable). These variables were used in the matching process to ensure a balanced comparison between treated and untreated groups.

In practice, the matching was conducted using the `MatchIt` package (Greifer, 2023). This process assigned weights to the observations, which were subsequently utilized in the regression models constructed with the aforementioned `plm` package (Croissant & Millo, 2023). This approach aims to provide a more reliable and accurate assessment of the effects of transnational investments on biodiversity.

The weights assigned to the control group (i.e. domestic locations) are summarized in figure 4.1. The treatment group (i.e. transnational locations) was assigned to a weight of 1, as the other observations were weighted to them.

4.1.7 Robustness Checks

To ensure the validity and robustness of the results, additional analyses were performed. Random effects models were constructed based on the methodology suggested by Baltagi (2021b, pp. 47, 50-55). These models were implemented again using the `plm` package (Croissant & Millo, 2023), specifically employing the `"random"` model.

Furthermore, Lagrange multiplier (LM) tests were utilized to determine whether significant random effects were present in the data, indicating the preference for

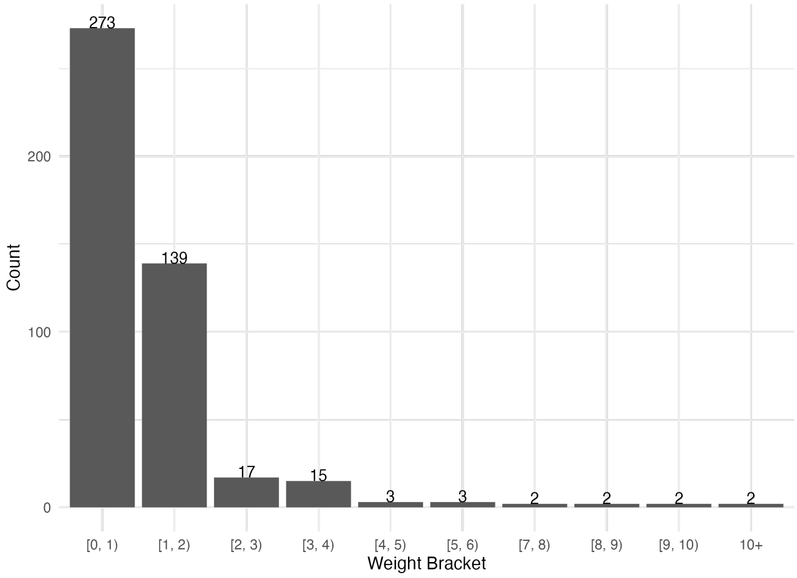


Figure 4.1: Weight brackets of the control group

models other than OLS. This approach follows the guidelines provided by Baltagi (2021a, pp. 81-82). Additionally, Hausman tests were conducted to validate the use of fixed effects models over random effects models, again as recommended by Baltagi (2021a, pp. 89-94). The tests were executed using the `plmtest()` and `phtest()` functions from the `plm` package in R (Croissant & Millo, 2023). These robustness checks aimed to confirm the appropriateness of the chosen modeling approaches and to ensure the reliability of the findings.

4.2 Results

This section will present the findings from each stage of the analysis, progressing from the basic OLS models to the matched FE models, ending with the robustness checks. For all models, clustered robust standard errors were calculated using the `vcovCR` function from the `clubSandwich` package in R (Pustejovsky, 2023). This approach follows the recommendations of Moody and Marvell (2020) to mitigate the potential effects of autocorrelation (also discussed in Mihálka, 2023). The regression tables were generated using the `Stargazer` R package by Hlavac (2022).

4.2.1 OLS Results

Full sample

The results for the full sample OLS models are presented in table 4.1, with the first column of table A.2 detailing the outcomes for model 1.6. The findings indicate a consistent and significant negative correlation between log(BHI) and the transnational dummy variable. The time dummy variable also exhibits a negative sign across all models; however, it loses statistical significance when human activities are incorporated in model 1.4.

Noteworthy results are observed for the DiD dummy variable. The coefficients for this variable are significant in all models, although their direction remains positive until country interactions are introduced. The exception to this pattern occurs with country-specific interactions, such as those for Brazil, Chile, and Mexico, where some interaction coefficients achieve statistical significance. This suggests that the DiD effect of transnational investments varies by country, highlighting the importance of country-specific factors in understanding the impact of these investments on biodiversity.

Table 4.1: OLS results

	Dependent variable:					
	log(BHI)					
	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(1.6)
Transnational Dummy	-0.18701^{***}	-0.07428^{***}	-0.07724^{***}	-0.08064^{***}	-0.08238^{***}	-0.04668^{***}
	(0.02225)	(0.02019)	(0.02060)	(0.02032)	(0.02016)	(0.01605)
Time Dummy	-0.08466^{***}	-0.08466^{***}	-0.08550^{***}	-0.04259	-0.02379	0.01313
	(0.00263)	(0.00263)	(0.00785)	(0.02671)	(0.02671)	(0.03964)
DiD Dummy	0.03214^{***}	0.03051^{***}	0.03672^{***}	0.03464^{***}	0.03510^{***}	-0.09457^{***}
	(0.00371)	(0.00391)	(0.00501)	(0.00538)	(0.00517)	(0.02350)
Interacted with countries	No	No	No	No	No	Yes
Constant	-0.70337^{***}	-1.43997^{***}	-2.21885^{***}	-2.02354^{***}	-2.03047^{***}	0.26978
	(0.01367)	(0.12871)	(0.24304)	(0.27973)	(0.27631)	(0.26117)
Controls for:						
Geography	No	Yes	Yes	Yes	Yes	Yes
Climate	No	No	Yes	Yes	Yes	Yes
Human Activity	No	No	No	Yes (GDP)	Yes (NTL)	Yes (NTL)
Observations	1,598	1,598	1,598	1,598	1,598	1,598
R^2	0.09111	0.38472	0.41478	0.43601	0.44068	0.70662
Adjusted R^2	0.08940	0.38045	0.40997	0.43066	0.43537	0.69792
Residual Std. Error	0.29072 (df = 1594)	0.23980 (df = 1586)	0.23402 (df = 1584)	0.22988 (df = 1582)	0.22893 (df = 1582)	0.16745 (df = 1551)

Note: Clustered robust standard errors are in parentheses
*p<0.1; **p<0.05; ***p<0.01

OLS with different sample subsets

The results for the limited sample OLS models are presented in table 4.2. For subsets of LSLAs implemented before 2015 (model 2.2) and 2010 (model 2.3), similar results to the full sample (model 2.1) are observed. In these models, the transnational and DiD dummies are significant and negative, while the time dummy remains insignificant.

However, when the data is limited to LSLAs implemented before 2005 (model 2.4), the main explanatory variables lose their significance. This change is likely attributable to the much lower number of observations (302), which potentially causes imbalances in this specific model.

Going forward, the subset used for model 2.3 will be employed in the research for several reasons. First, the sample size remains large enough, with 850 observations, as anything smaller would be insufficient. Second, this subset introduces a ten-year difference between the last year of implementation and the state of the BHI in 2020. Third, the coefficients, with the country interactions further discussed in column 2 of table A.2, continue to show meaningful results relevant to the research.

Table 4.2: OLS results with different subsets

	Dependent variable:			
	log(BHI)			
	(2.1)	(2.2)	(2.3)	(2.4)
Transnational Dummy	-0.04668^{***}	-0.05338^{***}	-0.06239^{**}	0.01506
	(0.01605)	(0.01754)	(0.02756)	(0.03478)
Time Dummy	0.01313	0.02683	−0.04806	−0.02488
	(0.03964)	(0.05278)	(0.07309)	(0.11986)
DiD Dummy	-0.09457^{***}	-0.09085^{***}	-0.10737^{***}	−0.02021
	(0.02350)	(0.02460)	(0.02783)	(0.02243)
Constant	0.26978	0.53463^{*}	0.05971	−0.87376
	(0.26117)	(0.30387)	(0.40175)	(0.65412)
Year of implementation between 2000 and:	2020	2015	2010	2005
Observations	1,598	1,252	850	302
R^2	0.70662	0.70399	0.70713	0.78776
Adjusted R^2	0.69792	0.69269	0.69150	0.75523
Residual Std. Error	0.16745 (df = 1551)	0.16535 (df = 1205)	0.16213 (df = 806)	0.13422 (df = 261)

Note: Clustered robust standard errors are in parentheses
All models contain geography, climate and human activity control variables
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

4.2.2 FE Results

The results for the initial FE models are presented in table 4.3, with the country interactions collected in the third column of table A.2. For these FE models, the data has been subset to include locations with a year of implementation before 2010. Additionally, observations from Ecuador have been excluded due to data structuring issues.

The results show that the transnational dummy is omitted in all cases, likely due to high autocorrelation with the newly introduced fixed effects for location IDs and years of observation. The time dummy retains its significance throughout all models, but interestingly, it flips to positive effects when human activity controls are introduced in model 3.4.

The DiD dummy is significant and positive across all models, with only some country interactions showing a negative relationship between the variable and log(BHI). The model fit has also improved significantly, with the R^2 increasing from 0.69 in the subsetted OLS model 2.3 to 0.79 in model 3.6. This improvement indicates a better explanatory power of the FE models, suggesting a more accurate depiction of the relationships between the variables under study.

Table 4.3: FE Results

	Dependent variable:					
	log(BHI)					
	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)	(3.6)
Transnational Dummy	*(omitted)*	*(omitted)*	*(omitted)*	*(omitted)*	*(omitted)*	*(omitted)*
Time Dummy	-0.09511^{***}	-0.09511^{***}	-0.01786^{***}	0.10610^{***}	0.10944^{***}	0.15684^{***}
	(0.00365)	(0.00365)	(0.00373)	(0.01599)	(0.01566)	(0.02217)
DiD Dummy	0.04170^{***}	0.04170^{***}	0.02602^{***}	0.02140^{***}	0.02084^{***}	0.03267^{***}
	(0.00478)	(0.00478)	(0.00368)	(0.00351)	(0.00350)	(0.00620)
Interacted with countries	No	No	No	No	No	Yes
Controls for:						
Geography	No	Yes	Yes	Yes	Yes	Yes
Climate	No	No	Yes	Yes	Yes	Yes
Human Activity	No	No	No	Yes (GDP)	Yes (NTL)	Yes (NTL)
Observations	846	846	846	846	846	846
R^2	0.71412	0.71412	0.84467	0.87297	0.87266	0.90022
Adjusted R^2	0.42620	0.42620	0.68675	0.74258	0.74197	0.79131

Note: Clustered robust standard errors are in parentheses. Sample is set to year of impl. between 2000 and 2010
Observations of Ecuador excluded due to data structuring issues
*p<0.1; **p<0.05; ***p<0.01

4.2.3 Matched FE Results

The results for the matched FE models are presented in table 4.3, with the country interactions collected in the fourth column of table A.2. In these models, human activity controls and the control for PCL are excluded, as they were used to construct the weights through the transnational dummy.

Once again, the transnational dummy has been automatically omitted, likely due to high autocorrelation with the fixed effects. The time dummy is mostly negative and retains its significance only until the climate control variables are introduced.

The DiD dummy is significant and positive in all cases, consistently producing negative correlations through country interactions. For model 4.4, this suggests a negative DiD effect for 10 of the countries included in the analysis. This finding implies that the impact of transnational investments on biodiversity is heterogeneous across different countries, with several exhibiting adverse effects.

The exclusion of specific controls in the matched models, while necessary for the construction of weights, does not appear to diminish the robustness of the findings. The significant and positive DiD dummy across all models underscores the reliability of the results, despite the variations in country-specific interactions.

4.2.4 Summary of Different Models

To summarize these results, table A.1 provides a comparison between the different stages of the analysis. While the transnational and time dummies vary in significance throughout these steps, the DiD dummy remains significant in all models. However, its overall direction changes when FE are introduced.

The initial OLS models show a significant negative correlation between log(BHI) and the transnational dummy. The time dummy, though consistently negative, loses statistical significance when human activities are controlled for. The DiD dummy, while significant in all cases, changes direction upon the introduction of country interactions, suggesting a country-dependent effect of transnational investments on biodiversity.

When examining the limited sample OLS models, similar results are observed for subsets of LSLAs implemented before 2015 and 2010. The transnational and DiD dummies are significant and negative, while the time dummy remains insignif-

Table 4.4: Matched FE Results

	Dependent variable:			
	log(BHI)			
	(4.1)	(4.2)	(4.3)	(4.4)
Transnational Dummy	*(omitted)*	*(omitted)*	*(omitted)*	*(omitted)*
Time Dummy	-0.08105^{***}	-0.08105^{***}	-0.00415	0.00619
	(0.00375)	(0.00375)	(0.00461)	(0.00490)
DiD Dummy	0.02764^{***}	0.02764^{***}	0.01326^{***}	0.01711^{**}
	(0.00485)	(0.00485)	(0.00383)	(0.00669)
Interacted with countries	No	No	No	Yes
Controls for:				
Geography	No	Yes	Yes	Yes
Climate	No	No	Yes	Yes
Observations	846	846	846	846
R^2	0.71080	0.71080	0.83963	0.87120
Adjusted R^2	0.41955	0.41955	0.67658	0.73193

Note: Clustered robust standard errors are in parentheses.
Sample is set to year of impl. between 2000 and 2010
Observations of Ecuador excluded due to data structuring issues
*p<0.1; **p<0.05; ***p<0.01

icant. However, for LSLAs implemented before 2005, the significance of the main explanatory variables is lost, likely due to the reduced number of observations. Consequently, the subset used for model 2.3, which includes LSLAs implemented before 2010, is employed going forward due to its sufficient sample size and meaningful results.

In the FE models, the transnational dummy is omitted, likely due to high autocorrelation with the fixed effects for location IDs and years of observation. The time dummy remains significant but flips to a positive effect when human activity controls are introduced. The DiD dummy is significant and positive throughout, with only some country interactions showing a negative relationship between the variable and log(BHI). The model fit improves significantly, with an R^2 increasing from 0.69 in the subsetted OLS model 2.3 to 0.79 in model 3.6.

The matched FE models continue to show the omission of the transnational dummy due to high autocorrelation. The time dummy remains mostly negative and loses significance with the introduction of climate control variables. The DiD dummy is consistently significant and positive, producing negative correlations through country interactions. This suggests a negative DiD effect for several countries included in the analysis, indicating heterogeneous impacts of transnational investments on biodiversity.

Overall, while the significance of the transnational and time dummies varies, the consistent significance of the DiD dummy highlights the robustness of the findings. The changes in the direction of the DiD dummy when FE are introduced underscore the importance of accounting for fixed effects in understanding the impact of transnational investments on biodiversity. The consistently high model fit, ranging from an adjusted R^2 of 0.69 in the OLS models to 0.79 in the initial FE model, further supports the reliability of the results.

4.2.5 Robustness checks

In table A.3 and table A.4, the results of the random effects models, which serve as the robustness check for the analysis, are shown. The DiD dummy remains consistently significant, except in model 6.4, where the random effects model introduces country interactions with the variable. These results further confirm the validity of the findings presented in the previous subsections.

The outcomes of the LM and Hausman tests, as displayed in table A.5, provide additional support. The LM test results show high chi-squared values and extremely low p-values, indicating the presence of panel-level effects in the dataset. Similarly, the Hausman test results suggest that the random effects estimator is inconsistent, favoring the fixed effects model.

In summary, the robustness checks validate the findings of the FE models, confirming that the results are both valid and robust.

4.3 Discussion

This research addresses a significant gap in the literature regarding the impact of TALSLAs on biodiversity compared to their domestic counterparts. In summary, the analysis, as illustrated in figure 3.7 and corroborated by the regression models, supports **H1**. Specifically, it was found that agricultural LSLAs have a significant negative effect on local biodiversity in Latin America and the Caribbean.

However, a more nuanced understanding was obtained for **H2**. While all comprehensive models indicated negative DiD coefficients in some countries, the findings were not universally applicable across all nations in the analysis. This led to a partial confirmation of **H2**: it was confirmed in some countries but not in others.

Therefore, in response to the research question, it can be concluded that LSLAs for agriculture generally have a detrimental effect on local biodiversity. Furthermore, transnational investments exacerbate these negative outcomes in specific countries, such as Bolivia and Colombia.

Even though these results might not confirm both hypotheses universally, this research still provides robust, interesting, and valuable findings, fills the gap in prior research discussed in chapter 2. One particularly intriguing discovery is that the initial biodiversity levels at locations of transnational LSLAs are worse off compared to their domestic counterparts. This disparity is hypothesized to be influenced by institutional settings, though identifying the precise reasons falls outside the scope of this research. Future studies should explore this aspect to gain deeper insights into why this might be the case, potentially leading to significant advancements in biodiversity protection in such areas.

Furthermore, the comprehensive regression models highlight that the DiD effect of transnational LSLAs is highly country-dependent. Surprisingly, some countries exhibit a positive DiD effect from transnational LSLAs, a result not anticipated before. This variation is likely attributable to institutional settings among other factors.

Additionally, the inclusion of controls for geography, climate, and other human activities has proven beneficial in researching changes in biodiversity, as evidenced by the improved model fit in most cases. Therefore, it is recommended that future quantitative research on biodiversity incorporates these variables to enhance the accuracy and relevance of their findings.

Future research should concentrate on case studies of different countries to uncover why the presence of transnational acquisitions impacts biodiversity so differently across the region. Such research should ideally begin with qualitative methodologies to identify the underlying mechanisms driving these variations. The insights gained from these qualitative studies can then inform more refined quantitative research, leading to a comprehensive understanding of the phenomena.

Moreover, the findings of this research have significant implications for policymakers. The demonstrated negative impact of transnational agricultural LSLAs on biodiversity in at least some cases suggests that greater caution is needed when permitting transnational companies to operate within their borders. Policymakers should ensure that robust institutional frameworks are in place to mitigate the potential harmful effects of agricultural practices on biodiversity. This includes developing and enforcing regulations that prioritize the protection of local ecosystems and biodiversity hotspots, ensuring that economic development does not come at the expense of environmental sustainability.

This research is, however, not without its limitations. First and foremost, approximating agricultural influence through the areas of LSLA is not ideal for several reasons. For instance, without knowing the exact size and precise locations of these LSLAs, the initial dataset might be spatially biased. Furthermore, the absence of a comprehensive database of areas entirely untouched by LSLAs prevents the establishment of a true control group for comparing domestic and transnational investments. Despite these limitations, due to the low data availability on this topic, these approximations still provide valuable insights until more detailed data becomes accessible.

Second, spatial autocorrelation is likely present in the final dataset, a factor that requires even more advanced methodologies to control effectively. Addressing spatial autocorrelation, for example through the usage of geographically weighted regressions could enhance the robustness of future research findings.

Third, due to data limitations, the analysis presented here cannot fully account for temporal variations. More comprehensive data sources and refined methodologies could help future research to better control for time-related biases, leading to more precise and accurate results.

In summary, while this research contributes valuable findings to the understanding of how transnational agricultural LSLAs affect biodiversity, addressing these limitations in future studies will be crucial for further refining and validating the results.

5

Conclusion

This thesis has sought to address the impact of transnational versus domestic large-scale land acquisitions (LSLAs) in agriculture on biodiversity in Latin America and the Caribbean. The study fills a critical gap in the literature by providing empirical evidence on how different types of investments influence local biodiversity, using a comprehensive quantitative analysis. The findings present a nuanced understanding of the relationship between LSLAs and biodiversity, offering valuable insights for policymakers and future researchers.

5.1 Summary of Data and Methodologies

The analysis utilized spatial panel datasets covering the BHI for the period 2000-2020, obtained from Harwood et al. (2022b). The areas affected by LSLAs were approximated using data from the Land Matrix databases (2024a & b). Various control variables were also incorporated to account for geographical, climatic, and human activity influences on biodiversity. The BHI was used as the primary measure to represent biodiversity comprehensively, aligning with the need for robust and detailed monitoring tools discussed in the literature review.

A DiD methodology was employed to isolate the impact of LSLAs on biodiversity. The analysis started with OLS models to establish baseline relationships. Subsequently, more refined models were used, including FE models to control for unobservable individual and time effects, and matched FE models based on full matching to address potential biases and discrepancies between domestic and transnational

LSLA locations. In constructing the regression models, control variables were incrementally introduced. Geography controls were added first, followed by climate controls, and then human activity controls. Finally, country interactions with the DiD variable were included to capture country-specific effects.

5.2 Summary of Key Findings

The research has confirmed that agricultural LSLAs generally have a significant negative impact on local biodiversity. This conclusion aligns with hypothesis **H1**, which posits that LSLAs, regardless of their origin, detrimentally affect biodiversity. This was evident from the consistent negative correlation observed between the Biodiversity Habitat Index (BHI) and the presence of LSLAs in the regions studied.

However, the investigation into hypothesis **H2**, which proposed that transnational LSLAs would have a more pronounced negative effect on biodiversity compared to domestic LSLAs, yielded mixed results. While the DiD dummy was significant across all comprehensive models, indicating some level of difference, the direction and magnitude of this effect varied by country. In some countries, transnational investments had a more negative impact, whereas in others, the effect was less pronounced or even positive. This suggests that the impact of transnational LSLAs is highly context-dependent, influenced by country-specific factors.

5.3 Theoretical and Practical Implications

Despite not fully confirming both hypotheses across all cases, the research provides robust and intriguing findings. One of the notable discoveries is that initial biodiversity levels in areas with transnational LSLA presence are worse than in areas with domestic investments. This could be due to various institutional settings, though uncovering the exact reasons lies beyond the scope of this study. Future research should delve into these institutional contexts to provide more insights, which could significantly advance biodiversity protection efforts in such areas.

Moreover, the country-dependent nature of the DiD effect underscores the importance of considering local contexts when evaluating the impacts of LSLAs. This research highlights the need for policymakers to carefully scrutinize and regulate transnational investments in agriculture. Institutions capable of protecting biodiversity from harmful agricultural practices must be established and strengthened to mitigate the adverse effects identified.

The study also reveals the importance of including controls for geography, climate, and human activities in quantitative research on biodiversity. These variables consistently improved model fit, suggesting their critical role in understanding biodiversity changes.

5.4 Future Research Directions

Future research should focus on conducting case studies in different countries to uncover why the impacts of transnational acquisitions vary so significantly across the region. These studies should begin with qualitative approaches to identify the underlying mechanisms, which could then inform more refined quantitative analyses.

Additionally, further exploration into the institutional settings that influence the outcomes of LSLAs is necessary. Understanding how different governance structures, regulatory environments, and socio-economic contexts affect biodiversity can provide deeper insights into mitigating the negative impacts of such investments.

From a methodological perspective, future studies should aim to address the limitations identified in this research. Enhancing data availability and quality is paramount. For example, obtaining more detailed information on the size and exact locations of LSLAs would allow for more precise spatial analysis and better control for potential biases. Furthermore, advanced methodologies to control for spatial autocorrelation and temporal variations should be employed to enhance the robustness of the findings.

5.5 Policy Recommendations

The findings of this research have significant implications for policymakers in Latin America and the Caribbean. It is crucial for governments to implement policies that regulate transnational agricultural investments more stringently. This includes establishing robust environmental protection frameworks and ensuring that agricultural practices do not compromise biodiversity. Policymakers should also consider creating incentives for sustainable agricultural practices and penalizing activities that harm local ecosystems.

Moreover, international cooperation and agreements may be necessary to manage the impacts of transnational LSLAs effectively. Countries in the region could benefit from sharing best practices and collaborating on strategies to protect biodiversity while promoting sustainable development.

5.6 Limitations

While this research makes significant contributions, it is not without its limitations. The approximation of agricultural influence through LSLA areas presents several challenges, such as potential spatial biases due to the lack of precise data on LSLA locations and sizes. Additionally, the absence of a comprehensive control group limits the ability to make definitive comparisons. The presence of spatial autocorrelation and the inability to fully control for temporal factors further constrain the robustness of the findings.

Despite these limitations, the approximations and methodologies used provide valuable initial insights, highlighting the need for more detailed data and advanced analytical techniques in future research.

5.7 Conclusion

In conclusion, this thesis has provided a comprehensive analysis of the impacts of agricultural LSLAs on biodiversity in Latin America and the Caribbean. While confirming that LSLAs generally harm biodiversity, it has also uncovered the complex and country-specific nature of these effects. The findings underscore the importance of robust institutional frameworks and context-specific policies to mitigate the negative impacts of transnational agricultural investments. Future research should build on these insights, addressing the identified limitations and further exploring the complex dynamics between LSLAs and biodiversity. By doing so, it can contribute to the development of more effective strategies for protecting biodiversity in the face of expanding agricultural investments.

6

Bibliography

Angrist, J. D., Pischke, J.-S. (2008). Mostly Harmless Econometrics: An Empiricist's Companion

Baltagi, B. H. (2021a). Test of Hypotheses with Panel Data, in Econometric Analysis of Panel Data, Springer Texts in Business and Economics, pp.75–108

Baltagi, B. H. (2021b). The Two-Way Error Component Regression Model, in Econometric Analysis of Panel Data, Springer Texts in Business and Economics, pp.47–74

Bucheli, M. (2008). Multinational Corporations, Totalitarian Regimes and Economic Nationalism: United Fruit Company in Central America, 1899–1975, Business History, vol. 50, no. 4, pp.433–454

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., Liu, Y. (2022). Global 1 Km $\times$ 1 Km Gridded Revised Real Gross Domestic Product and Electricity Consumption during 1992–2019 Based on Calibrated Nighttime Light Data, Scientific Data, vol. 9, no. 1

Croissant, Y., Millo, G. (2023). Linear Models for Panel Data, R, Available Online: https://cran.r-project.org/web/packages/plm/plm.pdf [Accessed 16 May 2023]

Dasgupta, P. (2021). The Economics of Biodiversity: The Dasgupta Review: Full Report, Updated: 18 February 2021., London: HM Treasury

Davis, K. F., Müller, M. F., Rulli, M. C., Tatlhego, M., Ali, S., Baggio, J. A., Dell'Angelo, J., Jung, S., Kehoe, L., Niles, M. T., Eckert, S. (2023). Transnational Agricultural Land Acquisitions Threaten Biodiversity in the Global South, Environmental Research Letters, vol. 18, no. 2

Dogan, B, (2022). Does FDI in Agriculture Promote Food Security in Developing Countries? The Role of Land Governance, Transnational Corporations, vol. 29, no. 2, pp.49–78

EarthEnv. (2024). Global 1,5,10,100-Km Topography, Available Online: https://www.earthenv.org/topography [Accessed 7 May 2024]

Ferrier, S., Guisan, A. (2006). Spatial Modelling of Biodiversity at the Community Level, Journal of Applied Ecology, vol. 43, no. 3, pp.393–404

Greifer, N. (2023). Package 'MatchIt', R, Available Online: https://cran.r-project.org/web/packages/MatchIt/MatchIt.pdf [Accessed 16 May 2024]

Hansen, B. B. (2004). Full Matching in an Observational Study of Coaching for the SAT, Journal of the American Statistical Association, vol. 99, no. 467, pp.609–618

Hansen, B. B., Klopfer, S. O. (2006). Optimal Full Matching and Related Designs via Network Flows, Journal of Computational and Graphical Statistics, vol. 15, no. 3, pp.609–627

Harwood, T., Ferrier, S., Ware, C. (2022a). Summarising CSIRO BILBI 30s Indicators by Region

Harwood, T., Ware, C., Hoskins, A., Ferrier, S., Bush, A., Golebiewski, M., Hill, S., Ota, N., Perry, J., Purvis, A., Williams, K. (2022b). BHI v2: Biodiversity Habitat Index: 30s Global Time Series, Available Online: https://researchdata.edu.au/bhi-v2-biodiversity-time-series/1959488

Hlavac, M. (2022). Stargazer: Well-Formatted Regression and Summary Statistics Tables, R, Available Online: https://CRAN.R-project.org/package=stargazer [Accessed 18 May 2023]

Li, X., Zhou, Y., Zhao, M., Zhao, X. (2020). A Harmonized Global Nighttime Light Dataset 1992–2018, Scientific Data, vol. 7, no. 1, pp.7-168

Li, X., Zhou, Y., Zhao, M., Zhao, X. (2023). Harmonization of DMSP and VIIRS Nighttime Light Data from 1992-2021 at the Global Scale, Figshare dataset, Available Online: https://figshare.com/articles/dataset/Harmonization_of_DMSP_and_VIIRS_nighttime_light_data_from_1992-2018_at_the_global_scale/9828827/8

Liu, Q., Sutton, P. C., Elvidge, C. D. (2011). Relationships between Nighttime Imagery and Population Density for Hong Kong, Proceedings of the Asia-Pacific Advanced Network, vol. 31, pp.79-90

Mann, H., Smaller, C. (2010). Foreign Land Purchases for Agriculture: What Impact on Sustainable Development?, Sustainable Development Innovation Briefs, no. 8

Mihálka, A. (2023). The Effects of Economic Development on Biodiversity: A Comparative Analysis of Latin America and Europe Using Panel Linear Regression Models, MSc Thesis, Department of Economic History, Lund University School of Economics and Management, Available Online: http://lup.lub.lu.se/student-papers/record/9134241

Moody, C. E., Marvell, T. B. (2020). Clustering and Standard Error Bias in Fixed

Effects Panel Data Regressions, Journal of Quantitative Criminology, vol. 36, pp.347–369

Natural Earth. (2024). Free Vector and Raster Map Data at 1:10m, 1:50m, and 1:110m Scales, Available Online: https://www.naturalearthdata.com [Accessed 1 March 2024]

Ortiz-Ospina, E., Beltekian, D., Roser, M. (2024). Trade and Globalization. Our World in Data, Available Online: https://ourworldindata.org/trade-and-globalization [Accessed 26 May 2024]

Österblom, H., Bebbington, J., Blasiak, R., Sobkowiak, M., Folke, C. (2022). Transnational Corporations, Biosphere Stewardship, and Sustainable Futures, Annual Review of Environment and Resources, vol. 47, no. 1, pp.609–635

Property Rights Alliance. (2024). International Property Rights Index, Available Online: https://internationalpropertyrightsindex.org [Accessed 7 May 2024]

Pustejovsky, J. (2023). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections, R, Available Online: https://cran.r-project.org/package=clubSandwich [Accessed 16 May 2024]

RDocumentation. (2024). Lm: Fitting Linear Models, Available Online: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/lm [Accessed 20 May 2024]

Ritchie, H., Roser, M. (2024). Land Use. Our World in Data, Available Online: https://ourworldindata.org/land-use [Accessed 26 May 2024]

Ritchie, H., Spooner, F., Roser, M. (2024). Biodiversity. Our World in Data, Available Online: https://ourworldindata.org/biodiversity [Accessed 26 May 2024]

Sändig, J. (2021). Contesting Large-Scale Land Acquisitions in the Global South, World Development, vol. 146

Santangelo, G. D. (2018). The Impact of FDI in Land in Agriculture in Developing Countries on Host Country Food Security, Journal of World Business, vol. 53, no. 1, pp.75–84

Schneider, J. M., Zabel, F., Mauser, W. (2022). Global Inventory of Suitable, Cultivable and Available Cropland under Different Scenarios and Policies, Scientific Data, vol. 9, no. 1

Sutton, P. (1997). A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States, Photogrammetric Engineering & Remote Sensing, vol. 63, no. 11, pp.1303–1313

The Land Matrix. (2024a). Deals, Available Online: https://landmatrix.org/list/deals/ [Accessed 1 March 2024]

The Land Matrix. (2024b). Map, Available Online: https://landmatrix.org/map/ [Accessed 1 March 2024]

WorldClim. (2022). Historical Monthly Weather Data, Available Online: https://www.worldclim.org/data/monthlywth.html [Accessed 15 May 2024]

Zoological Society of London, WWF. (2023). Living Planet Index, Available Online: https://www.livingplanetindex.org [Accessed 26 May 2024]

Appendix A

Figures and Tables

Table A.1: Summary Table of Different Models

	Dependent variable:			
	log(BHI)			
	OLS		*FE*	*Matched FE*
	(1.6)	(2.3)	(3.6)	(4.4)
Transnational Dummy	-0.04668^{***}	-0.06239^{**}	*(omitted)*	*(omitted)*
	(0.01605)	(0.02756)		
Time Dummy	0.01313	-0.04806	0.15684^{***}	0.00619
	(0.03964)	(0.07309)	(0.02217)	(0.00490)
DiD Dummy	-0.09457^{***}	-0.10737^{***}	0.03267^{***}	0.01711^{**}
	(0.02350)	(0.02783)	(0.00620)	(0.00669)
Constant	0.26978	0.05971	-	-
	(0.26117)	(0.40175)		
Year of implementation between 2000 and:	2020	2010	2010	2010
Observations	1,598	850	846	846
R^2	0.70662	0.70713	0.90022	0.87120
Adjusted R^2	0.69792	0.69150	0.79131	0.73193

Note: Clustered robust standard errors are in parentheses
All models contain geography and climate control variables
Models 1.6, 2.3 and 3.6 contain human activity control variables
*p<0.1; **p<0.05; ***p<0.01

Table A.2: Country Interaction Coefficients

	Dependent variable:			
	log(BHI)			
	OLS		*FE*	*Matched FE*
	(1.6)	(2.3)	(3.6)	(4.4)
DiD Dummy	-0.09457^{***}	-0.10737^{***}	0.03267^{***}	0.01711^{**}
	(0.02350)	(0.02783)	(0.00620)	(0.00669)
Interacted with:				
Bolivia	-0.02459	0.04427	-0.05940^{***}	-0.07933^{***}
	(0.06624)	(0.06963)	(0.00655)	(0.00665)
Brazil	0.17570^{***}	0.19319^{***}	-0.01984^{**}	0.01525^{*}
	(0.03148)	(0.03937)	(0.00926)	(0.00808)
Chile	0.13230^{**}	0.18213^{**}	-0.00732	0.02558^{***}
	(0.05429)	(0.07838)	(0.00841)	(0.00733)
Colombia	0.15996^{***}	0.25381^{***}	-0.06375^{***}	-0.06487^{***}
	(0.03971)	(0.07062)	(0.00839)	(0.00886)
Costa Rica	0.11690	0.21300	-0.07054^{***}	-0.04103^{***}
	(0.09363)	(0.13499)	(0.01421)	(0.01548)
Guatemala	0.10877	0.14103^{*}	-0.02167	-0.05895^{***}
	(0.06097)	(0.07240)	(0.02192)	(0.02068)
Honduras	0.04911	0.08829^{***}	-0.02892^{***}	-0.03382^{***}
	(0.02253)	(0.02689)	(0.00636)	(0.00650)
Mexico	0.22502^{***}	0.47914	-0.13233^{***}	-0.04481^{***}
	(0.11690)	(0.32717)	(0.01310)	(0.00766)
Nicaragua	0.16868^{**}	0.20197^{***}	-0.04153^{***}	-0.04266^{***}
	(0.05564)	(0.07711)	(0.00934)	(0.01098)
Panama	0.07231^{***}	0.10768^{***}	-0.08947^{***}	-0.03149^{**}
	(0.03158)	(0.04012)	(0.01278)	(0.01234)
Paraguay	-0.00128	0.05145	-0.00373	-0.04878^{***}
	(0.05164)	(0.05399)	(0.01249)	(0.01027)
Peru	0.24231	0.36375^{***}	-0.03057^{***}	-0.04446^{***}
	(0.07256)	(0.08976)	(0.00766)	(0.00975)
Uruguay	0.17463	0.22345^{***}	0.01893^{***}	0.00449
	(0.02773)	(0.02992)	(0.00708)	(0.00742)
Venezuela	0.03688			
	(0.05212)			
Observations	1,598	850	846	846
R^2	0.70662	0.70713	0.90022	0.87120
Adjusted R^2	0.69792	0.69150	0.79131	0.73193

Note: Clustered robust standard errors are in parentheses.
Sample is set to year of impl. between 2000 and 2010 for models 2.4, 3.6 and 4.4
Interaction variables with Dominican Republic, Ecuador, and Jamaica
have been omitted in all cases
*p<0.1; **p<0.05; ***p<0.01

Table A.3: Random Effects Models

	Dependent variable:					
	log(BHI)					
	(5.1)	(5.2)	(5.3)	(5.4)	(5.5)	(5.6)
dummy_tn	−0.26454***	−0.14371***	−0.18431***	−0.18726***	−0.18970***	−0.08203***
	(0.02749)	(0.03317)	(0.03516)	(0.03656)	(0.03624)	(0.02922)
dummy_t4	−0.09530***	−0.09530***	−0.03846***	0.05415***	0.05751***	0.13162***
	(0.00366)	(0.00366)	(0.00394)	(0.01524)	(0.01519)	(0.02343)
dummy_did_4	0.04188***	0.04188***	0.03098***	0.02705***	0.02693***	0.02701***
	(0.00478)	(0.00478)	(0.00381)	(0.00358)	(0.00362)	(0.00642)
Constant	−0.65620***	−1.14980***	0.60514***	1.08930***	1.10102***	2.12096***
	(0.01601)	(0.19693)	(0.21435)	(0.22369)	(0.21766)	(0.28712)
Observations	844	844	844	844	844	844
R^2	0.57584	0.60022	0.70510	0.73713	0.73776	0.80988
Adjusted R^2	0.57433	0.59494	0.70049	0.73237	0.73301	0.80016
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.4: Matched Random Effects Models

	Dependent variable:			
	log(BHI)			
	(6.1)	(6.2)	(6.3)	(6.4)
dummy_tn	−0.21100***	−0.12093***	−0.14995***	−0.15053***
	(0.03019)	(0.03044)	(0.03235)	(0.03202)
dummy_t4	−0.08136***	−0.08110***	−0.03031***	−0.02652***
	(0.00432)	(0.00437)	(0.00426)	(0.00458)
dummy_did_4	0.02794***	0.02768***	0.01858***	0.01000
	(0.00530)	(0.00535)	(0.00414)	(0.00705)
Constant	−0.70974***	−1.29976***	0.44995***	0.32184*
	(0.02031)	(0.14044)	(0.16688)	(0.18013)
Observations	844	844	844	844
R^2	0.57052	0.58163	0.68794	0.71122
Adjusted R^2	0.56899	0.57660	0.68343	0.70239
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.5: LM and Hausman tests

		(3.6) & (5.6)	(4.4) & (6.4)
LM	*chi-sq*	327.31	268.05
	p-value	2.2×10^{-16}	2.2×10^{-16}
Hausman	*chi-sq*	849.69	242.36
	p-value	2.2×10^{-16}	2.2×10^{-16}

Appendix B

General Regression Formula

$$log(BHI_{it}) = \beta_0 + \beta_1 D_{Tn,it} + \beta_2 D_{Time,it} + \beta_3 D_{DiD,it} + \beta_4 (D_{DiD,it} \times \text{Country}_{it})$$

$$+\beta_5 GEO_{it} + \beta_6 CLIM_{it} + \beta_7 HUM_{it} + \eta_i + \lambda_t + \epsilon_{it}$$

where:

- BHI_{it} represents the BHI at location i and time t,
- $D_{Tn,it}$ is a dummy variable indicating the presence of a transnational LSLA,
- $D_{Time,it}$ is a time dummy variable,
- $D_{DiD,it}$ is the interaction term for the DiD analysis,
- Country_{it} is the categorical country variable,
- GEO_{it} includes geography control variables[1],
- $CLIM_{it}$ includes climate control variables[2],
- HUM_{it} includes human activity control variables [3],
- η_i represents location fixed effects,
- λ_t represents time fixed effects, and
- ϵ_{it} is the error term.

[1]The geography control variables are log(location size), latitude, longitude, log(elevation), log(slope), and PCL

[2]The climate control variables are average maximum temperature and precipitation

[3]The human activity control variables are IPRI and either log(GDP) or log(NTL)