Essays in Urban Economics and International Trade

by

Michelle S. Lam

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2023

Doctoral Committee:

Professor Andrei Levchenko, Co-Chair
Associate Professor Sebastian Sotelo, Co-Chair
Professor John Bound
Professor Jagadeesh Sivadasan

Michelle S. Lam
msll@umich.edu
ORCID iD: 0000-0003-1094-5246

To anyone who ever identified with Sisyphus.

ACKNOWLEDGEMENTS

There have been many people who have contributed to my journey to complete a Ph.D., and I am very thankful for all the help and support that I have received.

I would like to thank my committee members, Andrei Levchenko, Sebastian Sotelo, Jagadeesh Sivadasan, and John Bound for their patience and guidance. I would like to thank other University of Michigan faculty members who have also generously given me guidance, including but not exhaustive: Dominick Bartelme, Hoyt Bleakley, and Charlie Brown. I would also like thank the Michigan economics department staff, the urban department at the Federal Reserve of Philadelphia, Brad Foster, Clint Carter, Michael Frachetti, and Morgan Kelly. I would also like to extend my sincere appreciation for my co-authors, Bunyada (Mos) Laoprapassorn and Dyanne Vaught.

I would like to give a special thank you to Wellesley College; the economics department, Akila Weerapana, Daniel Sichel; the music department, Jenny Tang, Margaret Angelini. Even when I was endlessly uncertain, they were endlessly encouraging.

I would like to thank Tiffany Ng, James Kibbie, and the School of Music, Theatre & Dance for being my second home at the University of Michigan. In the face of adversity, it is always to music I return.

I would like to thank my friends and family for their support. To Yuen Lam, Ting Lam, David Lam, and Patrick Wu: thank you.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES vii

LIST OF TABLES viii

LIST OF APPENDICES x

ABSTRACT xi

CHAPTER

I. Implications of Zoning for Urban Gentrification in New York City 1

1.1 Introduction 1
1.2 Background and Data 7
1.2.1 Background 7
1.2.2 Data 9
1.3 Empirical Analysis 11
1.3.1 Summary Statistics 11
1.3.2 Reduced Form 18
1.4 Spatial Equilibrium Model of a City 23
1.4.1 Setup 23
1.4.2 Individuals 23
1.4.3 Firms 27
1.4.4 Developers 28
1.4.5 Spillovers 29
1.4.6 Equilibrium 30
1.4.7 Parameters 30
1.4.8 Model Calibration 35
1.5 Counterfactuals and Welfare Decomposition 38
1.5.1 Counterfactuals: Bloomberg Rezoning 38

1.5.2 Welfare Decomposition . . . 44
1.6 Conclusion . . . 45

II. The Impact of Trade on Development: Evidence from Pastoralist Practices on the Ancient Silk Road . . . 46

2.1 Introduction . . . 46
2.2 Background . . . 50
2.2.1 Definitions . . . 52
2.2.2 History . . . 54
2.3 Data . . . 55
2.3.1 Data Sources . . . 56
2.3.2 Data Construction . . . 60
2.4 Specification and Results . . . 62
2.4.1 Standard Errors . . . 64
2.4.2 OLS Approach . . . 65
2.4.3 Instrumental Variable Approach . . . 70
2.5 Mechanisms . . . 82
2.6 Conclusion . . . 85

III. Corporate Influence in Trade Agreements: Approaches Using Machine Learning . . . 87

3.1 Introduction . . . 87
3.2 Data . . . 90
3.2.1 Texts of Trade Agreements . . . 90
3.2.2 Supplement to TOTA . . . 90
3.2.3 Trade Flows . . . 91
3.2.4 Design of Trade Agreements . . . 91
3.2.5 Corporate 10-K Filings . . . 91
3.2.6 Loughran-McDonald Sentiment Word Lists . . . 92
3.3 Models . . . 92
3.3.1 DESTA . . . 93
3.3.2 Sentiment Analysis . . . 96
3.3.3 Topic Modeling . . . 100
3.3.4 doc2vec . . . 107
3.3.5 Transformers . . . 110
3.4 Comparing Methods . . . 114
3.5 Conclusion . . . 116

APPENDICES . . . 118
A.1 Tables . . . 118
A.2 Figures . . . 122
B.1 Data . . . 124
B.1.1 Constructing Flow Accumulation . . . 124

B.2 Robustness Checks 125
C.1 DESTA 130
C.1.1 Variables Included in the Corporate Influence Index 130
C.2 Topic Modeling 132
C.2.1 LDA Topic Words 132
C.2.2 STM Topic Words 133
C.3 Sentiment Analysis 135
C.4 doc2vec 138

BIBLIOGRAPHY 139

LIST OF FIGURES

Figure

1.1	Upzoned and Downzoned Tax Lots in NYC	12
1.2	East Harlem	13
1.3	Changes in New York City from 2002 to 2020	14
1.4	Amenities	36
1.5	Prices	37
1.6	Counterfactual 1	41
1.7	Counterfactual 1: Measures of Gentrification	42
1.8	Counterfactual 2: Measures of Gentrification	43
2.1	Silk Road Routes	51
2.2	Silk Road Sites in Inner Asia Mountain Corridor	57
2.3	Shortest Distance to Silk Road site	61
2.4	Variogram	66
2.5	Simulated Herding Path	73
2.6	Simulated Herding Path and Silk Road Sites	74
2.7	Shortest Distance to Random Points	79
3.1	Corporate Influence Index (DESTA)	94
3.2	Trade Agreements: Sentiment Word Lists Mean Proportions Over Time	97
3.3	10-Ks: Sentiment Word Lists Mean Proportions Over Time	98
3.4	LDA Coherence Scores	103
3.5	LDA Topics' Mean Proportions Over Time	104
3.6	STM Topics' Mean Proportions Over Time	105
3.7	doc2vec Similarity Cosine Scores Over Time	109
3.8	BERT Similarity Cosine Scores Over Time	112
A.1	Zoning Categories in 1961	122
A.2	Counterfactual 2	123
B.1	No. of Silk Road cells within 100km	126
C.1	Components of the Corporate Influence Index	131
C.2	Trade Agreements: Sentiment Word Lists Mean Proportions Over Time	135
C.3	10-Ks: Sentiment Word Lists Mean Proportions Over Time	136

LIST OF TABLES

Table

1.1	Summary Statistics by Zoning Change	16
1.2	Changes in Residential Area by Zoning Change	16
1.3	Quartiles of Price Per Square Foot in 2020	17
1.4	Quartiles of Δ Log(Price Per Square Foot) from 2002 to 2020	18
1.5	Dependent Variable: Change in ln(Price per Square Foot)	22
1.6	Estimating θ_g	33
1.7	Calculating θ_g	33
1.8	Estimating ρ_g, η_g	34
1.9	Counterfactual Results	38
1.10	Welfare Decomposition Example	44
2.1	Count of Silk Road Sites by Country and Node Types	57
2.2	Summary Statistics of Development Measures	63
2.3	OLS Results	67
2.4	OLS with Alternative Measures of Development	69
2.5	IV Results	77
2.6	Robustness Checks	80
2.7	Mechanisms	84
3.1	Gravity Equations with DESTA Country Influence Index	95
3.2	Gravity Equations with Loughran-McDonald Sentiment Word Lists	99
3.3	Gravity Equations with Loughran-McDonald Sentiment Word Lists (Manufacturing)	100
3.4	Gravity Equations with LDA Topics	106
3.5	Gravity Equations with STM Topics	107
3.6	Gravity Equations with Doc2vec Similarity Cosine Scores	110
3.7	Gravity Equations with BERT Similarity Cosine Scores	113
A.1	LODES in New York City	118
A.2	NYC PLUTO Summary Statistics	119
A.3	Dependent Variable: ln(Price per Square Foot)	120
A.4	Dependent Variable: ln(Price per Square Foot)	120
A.5	Parameter Summary	121
B.1	Robustness Checks	127
B.2	IV Results Using Different Cutoffs to Construct an Instrument	129

C.1 Correlations Between Sentiment Word List and LDA Topic Proportions 137
C.2 Correlations Between Sentiment Word List and STM Topic Proportions 137
C.3 Gravity Equations with Doc2vec Similarity Cosine Scores 138

LIST OF APPENDICES

Appendix

A. Appendix for Chapter I 118

B. Appendix for Chapter II 124

C. Appendix for Chapter III 130

ABSTRACT

This dissertation contains three independent essays in urban economics and international trade, with the aim of understanding geography's impact on spatial location and trade.

The first chapter estimates the effect of the sweeping rezoning program initiated by the Bloomberg administration in New York City from 2002 to 2013. Using an instrumental variable strategy incorporating regime change from the 1961 Zoning Ordinance, this paper finds a 10% increase in the floor area ratio (a physical change), defined as the building floor area to the lot size, decreases price per square foot by 1.7% and a 10% increase in the maximum floor area ratio (a policy change) results in a 2.4% decrease in price. The paper uses a spatial general equilibrium model to quantify the aggregate impact of the rezoning policy on low- and high-skilled workers. Counterfactual analysis finds that as a result of the Bloomberg rezoning, GDP increases and price per square foot for residential floorspace drops, contrasting with an increase in residential segregation and increases in welfare only for high-skilled workers.

The second chapter, co-authored with Bunyada (Mos) Laoprapassorn, studies the long-term effect of trade on development. We approach this question in the context of the ancient Silk Road, examining whether the locations along the highland Silk Road continue to be relatively more developed than other locations in the highland region along the Inner Asia Mountain Corridor that were not on the ancient Silk Road. We proxy for modern development using high-resolution satellite imagery. To

provide a causal effect between proximity to the Silk Road and modern development, we adopt a novel instrumental variable, using a simulated seasonal mobility pattern of the nomadic pastoralists from Frachetti et al. (2017) as an instrument for the locations of the Silk Road sites. We find a significant and robust positive relationship between proximity to Silk Road sites and modern development measures; an increase in the distance to the Silk Road by one standard deviation decreases the night lights intensity by 10.0%. Based on the elasticity of night lights with respect to GDP in the literature, this corresponds to a decrease in GDP of about 4.1%-9.7%.

The third chapter, co-authored with Dyanne Vaught, studies the evolution of corporate influence over time on trade agreements and its impact on trade flows. We approach this question using a variety of techniques, from traditional indicators to machine learning methods, and find that corporate influence has been increasing over time in trade agreements. A 10% increase in the proportion of Services and Investment topic in the trade agreements results in 0.19% to 0.92% increase in trade flows. Next, this paper finds that to multinational corporate language has a differential effect on trade flows. Using deep learning methods to score similarity, a 10% increase in similarity to multinational corporation language has a negative effect on trade flows, ranging from -0.3% to -1.2%.

Chapter I

Implications of Zoning for Urban Gentrification in New York City

1.1 Introduction

Over the past few decades, the debate over gentrification and the role of zoning regimes and changes have captured worldwide public attention. In the United States, each locality has significant discretion to make its own zoning policy, resulting in significant differences in how decisions are made within the community and how the land is subsequently developed. New York City represents an interesting case due to its status as the largest city in the United States and the diversity of neighborhood typologies within the five boroughs, ranging from tall skyscrapers, to middle height mixed-used districts, and low-height residential areas.

New York City has had just two main zoning ordinances: 1916 and 1961. Current zoning policy still draws on the 1961 ordinance, which changed the main tool of zoning to be the maximum floor area ratio (MFAR). The FAR is defined as the ratio of the building floor area to the lot area. In this paper, I differentiate between the policy, MFAR, and the physical reality, FAR; changes in MFAR may not be reflected in FAR. Mayor Michael Bloomberg took office after a period of decline in the 1980s and the destabilizing events of 9/11 (Stein, 1981; Von Hoffman, 1990). As

a businessman himself, his administration's goals were to grow the city and increase its competitiveness to attract firms and workers from all over the world. To do this, his administration set about a massive rezoning of almost 40% of land in New York City from 2002–2013 with the stated aim to increase housing supply (City, 2007).

In this paper, I develop a framework to study the effect of this massive rezoning effort. I assemble a detailed geographic dataset of Decennial Census, American Community survey, and New York City datasets both from public and confidential sources.[1] Depending on the context and the underlying data source used, the geographic unit will either be New York City tax lot, Census block group, or Census tract. I first use the data to present basic statistics on changes in price per square foot and quantities to help motivate the paper and provide direction for the model. I find that upzoned areas, i.e. those that have had an increase in MFAR, had both higher levels of prices in 2020, as well as greater increase in prices between 2002 and 2020. Consistent with this finding, the quartiles of city block groups that had the highest 2020 price levels and price changes also had higher percentage of upzoned block groups and higher median income of households. These results suggest that certain neighborhoods were targeted due to their access and proximity to productive areas, which is supported by contemporaneous accounts observing that upzoned areas attracted higher-income workers and spurred gentrification. I also find, somewhat surprisingly, that the percentage of lots that had an increase in residential area is not materially different between downzoned and upzoned lots, and that in upzoned lots, the actual FAR, on average, is only about 47% of the permitted MFAR. These results suggest that an evaluation of the effects of the zoning policy requires strategies to deal with endogeneity of how zoning changes are targeted, and must account for

[1]Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2212. (CBDRB-FY23-P2212-R10103)

mechanisms that discourage development from occurring to the maximum possible threshold.

Next, I present more rigorous empirical analysis that adds controls for confounding factors, and adopts an instrumental variable strategy to address concerns from the apparent targeting of zoning changes to evaluate the impact of changes in MFAR and FAR on housing prices. Echoing the observations from the previous paragraph, the results from the OLS without controls indicate a similar story in that an increase in the FAR is associated with an increase in prices. However, once controls are added and an instrumental variable strategy is implemented, the signs on the two main variables are flipped, indicating that more permissive zoning is associated with price decreases over time. In the preferred 2SLS specification, a 10% increase in the FAR results in a 1.7% decrease in price, and a 10% increase in the MFAR results in a 2.4% decrease in price. The instrumental variable strategy relies on the change in the zoning ordinance in 1961, where the New York City government shifted from using bulk regulations to floor area ratio as the primary instrument of zoning. Due to the inconsistency in converting bulk regulations to floor area ratio, there exists random variation in how floor area ratio and zoning categorization was determined in 1961. Moreover, the likelihood of an area being rezoned depends on the level of the 1961 zoning and the subsequent changes that occurred between 1961 and 2002. The more changes that occurred between 1961 and 2002, the less likely MFAR is to be changed during the Bloomberg administration and more likely for the building to reach its MFAR during the 2002 to 2020 period. These two instruments, the level of MFAR in 1961 and the change between MFAR in 1961 and the FAR in 2002, are used to instrument for the policy change and the actual physical change, assuming that developers and policymakers in 1961 through 2002 did not anticipate future zoning changes that the Bloomberg administration eventually carried out.

I follow with a quantitative spatial equilibrium model with a commuting gravity

equation. The model is a closed city in which the population may only move within the city and the population numbers are fixed. There are two groups, high-productivity and low-productivity residents that are differentiated by characteristics such as their commuting preferences and taste for density. Each resident is a renter, and shares the profits from the land equally. There are firms in each geographic unit which use labor as an input. Next, there are developers that build living space taking as given the lot size, MFAR, and the convex construction costs based on the FAR and the MFAR. The convex building costs in particular help to disincentivize developers from building to the maximum allowed floorspace. Last, there are spillovers on amenities, with a location becoming more attractive based on the density, i.e. number of residents divided by the floor area. I then calibrate or structurally estimate parameters in order to be able to calculate an equilibrium from observable data such as number of residents and workers for each geographic unit and back out unobservable data such as amenity values and productivity values for each geographic unit.

With the parameters of my structural model in hand, I perform two counterfactuals to analyze effects of Bloomberg era zoning changes, under alternative assumptions about whether or not there is an increase in high-productivity residents. First, I examine the implied changes in rents and location choices of different resident types in the model from the observed change MFAR, while holding the city employment fixed, i.e., not allowing for gentrification induced by movement of high-productivity residents into the city. In a second counterfactual, I allow for a 10% increase in high-productivity residents along with the proposed change in MFAR. This scenario mirrors the growth of the NYC population from 8 million to 8.8 million in the period spanning 2000 to 2020. The results from both counterfactuals are qualitatively similar: GDP increases, rent decreases, inequality increases, and segregation increases.[2]

[2]Residential segregation is measured by the dissimilarity index. It is calculated using the following formula: $\frac{\sum_{i=1}^{n} t_i|p_i - P|}{2TP(1-P)}$. n is the number of areas, t_i us the total population of area i, p_i is ratio of the minority population of area i to t_i, and P is the ratio of the overall minority population to the overall area population. Dissimilarity then measures the percentage of a group's population that

High-productivity residents are able to move to places that both have higher group-specific amenity values and are closer in proximity to higher-paying jobs, resulting in an increase in their welfare. On the other hand, low-productivity residents face the opposite situation and move further away from productive locations and relocate to less dense locations, resulting in a decrease in their welfare. The overall effectiveness of the rezoning then depends on the social planner's outlook on whether an increase in GDP and decrease in rent at the expense of increasing residential isolation and inequality is acceptable.

This paper contributes to several strands of literature. First, this paper adds to the burgeoning body of literature that uses quantitative spatial equilibrium models to explore city development. Ahlfeldt et al. (2015) and Tsivanidis (2019) develop spatial equilibrium models using fine spatial detail to study the impact of city policy changes (division and reunification of Berlin and improvement of public transit in Colombia, respectively) on the spatial organization and welfare of its residents. Papers such as Couture et al. (2019) and Hoelzlein (2020) further explore the implications of rising incomes and prices to the spatial organization of cities and present measures to assess whether a neighborhood has gentrified. To add to this literature, I study the impact of zoning policy changes, including a developer problem with convex housing costs and allow for two groups of differently skilled workers. Within this strand of literature, Acosta (2020) also studies the impact of zoning changes, focusing on Chicago and increasing the share of residential zoning, while I study New York City whose rezoning processes are significantly different, and my analysis focuses on changing the MFAR instead.

Next, this paper contributes to the literature on zoning policy's effect on housing supply, beginning with Arnott and MacKinnon (1977)'s theoretical model. More

would have to change residence for each neighborhood to have the same distribution of the group as the overall area; 0 represents complete integration, 1 represents complete segregation. The minority group is considered to be the high-skilled workers.

contemporary studies include Bertaud and Brueckner (2005), Gyourko and Molloy (2015), Bunten (2017), Freemark (2020), and Tan et al. (2020). The causal inference strategy is usually dependent on city studied and geographic unit. The general results from such papers finds that the more restrictive the zoning policy, the more expensive housing becomes. My contribution to this literature is the addition of an instrumental variable strategy specific to New York City based on its 1961 regime change in zoning ordinance.

Last, this paper contributes to the literature on the spatial organization of New York City. Glaeser et al. (2005) was one of the first to argue that land use restrictions in New York City were the primary driving factor in limiting the supply of housing, and subsequently, the main culprit of high prices observed. Been et al. (2009) studies underdeveloped lots and find that most are low-density single-family housing, suggesting that regulatory constraints are too restrictive for the city's needs. Wu (2019) examines the impact of the 1961 ordinance, concluding that the zoning changes resulted in residential segregation of white and black residents, something that is echoed in my paper's finding that the Bloomberg era zoning changes resulted in more residential segregation of high and low-skilled residents. Davis (2021) evaluates the recent rezonings of the Bloomberg administration from 2000 to 2010, finding that the upzonings in particular increased white resident percentage. To this literature, I add a new panel dataset derived from confidential and publicly available detail at a fine geospatial scale, along with a framework to explore the impact of these zoning changes directly on price.

The rest of the paper is organized as follows. Section 1.2 provides background information and introduces the datasets used. Section 1.3 provides motivations, summary statistics, and a reduced form analysis. Section 1.4 outlines the spatial equilibrium model used. Section 1.5 discusses counterfactual results. Finally, Section 1.6 concludes.

1.2 Background and Data

1.2.1 Background

Zoning in New York City Prior to 2002 Land use regulation in New York City started with the Commissioners' Plan of 1811 which laid out the North-South street grid and established parcel sizes. This was followed by the 1916 ordinance which restricted height, setbacks, and usage for every lot, and was the first comprehensive land use regulation of its kind in the United States. According to this ordinance, developers could typically build out the entire lot size, but after a certain height, had to setback a certain amount, resulting in many "wedding cake" style buildings. While the 1916 ordinance was more motivated by a desire to regulate commercial interests, the following major revision in 1961 was motivated by interests to improve residential life, influenced by both the Garden City and Modernist urban planning movements. Replacing the previous three categories of land use (residential, commercial, unrestricted), the new ordinance now had more than forty categories. Most importantly, the 1961 zoning ordinance removed height and setbacks as the main tool of controlling density, instead replacing it with the floor area ratio (FAR), the building's floor area divided by the lot area. The wedding cake style buildings were now increasingly replaced by setback slender towers with the free lot space transformed to community green space. This change in policy and the resulting changes in floor area ratios informs the main instrumental variable strategy.

The 1961 resolution has not been replaced, though the city and urban planning attitudes have changed drastically since then. Both the 1916 and 1961 resolutions both mostly accepted the status quo of the land use distribution and existing densities, and were not oriented towards accommodating future growth and changing development requirements. Residents in the peripheral low-lying neighborhoods complained the ordinance allowed for too much density and thus, when new buildings were built ac-

cording to the bulk regulations, ruined neighborhood character. The 1961 zoning code also protected valuable, centrally located commercial properties from redevelopment.

Bloomberg Administration Rezoning Following a booming real estate market in the 1990s, Michael Bloomberg became mayor of New York City from 2002 to 2013. However, by 2002, most empty lots had been built out, but the city was still continuing to grow in population. To accommodate Bloomberg's vision of "enhancing competitiveness" and using zoning to "jump-start and complement private sector investment," the Department of City Planning (DCP) initiated and implemented over 120 rezoning plans, resulting in over 40% of the city being rezoned. The city-planner approach was in stark contrast to previous mayoral administrations, in which developers would usually be the ones initiating rezoning requests. Generally, the Bloomberg rezonings fell into three categories: (1) increasing the residential FAR, (2) increasing the residential and commercial FAR in high-demand corridors while restricting development in mid-blocks, and (3) decreasing the residential FAR in historically important residential neighborhoods.

In the current era, changes to the zoning code are ultimately decided by the mayoral administration, with most projects submitted through the Uniform Land Use Review Procedure (ULURP). The ULURP is a seven-month review process that first gathers feedback and nonbinding recommendations from the neighborhood community board and borough president. The mayor, the City Planning Commission–with a majority of members appointed by the mayor, and the City Council (depending on the context) then vote on the proposal. Using this process, more than 90% of the Bloomberg proposed rezoning plans were passed.

Since Bloomberg's administration, reactions to the impact of its comprehensive rezoning have been mixed, with most media focusing on the negative residential outcomes with headlines such as "Despite Much Rezoning, Scant Change in Residential Capacity," "As Bloomberg's New York Prospered, Inequality Flourished Too," "A

Luxury Apartment Rises in a Poor Neighborhood. What Happens Next?," (*New York Times*) and "Vanishing New York's Death by Gentrification" (*Bloomberg*). As such, this paper seeks to understand how the rezonings affected residential floorspace, the subsequent impact on the spatial distribution of residents, and whether the spatial reorganization contributed to inequality in the city.

1.2.2 Data

The three primary geographic units used in this paper are New York City tax lots, Census block groups, and Census tracts. Defined by New York City, a tax lot is a parcel of land identified with a unique borough, block and lot number for property tax purposes. Generally, a tax lot may be thought of as a building. New York City has five boroughs or counties: Manhattan, Queens, Brooklyn, Bronx, and Staten Island, which are further broken into blocks, defined as portion of land bounded on all sides by streets or a combination of streets, public parks, waterfront, and other boundaries. A New York City block does not necessarily correspond to the block as defined by the Census Bureau. The block group used in this paper is, however, defined by the Census Bureau, and are statistical divisions of Census tracts that contain between 600 and 3,000 people. Finally, the Census tract is a subdivision of a county that is made up of at least one block group, and usually has a population between 1,2000 to 8,000 people.

NYC PLUTO and MapPLUTO are the main sources of New York City zoning data, and they contain extensive land use data at the tax lot level. PLUTO and MapPLUTO both start in 2002 and are released annually. Along with important zoning data such as land use classification, maximum floor area ratio, floor area, and lot area, the dataset also includes assessed land value and assessed total value which are calculated by the Department of Finance by multiplying the tax lot's full market value as of January 5th by a uniform percentage for the property's tax class. This is corrected

for in the following analyses where price per square foot is concerned. PLUTO also provides mapping to Census block designations, which allows for matching to Census data products.

New York City also provides other datasets that were used in this project. First, the Department of Consumer Affairs' Legally Operating Businesses data tracks the businesses and individuals that hold a license to operate in the city. While this dataset excludes certain businesses, such as restaurants, and only one license may be granted to a business with many locations, it still provides a broad overview of the density of services and amenities in the city. Second, the NYC Parks and Recreation staff conducted Tree Censuses in 1995 and 2015, which are also used to proxy for the amenity density in the city. The New York City Department of City Planning, Technical Review Division also provided a digitized map of the changed zoning regulations in 1961, which was critical to implement the instrumental variable strategy. By aggregating data from NYC government, the NYU Furman Center provides data on housing subsidies and tax breaks at the tax lot level through its Subsidized Housing Database.

The population and demographic data come from a mixture of public and confidential Census data.[3] The 1990 and 2000 data come from the Decennial Census, and what is termed the 2020 data in this analysis is from the 2016–2019 American Community Surveys (ACS). The actual 2020 ACS data collection was disrupted by the pandemic and subsequently has not been used. The Master Address File Extract (MAFX) was then used to match Census datasets across changing geographies over time.

Depending on the analysis context, the commuting data comes from the con-

[3]Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2212. (CBDRB-FY23-P2212-R10103)

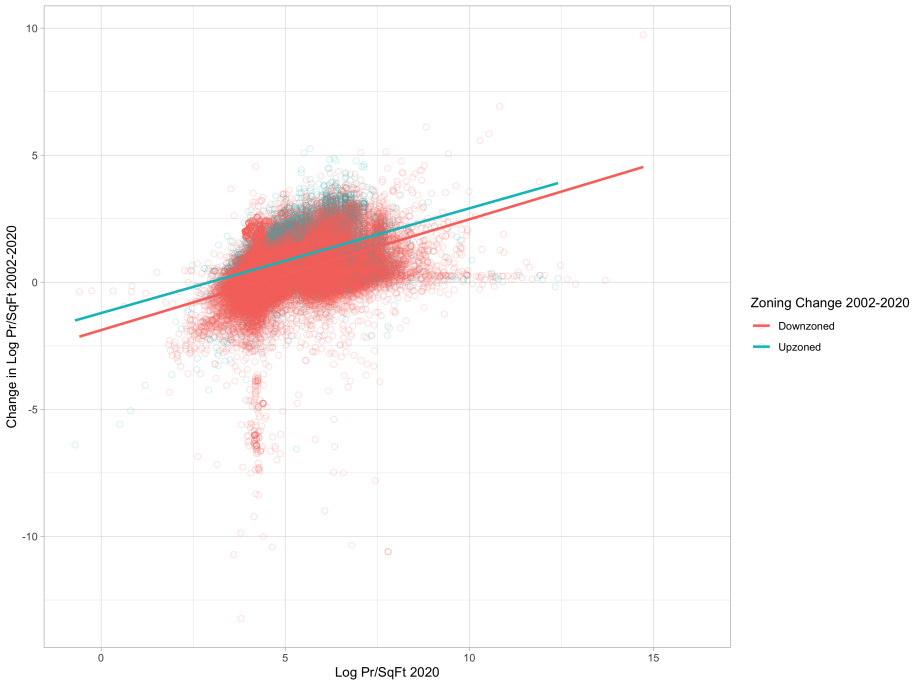
fidential 2000 Decennial Census and the publicly available Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES). Both datasets are available at the block level. Census data relies on self-reporting of residence and workplace, and as such, has issues with missing and incomplete data. The LODES data relies on administrative data and thus has less issues with incomplete data but less flexibility on how to categorize groups of residents. LODES data is released annually, starting from 2002 to 2019. To calculate the commuting distance and time by driving and walking, I use OpenStreetMap data with the Open Source Routing Machine (OSRM) algorithm. The summary statistics for LODES are located in Table A.1.

1.3 Empirical Analysis

1.3.1 Summary Statistics

In this section, I present two observations from the data to motivate the paper and my choice of model. I show that (1) upzoned areas are higher in both 2020 price levels and change in price levels from 2002 to 2020 compared to downzoned areas and that (2) areas that are higher in price levels and change in price levels exhibit characteristics that observers may associate with gentrification.

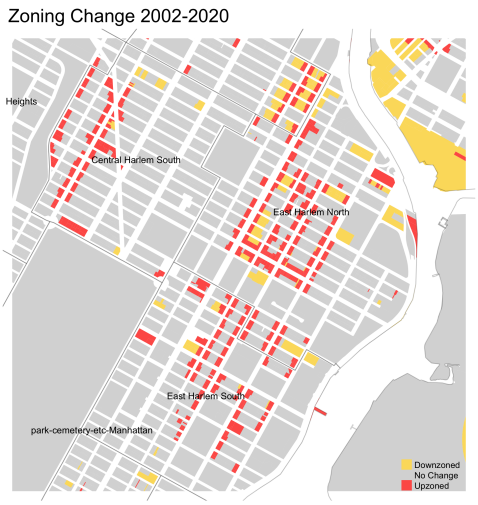
Figure 1.1: Upzoned and Downzoned Tax Lots in NYC



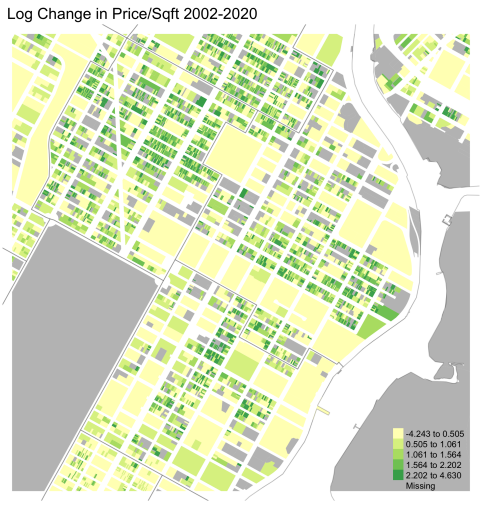
Notes: Data from NYC PLUTO.

Figure 1.2: East Harlem

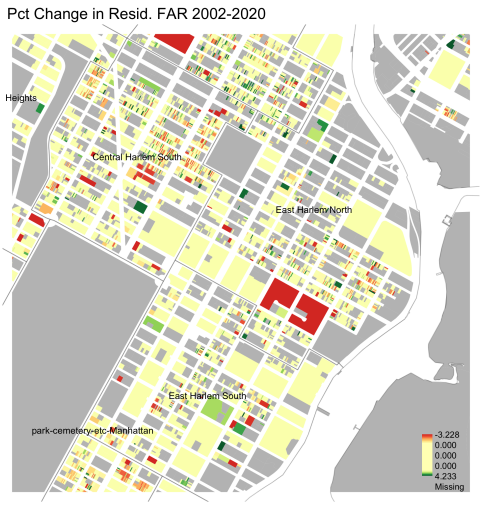
(a) Zoning Change 2002 to 2020



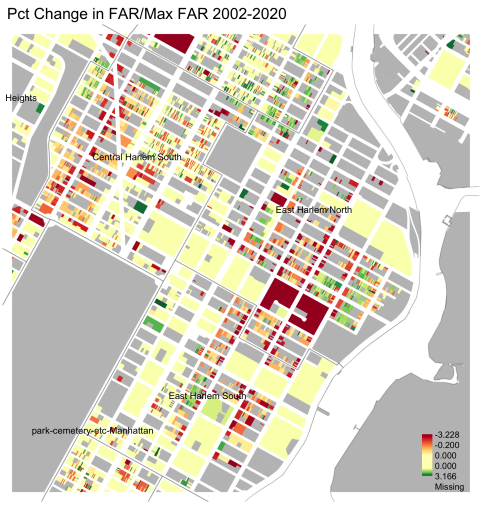
(b) Change in Log Price Per Square Foot 2002 to 2020



(c) Change in Log Res. FAR 2002 to 2020



(d) Change in Log FAR/Max FAR 2002 to 2020

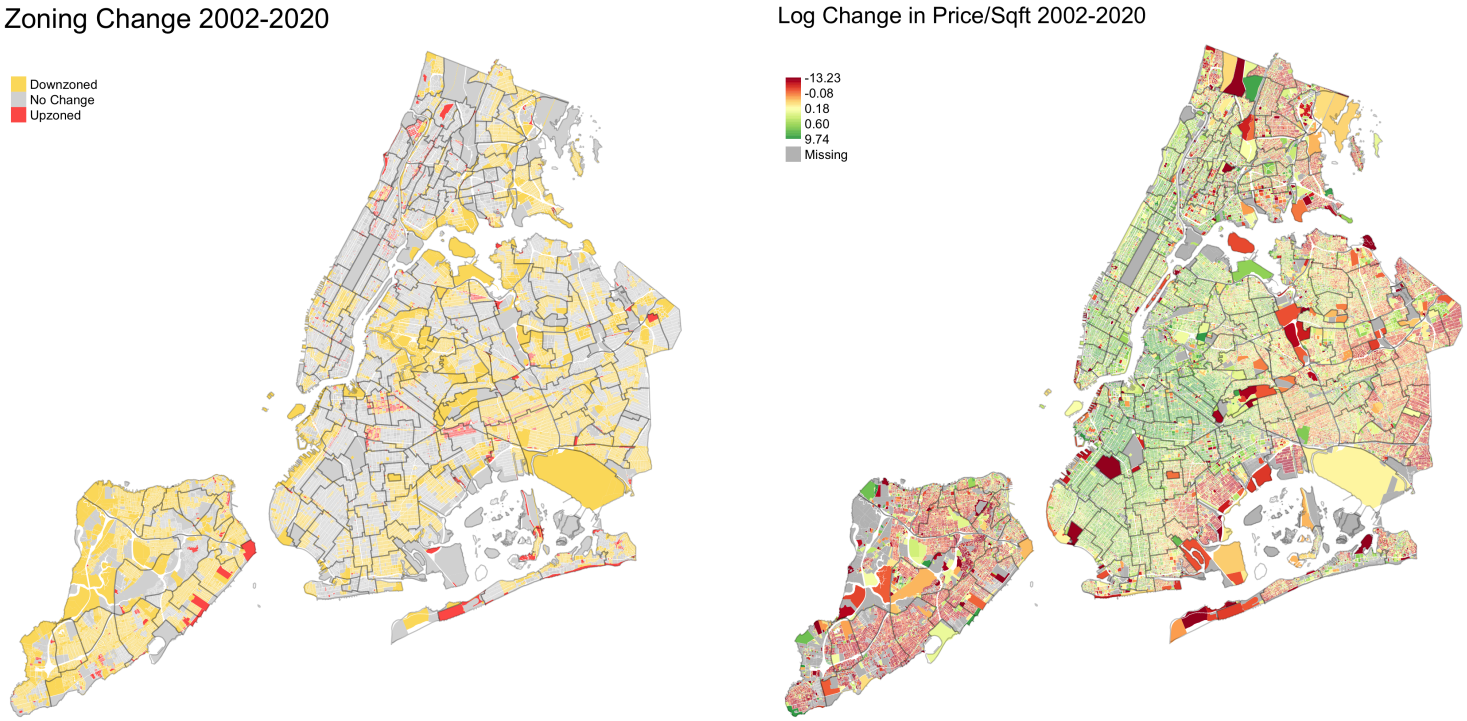


Notes: Data from NYC PLUTO at the tax lot level.

Figure 1.3: Changes in New York City from 2002 to 2020

(a) Zoning Change from 2002 to 2020

(b) Log Change in Price Per Square Foot from 2002 to 2020



Notes: Data from NYC PLUTO at the tax lot level.

1.3.1.1 Price Levels and Policy Changes

First, Figure 1.3 shows maps of New York City and the changes it has undergone from 2002 to 2020. While there was much emphasis placed on upzoning, the map shows that the majority of land rezoned during the Bloomberg administration was downzoned, with specific neighborhoods targeted for upzonings. The graph on the right shows the change in prices, which generally illustrate a pattern that the outer neighborhoods saw price decreases while the neighborhoods closer to Manhattan were increasing in price. The overall relationship between policy change and price is hard to discern at this scale. Figure 1.2(a) zooms into East Harlem, where the corridors were upzoned, and the midblock streets were downzoned. Figure 1.2(c) shows the change in residential FAR. It is difficult to see a clear correlation between policy change and change in the floor area. In Figure 1.2(d), when taking observing the change in the ratio between the FAR and the MFAR from 2002 to 2020, a correlation with the zoning policy change begins to be visible a pattern starts to emerge: the upzoned areas tend to show a decrease in this ratio, i.e. relative to the MFAR allowed, developers are using less, percentage-wise, of the allowable space to build.

As in Table 1.1, Figure 1.1 shows that upzoned areas are, on average, higher in both 2020 price levels and change in price levels from 2002 to 2020. We can see that the simple regression line of the upzoned data through the scatter plot data is above and nearly parallel to that of the downzoned data, indicating that upzoned areas, in general, are associated with higher price per square foot in levels and in changes than downzoned areas.

Table 1.1 shows the aggregated statistics of levels and changes broken out by policy change. The upzoned areas are associated with the highest change in price per square foot and change in price per square foot levels. While the mean change in the maximum floor area ratio ranges from -24% to 38%, the mean change in physical floor area ratio is much more restricted, ranging from -1% to 1%. In fact,

the ratio between the FAR and MFAR becomes higher in the downzoned lots, and lower in the upzoned lots, upholding the observation from Figure 1.2(c). On average in 2020 levels, the FAR in upzoned areas is only about 47% of its potential maximum. Table 1.2 shows the percentage of each policy change and what actually occurred; in all policy changes, the majority of lots had no change. Upzoned lots had the highest percentage of increases in residential area (10%), followed by lots that had no change (8%), and lots which were downzoned (7%).

Table 1.1: Summary Statistics by Zoning Change

	Downzoned	No Change	Upzoned	Total
2002-2020				
Δ ln(Res. Max FAR)	-0.237	-0.016	0.381	-0.114
Δ ln(FAR)	-0.011	-0.001	0.013	-0.005
Δ ln(Pr/Sqft)	0.265	1.55	3.308	0.971
Δ ln(FAR/Max FAR)	0.160	-0.060	-0.438	
2020				
Res. Max FAR	0.644	1.808	2.802	1.275
FAR	0.739	1.424	1.883	1.108
Pr/Sqft	107	222	238	168
FAR/Max FAR	0.977	0.662	0.471	0.803
N	408513	429928	18123	856564
Pct of Total (N)	0.477	0.502	0.021	1.000

Notes: Data from NYC PLUTO, at the tax lot level.

Table 1.2: Changes in Residential Area by Zoning Change

	Change in Policy			
Actual Change	Downzoned	No Change	Upzoned	Total
Pct Increasing in Res. Area	0.069	0.083	0.101	0.076
Pct Decreasing in Res. Area	0.310	0.320	0.286	0.315
Pct No Change in Res. Area	0.622	0.597	0.613	0.609
N	343180	363011	11873	718064

Notes: Data from NYC PLUTO, at the tax lot level.

1.3.1.2 Characteristics of Neighborhoods

Table 1.3 and Table 1.4 are split into quartiles of price per square foot and change in the log of price per square foot, respectively. In general, a higher price per square foot is associated with markers of what some may call a more well-to-do neighborhood: higher median income, higher residential floorspace, and greater proximity to businesses. Higher price per square foot block groups are also associated with a higher percentage of lots being upzoned and less so with being downzoned. A higher change in price per square foot, echoing Table 1.3, is associated with a higher change in median household income, higher change in residential floorspace, and greater proximity to businesses. It is also associated with a higher percentage of lots being upzoned and again, less so with being downzoned. In particular, as upzoned areas are often in the top two quartiles of change in price per square foot, upzoned areas also are associated with higher changes in median household income, what some may consider a sign of gentrification—a common complaint against Bloomberg era upzoning changes.

Table 1.3: Quartiles of Price Per Square Foot in 2020

	1	2	3	4	Total
Median Household Inc (000s)	63.69	57.80	59.26	127.10	72.53
FAR/Max FAR	0.74	0.67	0.58	0.50	0.6167
Residential FAR	1.06	1.32	1.31	2.68	1.591
Percent Upzoned 2002-2020	0.016	0.028	0.043	0.040	0.032
Percent Downzoned 2002-2020	0.462	0.277	0.234	0.162	0.284
Num. of Businesses Within 0.5mi	173	395	653	1333	494
Percent Homeowner in 2000	0.540	0.358	0.296	0.276	0.367
Percent White in 2000	0.432	0.442	0.451	0.628	0.488
Percent College-Educated in 2000	0.139	0.140	0.148	0.394	0.205
Median Household Inc (000s) in 2000	43.79	38.44	36.55	86.86	51.36
N					6100

Notes: Statistics aggregated from the ACS 2016-2019 and NYC PLUTO, at the block group level.

Table 1.4: Quartiles of Δ Log(Price Per Square Foot) from 2002 to 2020

	1	2	3	4	Total
Δ ln(Median Household Inc)	0.476	0.483	0.582	0.800	0.585
Δ ln(FAR/Max FAR)	0.111	0.077	0.086	0.089	0.617
Δ ln(Residential FAR)	0.030	-0.001	0.012	0.036	1.591
Percent Upzoned 2002-2020	0.010	0.019	0.039	0.061	0.032
Percent Downzoned 2002-2020	0.438	0.340	0.222	0.135	0.284
Num. of Businesses Within 0.5mi	135	455	907	1130	494
Percent Homeowner in 2000	0.532	0.377	0.303	0.257	0.367
Percent White in 2000	0.515	0.488	0.498	0.451	0.488
Percent College-Educated in 2000	0.191	0.207	0.213	0.209	0.205
Median Household Inc (000s) in 2000	53.02	53.13	51.96	47.34	51.36
N					6100

Notes: Statistics aggregated from the ACS 2016-2019 and NYC PLUTO, at the block group level.

1.3.2 Reduced Form

Having shown that upzoned areas are associated with high price levels and high price increases despite an overall increase in residential housing supply, I now provide reduced form evidence separating the effect of zoning policy from the increase in residential housing policy on the change in prices. For the sake of brevity, unless otherwise noted, each variable discussed is the log difference. The reduced form equation that I estimate is:

$$\Delta \ln P_i = \beta_0 + \beta_1 \Delta \ln \left(\frac{f_i}{\bar{f}_i} \right) + \beta_2 \Delta \ln f_i + \mathbb{Z}_i + \mathbb{F}_{\text{borough}} + \xi_i \tag{1.1}$$

The unit i is either the tax lot or the aggregated block group. P_i is then the price per square foot, f_i is the FAR, $\bar{f}_i$ is the MFAR, $\mathbb{Z}_i$ are a set of standard controls such as building and neighborhood characteristics both contemporary and past, and $\mathbb{F}$ is a set of borough fixed effects. I allow for spatial correlation by using standard errors

following Conley (1999a). The fraction $f_i/\bar{f}_i$ directly represents the ratio captured in Figure 1.2(c). Thus, the effect of the policy is $\alpha_1 \equiv -\beta_1$ and the effect of the change in FAR is $\alpha_2 \equiv \beta_1 + \beta_2$.

Table 1.5 in columns 1 and 2 present the OLS regressions, without and with controls, respectively. Within this table, there are three panels, A: All Block Groups, B: Block Groups with Zoning Change Prior to 2010, and C: All Tax Lots. Panel A and Panel B use Census block-group level data and thus are able to include detailed demographic controls. Panel C uses the public NYC PLUTO data, and while it is more fine-grained at the tax lot level, does not have detailed demographic controls. First, looking at Panel A, Column 1, there is a positive relationship between FAR and the price, and a negative relationship between the FAR/Max FAR ratio and price, i.e. as the FAR becomes closer to the MFAR, the price goes down. The policy effect and the FAR effect are both positive. To account for other factors that may affect price, I control for average building characteristics of tax lots in the block group, proximity to businesses and trees, 2000 demographics, and the fraction of the block group upzoned and downzoned. When adding controls, the signs on the first two variables flip, and are still both significant. The effect of a 10% increase in the MFAR corresponds to a 0.62% decrease in prices, and the effect of a 10% increase in the FAR results in a 1.6% decrease in prices. Panel B has similar results, but the results are not significant. In Panel C, the effect of a 10% increase in the MFAR corresponds to an increase of 1.8% in price, and the effect of a 10% increase in the FAR results in a 6.2% decrease in price. All three panels show a negative relationship between FAR and price.

However, despite the extensive controls, there may still be unobserved heterogeneity. Areas that received changes in zoning policy were selected for a variety of reasons, such as attractiveness to developers, highest potential for future growth, least opposition to in the voting populace, etc. Thus, the FAR and MFAR are endogenous. Given this, I establish an instrumental variable strategy to establish a causal relationship

between prices and policy changes.

I take advantage of the change in zoning policy in 2002 starting with the Bloomberg administration who planned to redevelop the entire city by implementing coordinated changes in zoning policy. Prior to 2002, New York City had two major revisions of its zoning code policy. The first zoning code was established in 1916 with bulk regulations (e.g., height, setbacks) as the main tools for controlling building size, and then was completely replaced in 1961 which expanded the set of zoning categories and replaced bulk regulations with FAR as the main tool for controlling building size. Both codes mostly codified what was already in place without an overall plan. I exploit the variation in 1961 zoning resulting from a change in zoning tools, and combine it with any pre-2002 FAR changes to predict the ratio of FAR to MFAR. The first stage is as follows:

$$\Delta \ln\left(\frac{f_i}{\bar{f}_i}\right) = \alpha_0 + \alpha_1[\ln(\text{FAR 2002}) - \ln(\text{Max FAR 1961})] + \alpha_2 \ln(\text{Max FAR 1961}) + \mathbb{Z}_i + \mathbb{F}_{\text{borough}} + \mu_i \tag{1.2}$$

The second stage uses the predicted measure of $\ln(f_i/\bar{f}_i)$ from the first stage. Overall, I expect OLS values to be biased towards 0. For example, areas that are upzoned are selected with anticipation that developers will be interested in building; developers will be interested in building given a high enough future return. From this line of reasoning, upzoned areas may be likely to be places where prices will rise relatively more in the future, and so the effect of raising the MFAR and the actual FAR will be dampened o

The intuition of the instrument is as follows: an area is more likely to be more constrained when the difference between the 1961 MFAR and the 2002 FAR is larger; i.e. a zoning change already occurred in the previous decades, buildings were changed

significantly, and thus during the 2002-2013 rezoning, the area would be less likely to have a change in zoning policy, and thus become relatively more constrained. The level of the MFAR in 1961 should also have a positive effect on the change in the ratio between FAR and Max FAR for similar reasons: as the higher the original level is, the less likely there is to be a policy change, and the more likely the area to become more zoning constrained. Therefore, the relationships between the instruments and instrumented variable should be positive. In Table 1.5, Column 3 of all panels, these instruments have a strong first stage and are positively correlated with the ratio. The exclusion restriction states that the two instruments cannot directly affect the change in price. Given appropriate controls and that these policy changes and changes to residential FAR were made before knowledge of Bloomberg's rezoning policies and the 1961 ordinance mostly reflected established norms, the instruments should not have a direct effect on the change in prices from 2002 to 2020.

Column 3 in Table 1.5 shows the first stage, in which all first stages are positive and with similar magnitudes to each other. Panel C includes an upzoned and downzoned indicator as it is at the tax lot level; Panel A and Panel B contain upzoned and downzoned percentages (not shown). Panel A and Panel B contain very similar results. Column 4, Panel A, contains the preferred specification. A 10% increase in the floor area ratio results in a 1.7% decrease in price, and a 10% increase in the maximum floor area ratio results in a 2.4% decrease in price. Similar results are echoed in Panel B, which only considers block groups which had a zoning change prior to 2010, allowing 10 years for building changes to occur. The magnitude of policy effect is somewhat larger and the magnitude of FAR effect is somewhat smaller. Panel C has a larger negative effect from the change in policy and a small positive effect from the change in FAR. The different results in Panel C may reflect the granularity of the data and that in a tax lot, an indicator of being upzoned or downzoned may be more significant than the magnitude of the change in policy and FAR.

Table 1.5: Dependent Variable: Change in ln(Price per Square Foot)

	(1) OLS	(2) OLS	(3) IV - First Stage	(4) IV - Second Stage
Panel A: All Block Groups				
ln(FAR/Max FAR)	-0.0722	0.0617		0.2405
	(0.0298)***	(0.0189)***		(0.0782)***
ln(FAR)	0.1007	-0.219		-0.4098
	(0.0475)***	(0.0537)***		(0.0987)***
ln(FAR 2000) - ln(Max FAR 1961)			0.2659	
			(0.0342)***	
ln(Max FAR 1961)			0.2893	
			(0.0329)***	
Policy Effect (α_1)	0.0722	-0.0617		-0.2405
FAR Effect (α_2)	0.0285	-0.1573		-0.1693
F-Stat			41.83	
N	6100	6100	6100	6100
Controls		X	X	X
Panel B: Block Groups with Zoning Change Prior to 2010				
ln(FAR/Max FAR)	-0.1139	0.0036		0.3576
	(0.0733)	(0.0477)		(0.1516)***
ln(FAR)	0.1326	-0.0964		-0.4628
	(0.0849)	(0.0553)*		(0.1587)***
ln(FAR 2000) - ln(Max FAR 1961)			0.2459	
			(0.0377)***	
ln(Max FAR 1961)			0.2062	
			(0.026)***	
Policy Effect (α_1)	0.1139	0.0036		-0.3576
FAR Effect (α_2)	0.0187	-0.0928		-0.1052
F-Stat			33.8	
N	3500	3500	3500	3500
Controls		X	X	X
Panel C: All Tax Lots				
ln(FAR/Max FAR)	-0.455	-0.183		0.492
	(0.071)***	(0.023)***		(0.115)***
ln(FAR)	0.116	-0.435	-0.008	-0.426
	(0.076)	(0.047)***	(0.007)	(0.048)***
Upzoned Indicator			-0.325	0.296
			(0.026)***	(0.073)***
Downzoned Indicator			0.293	-0.306
			(0.017)***	(0.034)***
ln(FAR 2000) - ln(Max FAR 1961)			0.278	
			(0.031)***	
ln(Max FAR 1961)			0.254	
			(0.025)***	
Policy Effect (α_1)	0.455	0.183		-0.492
FAR Effect (α_2)	-0.339	-0.618		0.066
F Stat			54.008	
N	712038	703693	701957	701957
Controls		X	X	X

Notes: All variables are first differences from 2002–2020 except for the instruments. Bandwidth for spatial HAC is equivalent to half mile, or about 10 North-South blocks in New York City. Regressions are at the block group level for Panel A and Panel C, and are at the tax lot level for Panel C.

1.4 Spatial Equilibrium Model of a City

1.4.1 Setup

The empirical results suggest that an increase in FAR results in lower price per square foot and a higher share of high-income residents, but the increase in FAR does not necessarily result in a building that is built to the MFAR. I build on the models of Ahlfeldt et al. (2015) and Tsivanidis (2019) to develop a spatial equilibrium model of a city to capture these observations in order to identify the citywide effect of zoning on residential and commercial development.

The model starts with a closed city; expected utility is endogenously determined. The city is made up of locations $i, j \in \{1, \ldots, L\}$ differing in lot area, floor area ratio, amenities and commercial productivity, and commute times to other locations. Each location is also affected by a density spillover, whereby increased density results in increased amenities. All residents are renters who choose where to live and where to work based on the location characteristics. Each location has one firm which produces a freely traded numeraire good, and demand for workers depends on the productivity of each firm in each location. Last, the class of developers supply residential floorspace with convex construction costs. All land is owned by the residents and all profits are redistributed equally to each resident.

1.4.2 Individuals

The city is populated by a fixed population $\bar{L}$ which is divided into two groups of workers based on income, low or high: $g \in G = \{n, h\}$. The total number of residents in each group is exogenously determined and denoted by $\bar{L}_g$. Residents are indexed by ω and supply one unit of labor inelastically. Residents choose a location i to live and a location j where to work. They consume a numeraire good $C_i(\omega)$ and consume floorspace $H_i(\omega)$ at cost r_i subject to non-homothetic preferences with

minimum amount of housing $\bar{h}$. The Cobb-Douglas parameter β is the fraction of income spent on housing. Each location has a group- and location-specific amenity u_{ig}, and each individual receives a location-specific preference shock $\xi_i(\omega)$ distributed Fréchet: $F(\xi_i) = \exp\left(-\xi_i^{-\eta_g}\right)$.

Residents receive income from their labor and land rents which are equally distributed amongst all residents (φ). Wage is dependent on location j and group g, is subject to a location-specific productivity shock distributed Fréchet: $F(\varepsilon_j) = \exp(-T_g\varepsilon_j^{-\theta_g})$, and is discounted by a commuting cost d_{ij}. The commuting cost is parametrized as $d_{ij} = \exp(\kappa t_{ij}) \geq 1$, where $\kappa > 0$ is a multiplicative constant scaling the size of costs, and $\tau_{ij} \in [0, \infty)$ is a measure of travel time or distance between locations i and j.

The utility of a worker ω is then:

$$\max_{C_i(\omega), H_i(\omega)} u_{ig} C_i(\omega)^{\beta} (H_i(\omega) - \bar{h})^{1-\beta} \xi_i(\omega) \tag{1.3}$$
$$\text{subject to: } C_i(\omega) + r_i H_i(\omega) = \frac{w_{jg}\varepsilon_j(\omega)}{d_{ij}} + \varphi.$$

Solving for optimal demand for the consumption good and housing results in the following expression for indirect utility:

$$U_{ijg}(\omega) = u_{ig}\left(\frac{w_{jg}\varepsilon_j(\omega)}{d_{ij}} - r_i\bar{h} + \varphi\right) r_i^{\beta-1}\xi_i(\omega) \tag{1.4}$$

Demand for housing for each location i may expressed as:

$$r_i = (1-\beta)\frac{X_i}{H_i - \beta\bar{h}L_i} \tag{1.5}$$

where X_i is the total spending on goods and housing from residents in i.

Decision Making The individual chooses first where to live and then where to work. The problem is solved via backward induction. I first consider the second stage in which the individual chooses the workplace location given that they have already chosen where to live. From the Fréchet distribution which depends on group, the worker draws a vector IID match-productivities for each location j. The scale parameter $T_g > 0$ determines the average productivity of workers in group g and the parameter $\theta_g > 1$ determines the dispersion of worker productivity in group g; a higher value of θ_g denotes a lower dispersion. From this draw, the worker chooses the work location with the highest effective wage: $\max\{w_{jg}\varepsilon_j(\omega)/d_{ij}\}$. The properties of the Frèchet distribution allows for the derivation of the probability a worker deciding to work in j given that they are in group g and live in i:

$$\pi_{j\,|\,gi} = \frac{(w_{jg}/d_{ij})^{\theta_g}}{\sum_k (w_{kg}/d_{ik}^g)^{\theta_g}} = \frac{(w_{jg}/d_{ij})^{\theta_g}}{\underbrace{RMA_{ig}}_{\text{Residential Market Access}}} = \frac{L_{ijg}}{L_{ig}} \tag{1.6}$$

Individuals are more likely to work in a location that pays a higher net wage relative to all other locations. The denominator, RMA_{ig} may be termed Residential Market Access, and is a measure of the location's proximity to high-paying jobs; the closer a location is to high-paying jobs, the larger RMA_{ig} will be. The parameter θ_g may also be thought of as a commuting elasticity determining the sensitivity of each group to commute costs.

In the first stage, individuals then choose their residential locations. They first draw a vector IID preference shocks over locations, drawn from the Fréchet distribution with dispersion parameter $\eta_g > 1$. Similar to θ_g, a larger η_g denotes a lower dispersion of shock values. Properties of the Fréchet distribution result in the following probability of living in location i given the individual is part of group g:

$$\pi_{i\,|\,g} = \frac{\left(u_{ig}(\bar{y}_{ig} - r_i\bar{h})r_i^{\beta-1}\right)^{\eta_g}}{\sum_{k=1}^{L}\left(u_{kg}(\bar{y}_{kg} - r_k\bar{h})r_i^{\beta-1}\right)^{\eta_g}} = \frac{L_{ig}}{\bar{L}_g} \tag{1.7}$$

where the expected income prior to the drawing of productivity shocks is $\bar{y}_{ig} = T_g^{1/\theta_g}\Gamma(1-\frac{1}{\theta_g})(RMA_{ig})^{1/\theta_g} + \varphi$. The following expression, $\Gamma(\cdot)$, is the gamma function. Workers are more likely to live in areas that feature a high level of amenities, are closer to potentially high net incomes, and have low rent.

From the derived probabilities and commuting market clearing condition which requires that the number of group g workers who commute to j must equal the number of workers in group g that work in j, I derive the number of residents that work in j:

$$L_{jg} = \sum_{i=1}^{L}\pi_{j\,|\,ig}L_{ig} = w_{jg}^{\theta_g}\sum_{i=1}^{L}\frac{L_{ig}}{d_{ij}^{\theta_g}RMA_{ig}} = w_{jg}^{\theta_g}\underbrace{FMA_{jg}}_{\text{Firm Market Access}} \tag{1.8}$$

where FMA_{ig} is defined as the Firm Market Access, which similarly to the Residential Market Access, is a measure of the access of the firms to type g workers in location i. FMA_{ig} becomes larger when there are more people nearby at a shorter commuting distance.

Welfare The average welfare in each location is equal to the expected utility of a type g worker and is expressed as:

$$\bar{U}_g = \Gamma\left(1 - \frac{1}{\eta_g}\right)\left[\sum_{i=1}^{L}\left(u_{ig}(\bar{y}_{ig} - r_i\bar{h})r_i^{\beta-1}\right)^{\eta_g}\right]^{1/\eta_g}. \tag{1.9}$$

1.4.3 Firms

Firms produce a single final good under perfect competition and constant returns to scale. The good is costlessly traded. The firm uses a CES combination N_j of effective low- and high-productivity labor $\tilde{L}_{jg}$ as inputs with elasticity of substitution σ and group labor share α_g where $\sum_g \alpha_g = 1$.

$$\begin{aligned} Y_j &= A_j N_j \\ N_j &= \left(\sum_g \alpha_g \tilde{L}_{jg}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \\ \tilde{L}_{jg} &= \bar{\varepsilon}_{jg} L_{jg} \end{aligned} \tag{1.10}$$

where A_j is location j productivity and $\bar{\varepsilon}_{jg}$ is the average productivity of an individual from group g who works in j. Maximizing their profits, firms set a price equal to marginal cost:

$$\begin{aligned} p_j &= \frac{W_j}{A_j} \\ W_j &= \left(\sum_g \alpha_g^{\sigma} w_{jg}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \end{aligned} \tag{1.11}$$

where W_j is the aggregate labor cost in location j. Demand for labor is given by:

$$\tilde{L}_{jg} = \left(\frac{\alpha_g W_j}{w_{jg}} \right)^{\sigma} N_j \tag{1.12}$$

and increases with the productivity of a location and decreases when the wage of workers increases.

1.4.4 Developers

Developers take the residential MFAR $\bar{f}_i$, lot area $\bar{Q}_i$ and convex building costs $C_i(f_i/\bar{f}_i)$ as given. Developers take the land cost per unit of lot area, R_i as given, and are in perfect competition. Floor area ratio f_i is produced in a Cobb-Douglas function with exogenous per lot building productivity A_i^H, freely traded capital K_i, and lot area with the land rents eventually redistributed equally to each resident. The owners of capital do not live in the city and the price of the capital is the same in every location.

The developer solves the following profit maximization problem:

$$\begin{aligned}
\max_{K_i} \pi_i &= \max_{K_i} r_i H_i - p_k K_i - C_i(\bar{f}_i)\bar{Q}_i - R_i\bar{Q}_i \\
\text{subject to: } & H_i = f_i\bar{Q}_i \\
& f_i = A_i^H \bar{Q}_i^{\alpha_h} K_i^{1-\alpha_h} \leq \bar{f}_i \text{ else } f_i = \bar{f}_i \\
& C_i(f_i/\bar{f}_i) = p_k \left[\frac{f_i^\nu/\bar{f}_i}{A_i^H \bar{Q}_i^{\alpha_h}} \right]^{\frac{1}{\nu(1-\alpha_h)}}, \nu \geq 1
\end{aligned}$$

where the convex building cost function is based on the variable cost function in Baum-Snow and Han (2019) and $\nu > 0$ is a building cost elasticity. To derive the bid-rent function, apply the zero profit condition and marginal cost pricing:

$$R_i = \frac{dC_i(f_i/\bar{f}_i)}{df_i}\frac{H_i}{\bar{Q}_i} - \frac{C_i(f_i/\bar{f}_i)}{\bar{Q}_i} - \frac{K_i}{\bar{Q}_i} \tag{1.13}$$

Demand for housing is then:

$$r_i = \nu_1 \left[A_i^H \bar{Q}_i\right]^{-\nu_1} \left(\frac{f_i}{\bar{f}_i}\right)^{\nu_1} f_i^{\nu_2} \tag{1.14}$$

where $\nu_1 > 0$ and $\nu_2 < 0$.[4] Thus, depending on the characteristics of the lot, there are two competing effects from the choice of FAR. First, as the ratio of FAR to MFAR increases, the price of housing increases, which can be thought of as the regulatory burden increasing as the building reaches its MFAR (Brueckner et al., 2017). The second competing effect comes from ν_2 where as the FAR increases, the housing price decreases, reflecting some economies of scale as more floor area is built (Chau et al., 2007). The housing cost decreases with housing productivity and size of lot area (Mohamed, 2009).

1.4.5 Spillovers

This model only features externalities in amenities. Amenities in a neighborhood depend on two components: (1) an exogenous $\bar{u}_{ig}$ which captures the underlying desirability of a location such as proximity to green space and closeness to subway station and (2) density of residents which is defined as number of individuals living in location i divided by the lot area $\bar{Q}_i$:

$$u_{ig} = \bar{u}_{ig} \left(\frac{L_i}{\bar{f}_i \bar{Q}_i}\right)^{\rho_g} \tag{1.15}$$

where ρ_g is the elasticity of density preference for each group. This agglomeration force acts to segregate the choices of low- and high-productivity workers even more depending on their differing preferences for density and valuation of the initial location.

[4] $\nu_1 = \frac{1}{\nu(1-\alpha_h)} > 0, \;\; \nu_2 = \frac{\nu - \nu^2(1-\alpha_h) - 1}{\nu(1-\alpha_h)} < 0$

1.4.6 Equilibrium

The equilibrium of the closed city model is defined as follows:

Definition Given parameters $\{\kappa, \sigma, \beta, T_g, \bar{h}, \eta_g, \rho_g, \theta_g, \nu\}$, exogenous city population $\{\bar{L}_g\}$, exogenous location-based characteristics $\{t_{ij}, \bar{f}_i, \bar{Q}_i, A_i, \bar{u}_{ig}\}$, the general equilibrium of the model is given by the endogenous vectors $\{[w_{jg}, L_{ig}, L_{jg}, \bar{U}_g]_{g\in\{n,h\}}$, $r_i, f_i, R_i, \varphi\}$ such that markets clear in labor (1.8, 1.12), markets clear in housing (1.14, 1.5), workers maximize their utility given their budget constraints (1.9), firms maximize their profits (1.11), developers maximize their profits given regulatory maximums and costs (1.13), and populations add up to the city total, i.e. $\bar{L}_g = \sum_i L_{ig}$ for $g \in \{n, h\}$.

1.4.7 Parameters

Table A.5 summarizes the parameters in the model and their values. Two parameters are taken from the literature, (1) κ, which governs the size of commuting costs, and (2) σ, the elasticity of substitution between high and low-productivity workers from Ahlfeldt et al. (2015) and Card (2009), respectively.

Calibration The fraction of expenditure on housing, β, is matched to the average fraction of housing costs (rent, mortgage, utilities, insurance, etc.) divided by total household income using the 2000 Decennial Census.

Following Tsivanidis (2019), the Fréchet parameter governing the level of worker productivity is normalized to 1 for the low-productivity worker. The wage premium in the model is then set to the wage premium in the data, resulting in $T_h = 4.14$.[5]

The minimum level of housing, $\bar{h} = 80$, is set to the NYC Administrative Code Section 27-2074 *Minimum room sizes*, which states "Every other living room of an apartment in a multiple dwelling erected after April eighteenth, nineteen hundred

[5]The wage premium is calculated as follows: Wage Premium $= \frac{T_h \sum_i RMA_{ih}^{1/\theta_h} \frac{L_{ih}}{L_i}}{\sum_i RMA_{in}^{1/\theta_n} \frac{L_{in}}{L_i}}$

twenty-nine shall contain eighty square feet."

Estimation The remaining parameters are all estimated using equations from the model.

The density spillover parameter ρ_g and living preference parameter η_g are estimated using Equation 1.7, which represents the supply of residents to an area (to live). After taking logs and differences, we arrive at the following equation:

$$\Delta \ln L_{ig} = \eta_g \rho_g \Delta \ln \frac{L_i}{f_i \bar{Q}_i} + \eta_g \Delta \left[\ln y_{ig} + (\beta - 1) \ln r_i \right] + \mathbb{Z}_i + \Delta \ln \varepsilon_{ig} \tag{1.16}$$

$\Delta \ln \varepsilon_{ig}$ represents unobserved changes in location, group-specific amenities. Because changes to density, income, and rent are endogenous, in order to identify η_g and ρ_g, I employ an instrumental variables strategy and use four instruments: the change in college-educated fraction from 1990 to 2000, the fraction of college-educated in 1990, the change in the count from 1990 to 2000, and the count in 1990. In particular, higher count level and count growth prior to 2000 should correspond to lower density change in 2000-2020, all else constant; higher fraction of college-educated and change in college-educated should correspond to higher income growth relatively due to the increasing gap between skilled and non-skilled earnings (Acemoglu and Restrepo, 2022). Table 1.8 shows the two preferred specifications with full controls using LIML (limited information maximum likelihood). LIML is chosen over 2SLS for its efficiency when dealing with multiple (weak) instruments (Blomquist and Dahlberg, 1999). From this, η_h is determined to be 1.082 and ρ_h is 1.275; η_n is 0.581 and ρ_n is 0.897. For the high-productivity, the higher value of η_h indicates a lower dispersion of living preferences, and the higher value of ρ_h indicates a greater preference for density.

The parameter governing the sensitivity to commuting costs, θ_g, is determined

using the commuting gravity equation (1.6). Results are shown in Table 1.7. LODES data were used instead of Decennial Census data in this case due to the incompleteness of self-reporting on Census data; the most prevalent issues being that some observations did not report work locations, and work locations reported were too general (e.g., reporting the city instead of street address). LODES defines three monthly income groups: earning $1250 or less, $1251 to $3333, and over $3000. I categorize the first two groups as low-income and the last group as high-income. On average, in New York City, this high-income group makes up 42% of each block group.

$$\ln \pi_{j \mid ig} = \text{Origin}_{ig} + \text{Destination}_{jg} - \theta_g \kappa t_{ij} + \varepsilon_{ijg} \tag{1.17}$$

Taking $\kappa = 0.01$ from literature, and calculating the fraction of people who primarily walk to work in NYC using the Decennial Census 2000, we arrive at a weighted value of θ_g for commuting costs measured in terms of duration and distance. In general, the low-income workers are more sensitive to commuting costs in New York City. The magnitudes are similar to Hsieh et al. (2019) and Acosta (2020), though neither estimate these values for New York City specifically. In my model, I choose to use the θ_g derived from duration estimates as commutes are more naturally characterized by time, not distances.

Last, the elasticity of variable costs with respect to FAR, ν is estimated in Section 1.3 based on Equation 1.14.

Table 1.6: Estimating θ_g

	High-Income	Low-Income
Driving Duration	-0.0126	-0.0146
	(0.0001)***	(0.0001)***
Driving Distance	-0.0124	-0.0138
	(0.0001)***	(0.0001)***
Walking Duration	-0.0013	-0.0015
	(0.0000)***	(0.0000)***
Walking Distance	-0.0118	-0.0132
	(0.0001)***	(0.0001)***
Block Group Fixed Effects	X	X
Year	2002	2002

Notes: Data from 2002 LODES.

Table 1.7: Calculating θ_g

	θ_h	θ_n
Duration	1.22	1.42
Distance	1.20	1.34
Fraction of Walk	0.03	

Table 1.8: Estimating ρ_g, η_g

	(1a) Second Stage Log High-Income Count	(1b) First Stage Log Density	(1c) First Stage Log Indirect Utility
Log Density ($\eta_g\rho_g$)	1.380*** (6.17)		
Log Indirect Utility (η_g)	1.082** (3.24)		
College Fraction 1990-2000		-0.165* (-2.14)	0.239*** (3.96)
College Fraction 1990		-0.0993 (-1.12)	0.336*** (4.96)
Count 1990-2000		-0.0000946*** (-6.14)	-0.0000219 (-1.63)
Count Fraction 1990		-0.0000945*** (-6.13)	-0.0000219 (-1.63)
N	4400	4400	4400
F-stat	7.475		
Hansen J Stat	0.245		
Hansen J P-Value	0.8848		
	(2a) Second Stage Log High-Income Count	(2b) First Stage Log Density	(2c) First Stage Log Indirect Utility
Log Density ($\eta_g\rho_g$)	0.521*** (4.12)		
Log Indirect Utility (η_g)	0.581*** (6.51)		
College Fraction 1990-2000		-0.0976 (-1.19)	0.673*** (8.85)
College Fraction 1990		-0.0360 (-0.40)	0.709*** (7.86)
Count 1990-2000		-0.0000912*** (-5.72)	-0.0000604*** (-4.16)
Count Fraction 1990		-0.0000910*** (-5.71)	-0.0000603*** (-4.15)
N	4400	4400	4400
F-stat	14.08		
Hansen J Stat	1.255		
Hansen J P-Value	0.5339		

Notes: All logged variables are first differences unless otherwise indicated. Values in parentheses are p-values, not standard errors. Instead of 2SLS, LIML is used for its better small sample properties.

1.4.8 Model Calibration

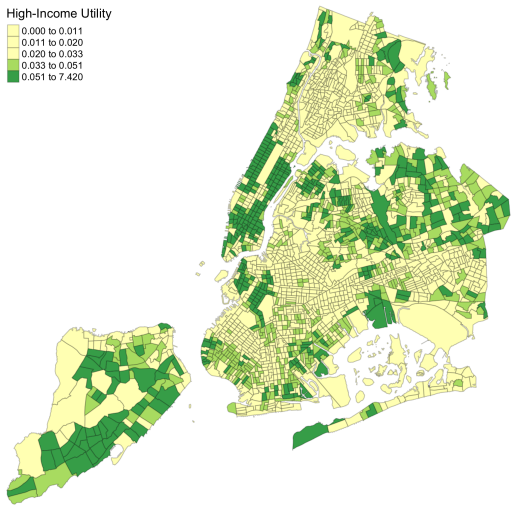
The model matches commuting shares, the number of residents, and the number of workers in each Census tract exactly. The model's housing supply is highly correlated with the NYC PLUTO's 2002 housing supply at a value of 0.6837. The aggregated wage W_i exhibits a 0.3277 correlation with the 2000 Decennial Census median wage.

In the model calibration, implicit amenity values by tract and income group were recovered. Figure 1.4 shows the geographic distribution. Neighborhoods such as Manhattan's Lower East Side, Upper West Side, and Brooklyn's Greenpoint are favored highly by both groups. The high-income group amenities are concentrated more strongly towards New York City's two central business districts (CBDs), Midtown and Lower Manhattan, and in neighborhoods such as Dumbo and Greenpoint which feature quicker subway access to the CBDs.

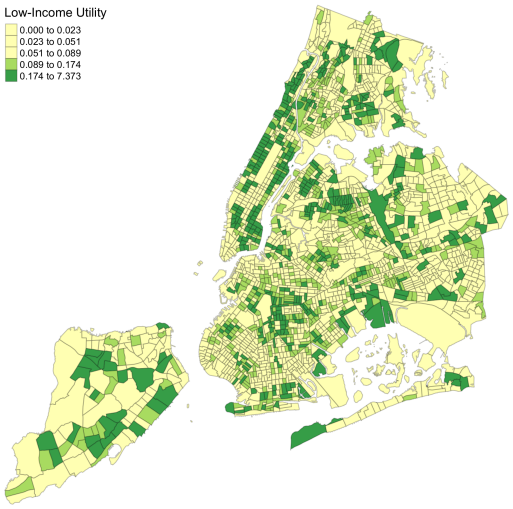
Figure 1.5 shows the geographic distribution of aggregated wage and the rent (price per square foot). As expected, the most expensive rent is concentrated on Lower and Midtown Manhattan, and parts of the boroughs closest with the easiest subway access to the CBDs. Comparing Figure 1.4(a) and Figure 1.5(a), there is a substantial overlap of tracts that have more high-income amenities and that have higher rent. Last, as expected, the aggregate wage is highest when near the CBDs.

Figure 1.4: Amenities

(a) High-Income



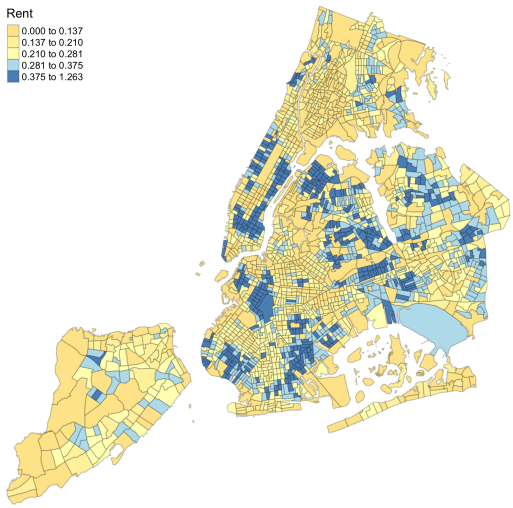
(b) Low-Income



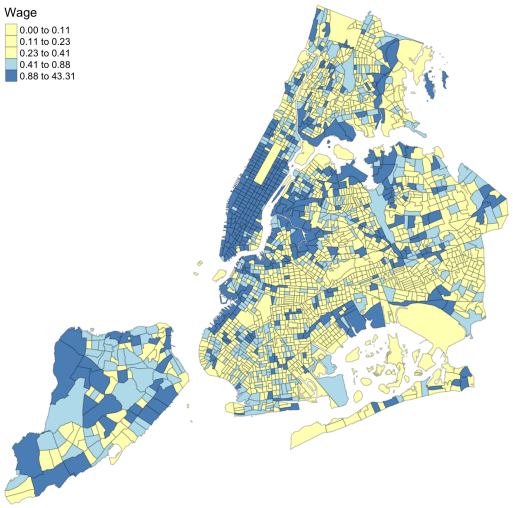
Notes: Data from baseline model simulation using 2002 LODES commuting shares and populations, and NYC PLUTO land characteristics from 2002.

Figure 1.5: Prices

(a) Rent



(b) Aggregate Wage



Notes: Data from baseline model simulation using 2002 LODES commuting shares and populations, and NYC PLUTO land characteristics from 2002.

These maps suggest that in terms of applying zoning policy, the city would benefit most from increasing residential capacity in neighborhoods such as the Upper West Side and Upper East Side which hold high amenity values for both groups, and are close to the productive areas of the city.

1.5 Counterfactuals and Welfare Decomposition

1.5.1 Counterfactuals: Bloomberg Rezoning

This section presents the results from two counterfactuals, which are summarized in Table 1.9. In the first counterfactual, New York City starts at the 2002 baseline values and shifts to the 2020 MFAR regime. In the second counterfactual, New York City starts at the 2002 baseline values, shifts to the 2020 MFAR regime, and also receives a 10% increase in high-productivity residents.

Table 1.9: Counterfactual Results

	Counterfactual 1 2002 to 2020 MFAR	*Counterfactual 2* 2002 to 2020 MFAR, 10% Increase in High-Income
GDP	0.010	0.023
Rent	-0.069	-0.030
Welfare H	0.011	0.013
Welfare N	-0.056	-0.052
Inequality	0.070	0.069
Avg Housing	-0.135	-0.162
Dissimilarity	0.207	0.290

Notes: Inequality is calculated as the ratio of high-productivity welfare to low-productivity welfare. Average housing is calculated as the total sum of residential housing divided by total residents.

Counterfactual 1 In the first counterfactual, GDP increases by 1% and average rent drops by -6.9%. Average housing consumed decreases by -13.5% and the dissimilarity index increases by 20.7%, indicating an increase of segregation between

the two groups of residents. Figure 1.6 presents a geographic overview of the population shifts. As seen in Figures 1.6(c) and 1.6(d), high-productivity individuals moved into areas with amenities attractive to their group and closer to the CBDs. With increased proximity to higher paying jobs, an average decrease in rents, and increase in amenities enjoyed, high-productivity residents experience a 1.1% gain in welfare. However, the low-productivity residents on average move further away from higher paying jobs and to less dense areas, resulting in a decrease in welfare of -5.6%.

From this information, I build two simple indicators of gentrification: first, any area with a change in the high-income fraction greater than 10%; second, any area where the number of high-income movers into an area exceed the number of low-income movers (who may also be leavers). Figure 1.7 shows the classifications on a map. By the first definition, gentrification occurred in only a few tracts (48 out of 2168 tracts). Some of the areas, such as Lower Manhattan, are already home to a significant population of high-income residents, and so some observers may not consider such change to be gentrification. With the second measure, 312 tracts were identified, including more areas anecdotally familiar as gentrified spaces: Greenpoint, Dumbo, and Hudson Yards.

This counterfactual suggests that the rezonings helped to reorganize the city such that workers could move closer to high productivity locations and decrease rents with an overall increase in housing supply. However, depending on the selected measure of gentrification, areas that were desirable to live in for both high and low-income residents were dominated by high-income residents, resulting in increased segregation.

Counterfactual 2 From the outset, the Bloomberg administration stated that the rezonings were chosen to attract high-productivity talent. Additionally, between 2000 and 2020, the population of New York City grew by almost 10%. To simulate this change, Counterfactual 2 introduces a 10% increase of high-productivity residents. The aggregate changes listed in Table 1.9 are qualitatively similar to the

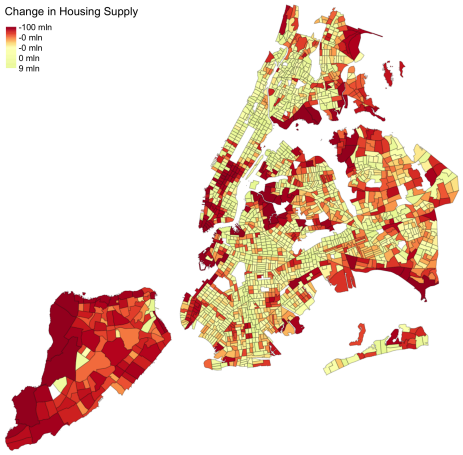
first counterfactual. GDP increases by 2.3% (1% in Counterfactual 1), and due to increased demand for housing, rent only decreases by -3% (-6.9% in Counterfactual 1). Dissimilarity increases by 29% (20.7% in Counterfactual 1), which indicates even more segregation between groups. High-income residents experience a gain in welfare of 1.3% and low-income residents experience a decrease in welfare of -5.2%. Compared to the first counterfactual, welfare for both groups see small increases, driven by increases in wages, and the mechanical increase in density.

Comparing population movements for both counterfactuals in Figure 1.6 and Figure A.2 reveal similar patterns, though the ranges in Counterfactual 2 are larger. In terms of the gentrification measures shown on Figure 1.8, the first measure qualitatively looks similar to its counterpart map from Counterfactual 1. For the second measure, there is some divergence with more high-income residents choosing to live further away from the CBDs.

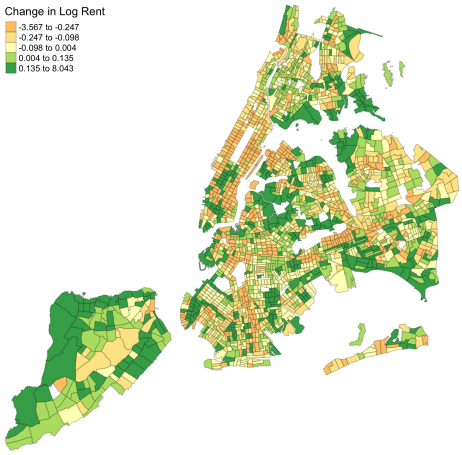
This counterfactual suggests that the pressure of additional high-productivity residents mediated the decrease in rents, contributed to an increase in segregation across the city, but also contributed to increased GDP growth and two small increases in welfare across groups.

Figure 1.6: Counterfactual 1

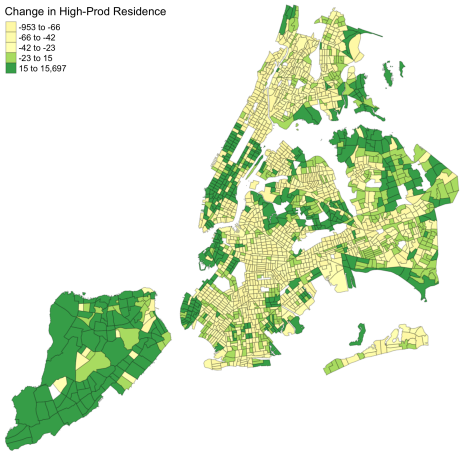
(a) Change in Housing Supply



(b) Percent Change in Rent



(c) Change in High-Productivity Residents



(d) Change in High-Productivity Fraction

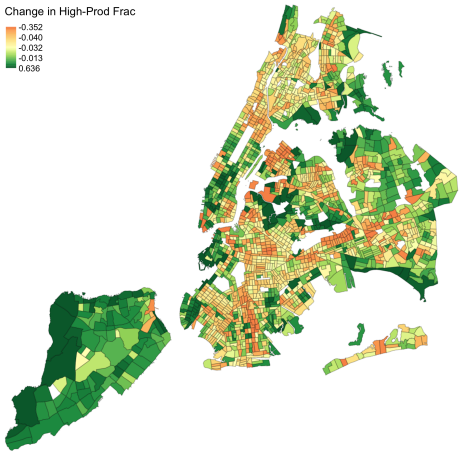
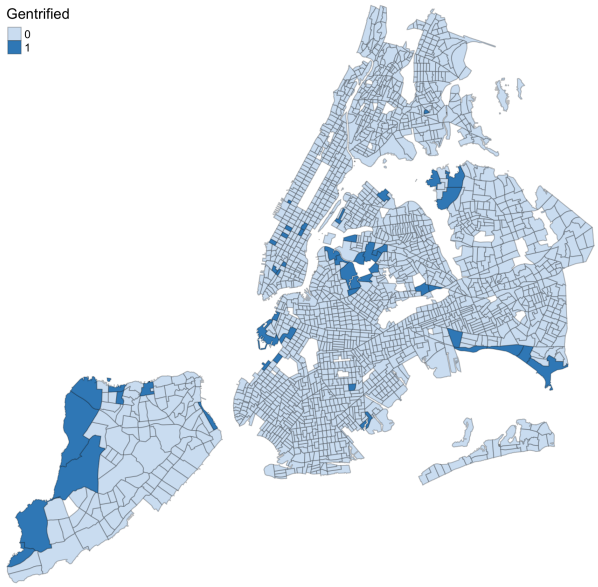


Figure 1.7: Counterfactual 1: Measures of Gentrification

(a) Δ High-Income Fraction $\geq 10\%$



(b) No. of High-Income Movers > No. of Low-Income Movers

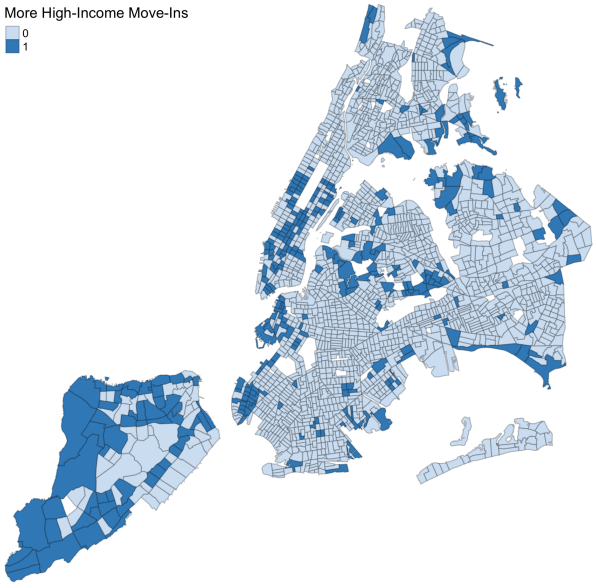
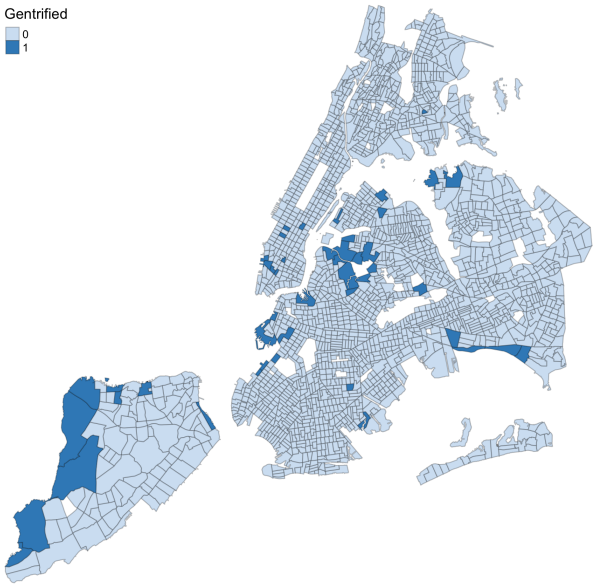
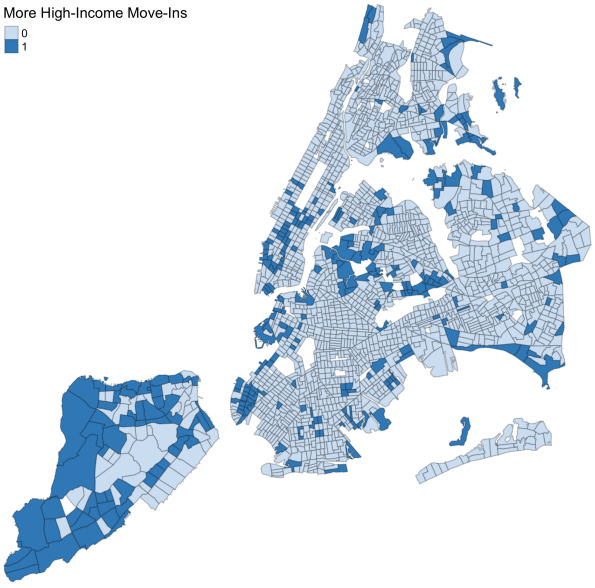


Figure 1.8: Counterfactual 2: Measures of Gentrification

(a) Δ High-Income Fraction $\geq 10\%$



(b) No. of High-Income Movers > No. of Low-Income Movers



1.5.2 Welfare Decomposition

Table 1.10 presents results for a sample welfare decomposition for the Brooklyn neighborhood, Dumbo. Using the equation 1.18, income-specific welfare changes may be decomposed into four parts: changes in housing prices, amenities, residential market access for jobs, and land rent income. We increase the Dumbo MFAR by 10 percent while holding housing supply constant in other tracts, resulting in an overall increase in welfare for high-income (0.0022%) and low-income (0.2769%) groups. Most of the welfare gains come from the increase in access to better-paying jobs, and the decrease in rent. Due to non-homothetic consumption of housing, decreases in rent disproportionately benefit low-income residents more and leads to a greater increase of welfare. Changes in amenities and land rent are negligible.

$$\delta_{\bar{U}_g,f_i} = \sum_{i=1}^{L} \pi_{ig} \left\{ \underbrace{(\beta-1)\delta_{r_i,f_i}}_{\text{Rent}} + \underbrace{\delta_{u_{ig},f_i}}_{\text{Amenities}} + \underbrace{\frac{T_g RMA_{ig}^{1/\theta_g}}{\theta_g \bar{y}_{ig}} \delta_{RMA_{ig},f_i}}_{\text{Market Access}} + \underbrace{\frac{\varphi}{\bar{y}_{ig}} \delta_{\varphi,f_i}}_{\text{Land Rents}} \right\} \quad (1.18)$$

Table 1.10: Welfare Decomposition Example

10% increase in MFAR For One Tract					
		Contribution to Welfare Change			
	Total Welfare Change	Rent	Amenities	RMA	Land Rent
Low-Income	0.2769%	31.87%	0.7000%	67.54%	1.34E-03%
High-Income	0.0022%	33.25%	0.8800%	65.99%	4.46E-04%

Notes: Example based on a 10% increase in MFAR for Kings/Brooklyn (FIPS County Code: 47), Dumbo (Tract: 21). Dumbo was chosen due to its accessible location to Manhattan's business districts and pre-zoning attractiveness to both high- and low-income groups. Housing supply in other tracts were kept the same before and after.

These results suggest that the bulk of welfare gains come from the reorganization of workers throughout the city, and their ability to locate closer to higher-paying

jobs. This is mirrored in the counterfactual results and the reorganization of workers closer to CBDs. If true, a direct policy implication would be to focus on rezoning areas that are well-connected transportation hubs and are able to absorb additional housing density.

1.6 Conclusion

While the Bloomberg administration rezonings resulted in aggregate increases in GDP and decreases in rent, the policy did so unevenly across space and across income groups. When selecting sites for changes in zoning policy, policymakers must balance these competing effects and implement concurrent policies to ensure the desired equitable outcome, such as development of better transit, encouraging development of satellite business districts, instituting affordable housing mandates, etc. Moreover, the COVID-19 pandemic has irrevocably changed commuting behavior, as noted in the *New York Times* article in May 2022 titled "N.Y.C. Companies Are Opening Offices Where Their Workers Live: Brooklyn." While the rezonings in the Bloomberg era concentrated on spurring development in places with better access to the traditional Manhattan CBDs, policymakers must now consider the adequacy of housing supply in further parts of the boroughs as commercial activity decentralizes.

Future work may focus on several avenues, such as varying geographic units, using an open city model, and modeling dynamics. In particular, the simulated counterfactuals were completed using Census tracts. Depending on perspective, this geographic unit may be too large or too small to describe neighborhood changes. While the reduced form evidence looked similar across Census block groups and New York City tax lots, finer granularity in the model simulation may yield additional insights into why particular blocks are developed even if the adjoining blocks are not.

Chapter II

The Impact of Trade on Development: Evidence from Pastoralist Practices on the Ancient Silk Road

Joint with Bunyada Laoprapassorn

2.1 Introduction

What is the long-term impact of historical trade on modern development? We approach this question in the context of the overland ancient Silk Road, the ancient trade route across Eurasia. Although the development of long-distance maritime technology in the sixteenth century has rendered the overland Silk Road trade route obsolete, path dependence could result in a persistent effect on economic activity along the ancient Silk Road trade route. Therefore, we seek to examine whether, five centuries after the decline of the overland ancient Silk Road trade, places in close proximity to the ancient Silk Road continue to be more developed relative to places that were not on the Silk Road. In particular, we focus on the long-term impact of the ancient Silk Road in the Inner Asia Mountain Corridor (IAMC), which spans modern-day Afghanistan, China, India, Kyrgyzstan, Kazakhstan, Pakistan, Tajikistan, and Uzbekistan.[1]

[1]We choose to study the IAMC as it contains unique conditions for the existence of nomadic pastoralism, which are keys to our identification strategy. We explain further details about the

Studying the impact of the ancient Silk Road trade on modern development along the IAMC is challenging because of the limited data availability. We overcome the data constraints by utilizing high-resolution satellite imagery. Our main analysis uses the intensity of night lights to proxy for the level of modern development. Remote sensing of nighttime light emissions from the Earth's surface has been widely used as a proxy for economic activity and economic development where conventional measures such as GDP are not available. In addition to its availability on a high spatial resolution, the night lights data allow researchers to circumvent concerns about the manipulation, censoring, and measurement errors in official statistics, which may be non-trivial issues in the context of the geographical area we study. Nevertheless, acknowledging the limitations of night lights data, we also use population density and data on urbanized lands as alternative measures of modern development in our supplement analysis.

To provide evidence of the causal relationship between proximity to the Silk Road and modern development level, we adopt a novel instrument for the locations of the Silk Road. We take advantage of the characteristics of the IAMC and its inhabitants to construct an instrument that provides exogenous variation in the locations of the Silk Road sites. Due to the harsh geographical conditions and the strong pastoralist tradition along the IAMC, highland Silk Road networks emerged in relation to seasonal mobility patterns of the nomadic herders (Frachetti et al., 2017). We use the simulated mobility patterns of the nomadic herders, which are generated based solely on the seasonal pasture quality, to provide variations in the placement of the Silk Road sites that do not arise from ease of travel. Conditional on the location's suitability for growing crops, our instrument only affects modern development level through the ancient Silk Road trade.

We conduct the main analysis in this paper using a grid of 0.167 degrees by

instrument in Section 2.4. Nomadism refers to not having a fixed place of abode, and pastoralism refers to people making their livelihood by herding animals.

0.167 degrees grid cells covering the highland region of the Silk Road. We find a negative and significant relationship between distance to the Silk Road and modern development level; as distance to the Silk Road increases by one standard deviation, the night lights intensity decreases by 10.0%. Based on the elasticity of night lights with respect to GDP in the literature, a 10.0% decrease in night lights intensity corresponds to a decrease in GDP of about 4.1%-9.7%.[2] Our results are robust across different measures of modern development, with a one standard deviation increase in distance to the Silk Road resulting in a 85.5% decrease in population density, and a 12.3% decrease in the percentage of land covered by cultivated land or artificial surfaces.

This paper makes three main contributions to the economics literature. The first contribution of this paper is adding to the sparse literature on how trade routes may have a persistent effect on development, even when the trade routes are no longer relevant to trade patterns in the period of development that is of interest. To do this, this paper also draws on the literature studying path dependence and the impact of durable built infrastructure, such as Rauch (1993), Bleakley and Lin (2012), Jedwab and Moradi (2015), Duranton et al. (2014), and Baum-Snow et al. (2017).

Second, this paper is one of the first studies on the ancient Silk Road using modern econometric techniques and detailed satellite data. Our paper adds to the small but growing economics literature that studies ancient civilizations using limited archaeological records: Garcia-López et al. (2015) study Ancient Roman roads, Dalgaard et al. (2018) also study Ancient Roman roads, Michaels and Rauch (2018) study British and French urbanization from 117 to 2012, Bakker et al. (2019) study the Ancient Phoenicians around ninth century BCE, and Barjamovic et al. (2019) study Assyrian merchants in nineteenth century BCE. Of particular interest is Ahmad and Chicoine (2021), who also study the persistence along the ancient Silk Road. To study the

[2]Literature has found the elasticity of night lights with respect to GDP to be around 0.41-0.97 (Henderson et al., 2012; Hu and Yao, 2021).

long-term consequences of the ancient Silk Road on economic activities, they regress the night lights intensity outcome variable per grid on the presence of the Silk Road in four bins: 0-50km away, 50-100km away, 100-150km away, and 150-200km away, and find there is a persistent and positive association in areas that are within 50km to the ancient Silk Road. While they have conducted exercises to provide suggestive evidence about the long-term impact of the ancient Silk Road on modern economic activities, the nature of the specification in their paper does not allow them to establish a causal relationship. Our study differs from theirs in two main respects: 1) our perspective that the Silk Road is a dynamic network of paths, leading us to avoid using static routes generated from least-cost paths calculation like those in Williams (2014); because of this, 2) we introduce a unique simulated instrument for distance to Silk Road sites to establish a causal link.

This paper is part of a strand of research that uses detailed satellite data, including papers such as Bleakley and Lin (2012), Michalopoulos and Papaioannou (2013), and Dingel et al. (2019) that use night light data to study development intensity. More recent papers have begun to use daytime imagery, often alongside machine learning, to extract features such as farmed land, building density, vegetation type, transportation, and roof types to study the economic development of areas that may otherwise have sparse data or data that is not easily comparable across borders. Such papers include Engstrom et al. (2017) Baragwanath et al. (2019), and Ahmad and Chicoine (2021).

The third contribution of this paper is introducing novel-to-economics mapping simulation techniques to generate a valid instrument for the impact of ancient trade routes on modern development. Lack of direct and accurate information is a major challenge when studying events from antiquity, but depending on the right conditions, a study may be able to use modern simulation techniques to generate patterns of movement that we can plausibly argue to be similar to those experienced in antiquity.

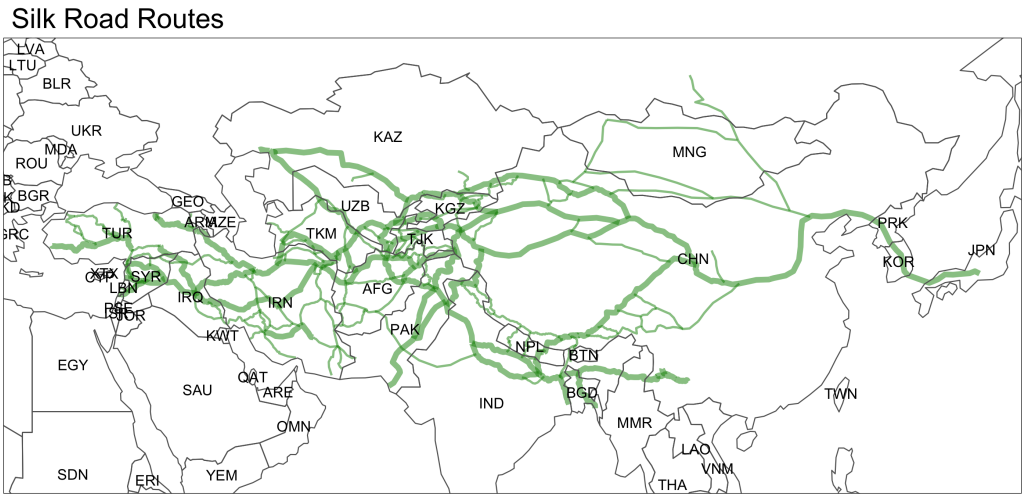
In this paper, we take advantage of the importance of the highlands area to the overland Silk Road trade and the highlands' continuing primary usage by herders and their flocks. Because of this, we are able to use modern satellite imagery to produce herding movement simulations comparable to those experienced in the ancient Silk Road heyday, and then use these calculated flows to instrument for the locations of the Silk Road sites.

Section 2.2 of the paper introduces the relevance to the background of the ancient Silk Road. Section 2.3 describes the data and how we construct data for each unit of observation. Section 2.4 details the empirical strategy, describing the instrument that we use and the various robustness checks we perform. Section 2.5 discusses possible mechanisms through which the ancient Silk Road impacts modern development. Section 2.6 concludes.

2.2 Background

Despite its name, the "ancient Silk Road" does not refer to a physical road, nor is there an unambiguous route that researchers agree to be definitively used during the centuries of overland Silk Road trade. Maps of the ancient Silk Road, such as Figure 2.1 pick important nodes at the researcher's discretion and often deploy least-cost methods based on terrain and elevation to determine where a likely path would lie. More accurately, one may think of the Silk Road as "skein of routes linking many entrepots" connecting East Asia to the Mediterranean and a phrase that represents the dynamic cultural phenomenon that connected Eurasia's people, ideas, and goods together Millward (2013); Spengler III (2019). Thus, when this paper mentions the ancient Silk Road, we are not referring to a physical road in particular, but to the dynamic network of paths identified through known historical sites.

Figure 2.1: Silk Road Routes



Notes: This figure presents hypothetical Silk Road routes compiled in Williams (2014). The thicker lines denote the main corridors; the thinner lines denote the sub-corridors. Williams first chose important nodes such as prominent cities and mountain passes, and then chose pathways considering smaller-scale sites and the ease of travel between segments. The scope of the routes spans from central China to the eastern Mediterranean, but does not address Mongolia and the routes between East Asia (China, Japan, Korea). The routes are also sparse in South Asia.

2.2.1 Definitions

The term Silk Road (in German, "Seidenstrasse") was coined in 1877 by a German geographer, Ferdinand Freiherr von Richthofen. At the time, this term narrowly referred to routes under the Han Empire (206 BCE to 220 CE) which commonly traded Chinese silk (Millward, 2013). However, since this initial usage, the Silk Road has evolved to describe diverse goods and ideas exchanges ranging over diverse geographies and long stretches of time–thousands of years and tens of thousands of kilometers. Travelers on the ancient Silk Road carried precious goods such as textiles, metals, stone, ceramics, perfumes, and horses, disseminated religious beliefs like Buddhism, Islam, Christianity, and technological advancements like paper, gunpowder, calendrical sciences, and medicine. As such, although the ancient Silk Road is often represented as a set of unchanging least-cost paths through important nodes as in Figure 2.1, the ancient Silk Road is more accurately thought of as a dynamic network that connected various geopolitical concerns at different settings in time (Williams, 2014). Therefore, in this paper, unless otherwise stated, the terms "road", "path" and "route" do not refer to an explicit, physical set of roads. Rather, we acknowledge the Silk Road dynamically evolved and shifted with the populations that inhabited the region, a trait that we use in this paper to arrive at an instrument.

This study exploits the mobility patterns of pastoral nomads to form an instrument. Following Wendrich and Barnard (2008) *The Archaeology of Mobility: Definitions and Research Approaches*, "pastoral nomadism" is the "general term for mobility centered on maintenance and welfare of flocks or herds." The word "pastoral" describes the herding, and the word "nomadism" describes the high mobility and impermanence of settlements. It is useful to make the distinction between pastoral nomads and hunter-gatherers, who are often thought to be similar but are distinct (Cribb, 2004). First, hunter-gatherers move towards resources for humans, whereas pastoral nomads move towards resources for flocks, independent of human resources.

Second, hunter-gatherers have typically varied mobility to secure different resources, whereas pastoral nomads are only interested in grazing resources. As such, the former's migration patterns are more complex compared to the latter's, which tend to be more predictable and can be simulated with greater confidence.

Seasonal migration, also referred to as transhumance, has two differentiating terms (Wendrich and Barnard, 2008). Vertical transhumance describes seasonal movements in mountainous areas where the snow in the winter forces the flock to move to the lowlands meadows, but in the warmer weather, the flocks move back to the highland pastures. Horizontal transhumance describes movement at around the same elevation even despite untoward weather. In this paper, following Frachetti et al. (2017), we define highland to be between 750m and 4,000m in elevation.

The area that we study in this paper is referred to as the Inner Asia Mountain Corridor (IAMC). The wide-ranging IAMC stretches from the Hindu Kush mountain range in present-day Afghanistan and Pakistan to the Altai Mountains in Siberia. The IAMC also spans the modern-day countries of China, India, Kyrgyzstan, Kazakhstan, Tajikistan, and Uzbekistan. Before the development of reliable, long-distance maritime routes in the sixteenth century, any trade between East Asia and the Mediterranean had to pass through the long overland routes in the IAMC. The IAMC's varied geography of deserts, steppes, and tall mountain ranges coupled with fertile valleys, oases, and inland deltas gave rise to both vibrant nomadic pastoralist and sedentary agricultural traditions (Frachetti et al., 2017). In the IAMC, vertical transhumance is the dominant seasonal migration pattern and, thus, the main focus of this paper. We use the small-scale, dynamic vertical transhumance routes of nomadic pastoralists on the highland steppes to study the impact of the ancient Silk Road on modern development.

In summary, using terms as developed by previous scholars across varied disciplines, we use the pastoral nomads' patterns of vertical transhumance in the Inner

Asia Mountain Corridor to establish a causal relationship between the ancient Silk Road and modern development.

2.2.2 History

Silk Road in Six Periods

Evidence of long-distance trade in Eurasia stretches as far back as the fourth millennium BCE, the period which ushered in the Bronze Age and in which writing was invented. One can divide the ancient Silk Road's history into roughly six periods (Millward, 2013). The first period, which spans from c. 3000 BCE to c. 300 BCE, is characterized by the expansion of farmers and herders out onto the steppe geographies in Central Asia by taking advantage of the new wheel and wagon technologies discovered c. 3500 BCE. These populations became the first nomadic pastoralists on the steppe.

The second period, lasting from c. 300 BCE to 300 CE, may be thought of as the "Classical Silk Road" period in which the broad area between the Mediterranean and China fell under the centralized control of a few empires. Zhang Qian, a government minister, was dispatched by the Han dynasty in 139 BCE to serve as an imperial envoy to Central Asia. Qian's trip has often been credited with stimulating the opening of different empires and Central Asia to transcontinental trade.

The third to fifth centuries CE was a "dark age" in some sense with the collapse of the western Roman Empire and the fall of the Han empire. However, the Silk Road trade still flourished, with Persians and other Central Asians taking control of the flows.

The fourth and fifth periods of the Silk Road occurred from the sixth to fifteenth centuries CE. They are characterized by the continued expansion of Persians and Arabs in Central Asia and the reunification of China again by the Sui and Tang dynasties and their cosmopolitan tastes. This is then followed by the domination of

the Mongol empire over Eurasia.

The last period began in the sixteenth century CE. This period had two major developments: improvement in long-distance maritime technology such that ships were more efficient than overland caravans, and increased security of traveling via the Black Sea. These two developments led to the decline of the overland Silk Road, leading to a simultaneous decline of nomad-steppe culture in Central Asia (Williams, 2014; Millward, 2013).

Nomadic Pastoralists on the Silk Road

Early scholarship on the ancient Silk Road in Central Asia focused on lowland oases where the agricultural tradition is strongest, because they are still heavily populated today and more easily accessible to study. However, with more archaeological work in more remote locations, historians in the past decades have identified more Silk Road sites in the highland mountains. In doing so, they continue to reaffirm the hypothesis that the nomadic pastoralists contributed significantly to the development of the highland Silk Road geography (Millward, 2013). Generally, pastoralists in the IAMC would herd their livestock in the mountainous regions to feast on the productive grassland in the summer months and retreat to the lowland oases to weather the winter months. As the grazing herd moved to the productive grassland through the mountainous regions, mobile pastoralists followed. Evidence suggests that these small-scale mobility patterns and subsequent paths formed influenced the formation of the macro-scale Silk Road network (Frachetti et al., 2017). Trade was most likely conducted in short stages: sold at one node, transported to the next node by following the paths influenced by the nomadic pastoralists, resold at that node, and so forth (Gorbunova, 1993).

2.3 Data

2.3.1 Data Sources

Examining the relationship between the ancient Silk Road trade and modern level of development requires data on the locations of the Silk Road and fine-level geospatial data on modern development and other characteristics of the areas. We outline below the primary datasets that we use.

Silk Road Sites: Data on locations of Silk Road sites come from Old World Trade Routes (OWTRAD) Project (Ciolek, 2014), a public-access aggregator of geo-referenced and/or chrono-referenced data of nodes between Eurasia and Africa ranging from 4,000 BCE to 1820 CE, and from Williams (2014), a study of the Silk Road sites done on behalf of International Council of Monuments and Sites. Compiled by historians and archaeologists, the OWTRAD Gazetteer and the ICOMOS study include locations such as current and past settlements, oases, temples, rest houses, markets, forts, river and mountain crossings. As of writing, these databases combined had over 16,000 entries. We follow Frachetti et al. (2017) and extract 258 sites by choosing only sites that are on the IAMC, collapsing duplicate observations which differ only in name, and removing sites that were entered solely to facilitate the generation of a Silk Road path. We further subset to include only the 254 sites situated between 750m and 4,000m in elevation.

Table 2.1 summarizes the final count of the Silk Road sites by country and node types. Figure 2.2 maps the locations of the Silk Road sites in the IAMC highlands.

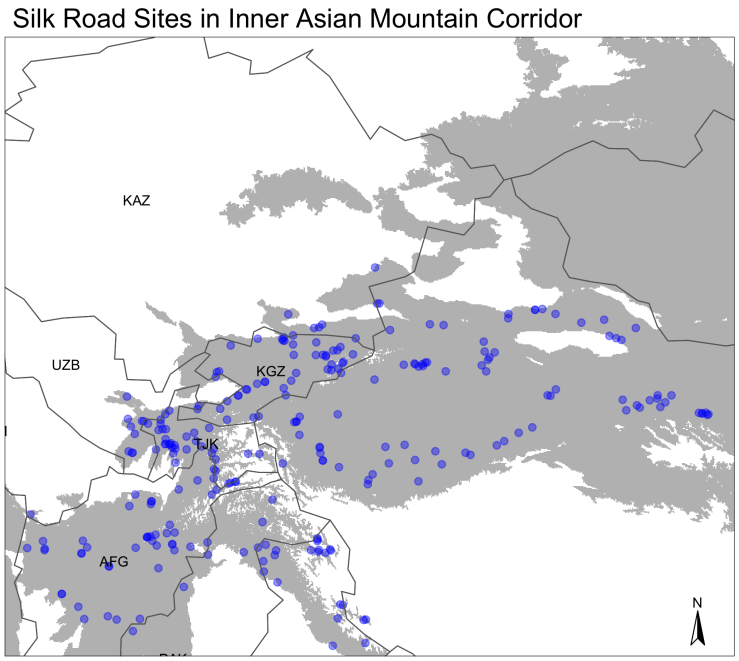
Nighttime Lights: Measuring development at fine spatial scales and across countries using existing surveys is a challenging task. Thus, we use high-resolution satellite imagery instead to proxy economic development. First, we use nighttime lights data from two sources: 1) the Version 1 Nighttime VIIRS Day/Night Band (VIIRS-DNB) Composites for the year 2016 from the Earth Observation Group (EOG) at NOAA/NCEI, and 2) the Satellite 18 Year 2013 Version 4 DMSP-OLS Nighttime Lights Time Series.

Table 2.1: Count of Silk Road Sites by Country and Node Types

	AF	AF/CN	AM	CN	IN	KG	KG/CN	KZ	PK	TJ	TR	UZ	Total
Current Settlement	18	0	1	41	7	17	0	3	0	29	2	0	118
Former Settlement	0	0	0	7	0	1	0	0	0	1	0	0	9
Gorge	0	0	0	1	0	1	0	0	0	0	0	0	2
Halting Place	1	0	0	0	7	0	0	0	0	0	0	0	8
Land Waypoint	0	0	0	0	0	1	0	0	0	0	0	0	1
NA	21	0	0	33	8	15	0	4	6	11	1	6	105
Pass	1	1	0	1	0	2	1	1	1	1	0	0	9
Stream	0	0	0	0	0	1	0	0	0	1	0	0	2
Total	41	1	1	83	22	38	1	8	7	43	3	6	254

Notes: This table summarizes the final count of the Silk Road sites by country and node types. Data is from ICOMOS/OWTRAD and Frachetti et al. (2017).

Figure 2.2: Silk Road Sites in Inner Asia Mountain Corridor



Notes: This figure presents Silk Road sites in the Inner Asia Mountain Corridor, as compiled by Old World Trade Routes (OWTRAD) Project (Ciolek, 2014) and Frachetti et al. (2017). The gray area represents the area of interest, i.e., the landmass of the IAMC that falls between 750m and 4,000m in elevation.

.

The DMSP-OLS dataset provides the annual cloud-free composites of the night lights intensity, available in 30 arc-second geographic grid cells. We use the "stable_lights.avg_vis" product, which filters out ephemeral events, such as fires. The data values range from 0 to 63. Meanwhile, the VIIRS-DNB dataset provides the annual composites of the night lights intensity, produced in 15 arc-second geographic grids. We use the "vcm-orm-ntl" product, which measures the cloud-free average radiance values, removing outliers to filter out the fire and ephemeral lights. The data contains the radiance values with units in nano Watts per square centimeter per steradian ($nanoWatts/cm^2/sr$). The VIIRS-DNB is considered to be superior to its predecessor DMSP-OLS, both in terms of spatial precision and low-light detection capabilities. We therefore consider the VIIRS-DNB dataset to be our preferred measurement.

The limitations with night lights data are well-known (Lowe, 2014). Light spills over into adjacent grid cells, some lights are a result of gas flares instead of economic activity, and the light measurements are usually not sensitive enough to pick up on less dense activities, such as agriculture. Because of the limitation of night lights data, we supplement our measure of modern-day development with two additional measures of development, as detailed below.

NDVI: We adopt the normalized difference vegetation index (NDVI) and the algorithm to simulate the seasonal mobility pattern of nomadic herders, both of which will be used for the construction of our instrumental variable, from Frachetti et al. (2017). The NDVI data used is the 7-day average NDVI values derived from multispectral eMODIS satellite imagery captured in August 2008. Given the Silk Road's history stretching back millennia, concerns arise from using modern satellite imagery to model herding paths over this long period. Using pristine ice core samples, modern paleoclimate models are able to capture broad changes in vegetation and weather over time. For the IAMC, these models generally find that there was a transition to

wetter climates from the first century to 250 BCE, after which the region's climate and vegetation was relatively stable (Hill, 2019; Aubekerov et al., 2003; Khotinskiy, 1984). These findings from the model correspond to current archaeological records of pollen and sediment composition (Frachetti, 2008), giving more confidence that modern satellite imagery provides a reasonable base for modeling herding patterns in Central Asia along the Ancient Silk Road over the past two millennia. We provide further details about the instrument in Section 2.4.3.

Other Development Measures: We first supplement our measure of modern-day development with data on population density, which we obtain from the Gridded Population of the World (GPWv4) provided by the NASA Socioeconomic Data and Applications Center (SEDAC). GPWv4 uses data from the 2010 round of Population and Housing Census to model the distribution of the human population at a spatial resolution of 30 arc-seconds. We use the estimate of the population density for the year 2015, which provides the number of persons per square kilometer.

In addition to nighttime lights and population density data, we use land cover data from GlobeLand30-2020 as an alternative measure of modern development. GlobeLand30's classification system consists of ten land cover types, namely cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bareland, and permanent snow and ice. Data is available at 30-m resolution.

Other Datasets: We supplement our data with other high-spatial-resolution data for our controls. The data for the controls are standard in the literature. We use the caloric suitability index from Galor and Özak (2016), which calculates the potential agricultural output based on crops that were available for cultivation in the time before 1500 CE. We also use the terrain ruggedness measure developed by Nunn and Puga (2012) as well since terrain ruggedness could affect how suitable an area is for settlement. Lastly, we use AquaMaps for yearly precipitation and shapefiles of major rivers. We also add a selection of variables from the Global Agro-

ecological Zones v3.0 (IIASA/FAO, 2012), such as the crop suitability index and total production capacity in terms of tons per hectares for barley, flax, foxtail-millet, pearl-millet, rice and wheat.

Finally, we have a few other spatial datasets that we use to test for mechanisms. We use the Global Map of Irrigation Areas from FAO's AQUASTAT, which reports the percentage of irrigated land per cell size of 5 minutes worldwide as of 2005. Additionally, we use the 2015 Global Exposure Database for GAR provided by the United Nations Office for the Coordination of Humanitarian Affairs. This dataset is primarily used to assess damage from disasters and includes estimates of various exposed capital stock worldwide at 1km spatial resolution. It includes the value of the capital stock, separated into categories such as housing, education, and health.

2.3.2 Data Construction

Since we are interested in the highland Silk Road along the Inner Asian Mountain Corridor, we limit the geographic extent of the study zone to 30 degrees to 55 degrees latitude and 60 degrees to 100 degrees longitude.[3] Additionally, we follow Frachetti et al. (2017) in limiting the study zone between the elevations of 750 m and 4,000 m. Within the region of interest, we construct a regular 0.167×0.167 degrees grid, which corresponds to approximately 19×19 kilometers. We treat each grid cell as an observation. We construct the variables as follows.

Shortest Distance to Silk Road Site: We define a cell as a Silk Road cell if it contains at least one Silk Road site. We capture the distance from each grid cell to the Silk Road sites by measuring the distance[4] between the cell's centroid to the centroid of the nearest Silk Road cell. We call this the shortest distance from the cell to the Silk Road site.[5] Figure 2.3 plots the shortest distance measure that we

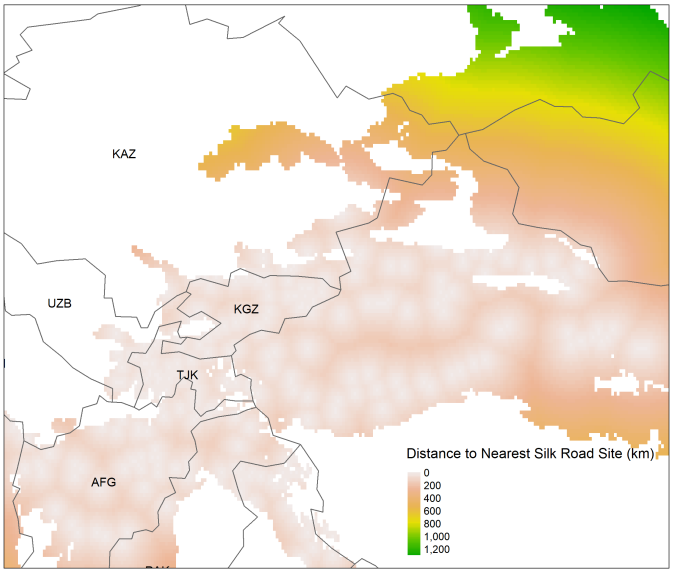
[3]The range of interest for our study area is dictated by our instrument.

[4]We use the shortest distance along the earth's surface, also known as"as the crow flies" distance.

[5]We also construct an alternative measure as a robustness check. See Appendix B.2 for details.

construct.

Figure 2.3: Shortest Distance to Silk Road site



Notes: This figure displays the shortest distance between the centroid of each grid cell to the centroid of the nearest Silk Road cell.

Our choice to define the Silk Road based on the Silk Road sites rather than the Silk Road routes as established by historians is intentional. Despite the name "Silk Road," there were no physical roads that dictated specific routes of the Silk Road. Traditionally, historians established Silk Road routes by calculating the least-cost paths between pre-identified historical Silk Road sites. By construction, the Silk Road routes as established by historians are inherently dictated by ease of travel, which is endogenous to modern development. In this respect, defining the Silk Road based on the Silk Road routes is problematic for the purpose of this paper. Therefore, we define the Silk Road locations based on the Silk Road sites without imposing any routes between them.

Measure of Modern Level of Development: Nighttime lights data and population density data consist of continuous values. As such, we compute the variable

value from those datasets by averaging the values across all pixels in a grid cell. The land cover dataset consists of categorical values and cannot be treated in the same manner. For each grid cell, we compute the percentage of the area that is classified as artificial surfaces, which are defined as "lands modified by human activities, including all kinds of habitation, industrial and mining area, transportation facilities, and interior urban green zones and water bodies, etc." A higher percentage of artificial surfaces signifies a higher level of urbanization and thus a higher level of development. Additionally, to account for the possibility that human activities in the region may consist primarily of agriculture, we construct an alternative variable that measures the percentage of the area classified as cultivated land or artificial surfaces.

Table 2.2 presents the summary statistics of the development measures at the 0.167×0.167 degrees grid cells level. Given the presence of grid cells with zero value, we use the inverse-hyperbolic-sine (IHS) transformation before taking log in our subsequent analyses. This approach is consistent with what has been adopted in the literature (Gibson et al., 2021; Bruederle and Hodler, 2018).

Other Variables: All other non-distance variables are computed by averaging the values across all pixels in each grid cell. We construct all distance-related variables in the same manner as how we measure the shortest distance to the Silk Road. Variables containing zero are transformed using IHS transformation before taking log in the subsequent analyses.

2.4 Specification and Results

The relationship between ancient Silk Road trade and modern development can be represented by the following regression:

$$Y_i = \beta_0 + \beta_1 DistancetoSilkRoad_i + \beta_2' Z_i + \gamma_i + \varepsilon_i \tag{2.1}$$

Table 2.2: Summary Statistics of Development Measures

	Mean	Std. Dev.	Min	Max	Skewness	Kurtosis
VIIRS	0.04	0.57	0.00	40.10	46.95	2,703.67
DMPS	0.38	2.04	0.00	62.92	13.03	262.54
Population Density	27.76	134.70	0.00	7,423.69	25.13	980.31
Land cover (artificial surfaces)	0.00	0.02	0.00	0.92	19.58	565.77
Land cover (both)	0.06	0.15	0.00	1.00	3.71	17.56

Notes: This table reports the summary statistics of measures of development measures. One unit of observation is a 0.167 degrees by 0.167 degrees grid cell. Population density refers to the number of persons per square kilometer. Land cover (artificial surfaces) refers to the percentage of a grid cell that is covered by lands modified by human activities. Land cover (both) refers to the percentage of a grid cell that is classified as cultivated land or artificial surfaces.

where Y_i is a measure of modern development of grid cell i, $DistancetoSilkRoad_i$ is the shortest distance from grid cell i to the ancient Silk Road site, Z_i represents other control variables, and γ_i represents the country fixed effects. We control for the geographical features that can influence both the placement of Silk Road sites and modern development. Specifically, we control for the potential agricultural output, terrain ruggedness, distance to the nearest river, and elevation.

Analyzing the standard errors of the coefficients from our regressions is particularly important given that we are working with spatial data. Below, we first describe how we calculate the standard errors before presenting the results from our OLS regressions and introducing our instrumental variable approach.

2.4.1 Standard Errors

The First Law of Geography, written by Tobler (1970), states that "everything is related to everything else, but near things are more related than distant things." Spatial data is autocorrelated and thus must be taken into account in the empirical strategy. Unlike the literature that addresses standard errors and autocorrelation through time, the literature that addresses standard errors in the context of spatial autocorrelation has less consensus on what methods are best to minimize the influence of correlated noise on the results.

The common starting point is Conley (1999b)'s paper, which outlines a process for generating spatial Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors. Conley (1999b)'s method is successful in curbing estimated t statistics, relative to robust standard errors, but is sensitive to the choice of kernel, bandwidth, and the unique spatial pattern of the data (Kelly, 2019). Related research has shown that the choice of kernel is less important relative to the choice of bandwidth. This is because asymptotically, different kernel densities tend to converge to similar results (Cameron and Trivedi, 2005). Thus, we focus on choosing a correct bandwidth, and

then perform robustness checks on our regressions to confirm that the standard errors are of the correct magnitude.

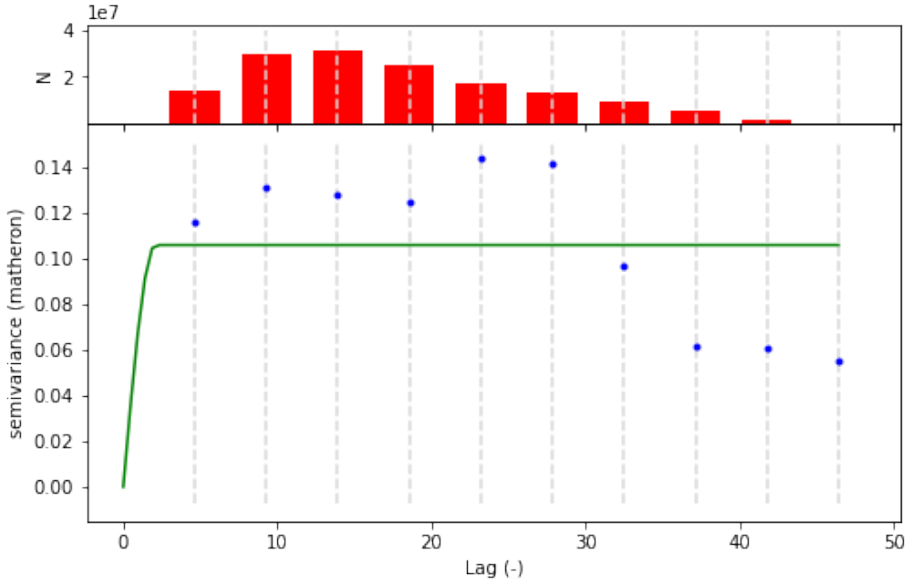
There are three methods we consider to choose an optimal bandwidth. The first method is a simple rule-of-thumb proposed by (Kelejian and Prucha, 2007). They suggest choosing a bandwidth b based on the following formula $b = n^{2\tau}$ where n is the number of observations, and $\tau \leq 1/3$. In our study, an observation is a grid cell, the size of which may be adjusted. Next, Lambert et al. (2008) suggest a data-driven cross-validation method performing the following minimization: $\min_b \sum_{i=1}^{n}(y_i - \hat{y}_{-i}(b))^2$ where $\hat{y}_i$ is the fitted value of y_i with location i omitted during the fitting process. Lambert et al. (2008) also suggest another data-driven approach of fitting a variogram to the data and setting the bandwidth as the generated range r, which is when the distance at which points are not significantly spatially correlated. Due to its straightforwardness and ease of interpretation, we use the last approach and fit a variogram (Figure 2.4), leading us to choose a bandwidth of 2 degrees.

2.4.2 OLS Approach

Preliminary results from the OLS regressions are shown in Table 2.3 and Table 2.4. Table 2.3 shows the results when we use the VIIRS night lights as a proxy for modern level of development. Distance to Silk Road site is standardized and all other variables are log-transformed. The first column shows the relationship between night lights and distance to the nearest Silk Road site, controlling for only the country fixed effects and no additional control variables. There is a significant and negative relationship; as we increase the distance to the Silk Road site by one standard deviation, the night lights level falls by 2.3% on average.

The second column shows the relationship when we control for other factors that may affect modern level of development. We control the caloric suitability index, the

Figure 2.4: Variogram



Notes: This figure presents the variogram that we generate to determine the bandwidth used to calculate spatial HAC standard errors. A variogram is a function that describes the degree of spatial dependence (y-axis) of a spatial field as the distance between two spatial units increases (x-axis). A variogram has three important parameters: the nugget n, the height of the variogram at the origin; the sill s (0), the spatial dependence value at which the curve begins to flatten (0.11); and the range r, the distance at which the curve begins to flatten (2.07). The range determines the bandwidth we will use when calculating the spatial HAC errors.

Table 2.3: OLS Results

	(1)	(2)	(3)	(4)
	Night lights (VIIRS)			
Distance to Silk Road site	-0.023***	-0.036***	-0.034***	-0.032***
	(0.006)	(0.008)	(0.008)	(0.008)
Caloric suitability index		-0.001	-0.001	-0.001
		(0.001)	(0.002)	(0.002)
Ruggedness		-0.010***	-0.015***	-0.013***
		(0.004)	(0.003)	(0.003)
Precipitation		0.040***	0.011	0.016*
		(0.008)	(0.008)	(0.009)
Distance to river		-0.007***	-0.005**	-0.004**
		(0.002)	(0.002)	(0.002)
Elevation		-0.051***	-0.029**	-0.042***
		(0.012)	(0.011)	(0.014)
Crop-Suitability-Indices	No	No	Yes	Yes
NDVI	No	No	Yes	Yes
Latitude and longitude	No	No	No	Yes
Country FE	Yes	Yes	Yes	Yes
N	17,675	17,675	17,125	17,125

Notes: This table reports results from the OLS regression following equation (2.1). Distance to Silk Road site is standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ruggedness of the terrain, the amount of precipitation, the distance to the nearest river, and the elevation. In the third column, we control for land fertility more directly using the crop-suitability-index of major crops from the FAO GAEZ dataset. Specifically, we control for the crop-suitability-index of wheat, rice, barley, flax, and foxtail millet. Additionally, we control for the NDVI value of each grid cell. The controls used in this column constitute our baseline specification. Kelly (2020) recommends accounting for the geographical location of the observations when studying the relationship between modern outcomes and the historical characteristics of the places in the past. In the fourth column, we control for the latitude and the longitude of our observation in addition to the controls in our baseline specification. Generally, we can see that adding in controls does not significantly affect our coefficient of interest; the relationship between the night lights and the distance to the Silk Road sites is always negative and significant at 1%.

The majority of existing literature estimates the elasticity between night lights and GDP using the DMSP-OLS data. Henderson et al. (2012) estimate the elasticity using country-level data and find that the structural effect of true income growth on lights growth varies between 1.03-1.72, implying the elasticity of GDP with respect to night lights between 0.58-0.97. Similarly, Hu and Yao (2021) find the cross-country estimate to be 0.76.[6] Among the limited literature that studies the VIIRS-DNB data, Chor and Li (2021) estimate the elasticity of GDP per capita with respect to night lights intensity using prefecture-level in China. They find that a 1% increase in night lights intensity corresponds to a 0.41-0.47% increase in GDP per capita.

Table 2.4 shows the results when we use alternative measures of modern development under our baseline specification. The negative relationship between distance to the Silk Road sites and modern level of development is significant across all measures. The first column shows that the coefficient of the distance to the Silk Road

[6]The paper finds that a one percentage point increase in GDP growth increases night lights growth by 1.317 percentage points, which implies an inverse elasticity of $1/1.317 = 0.76$.

Table 2.4: OLS with Alternative Measures of Development

	(1)	(2)	(3)	(4)
	Night lights (DMPS)	Population density	Land cover (artificial surfaces)	Land cover (both)
Distance to Silk Road site	-0.157***	-0.689***	-0.004***	-0.064***
	(0.028)	(0.099)	(0.001)	(0.008)
Caloric suitability index	-0.013*	0.055**	-0.000	-0.007***
	(0.007)	(0.027)	(0.000)	(0.002)
Ruggedness	-0.068***	-0.048	-0.002***	-0.043***
	(0.011)	(0.047)	(0.000)	(0.005)
Precipitation	0.042	0.420***	0.001	0.019**
	(0.029)	(0.112)	(0.001)	(0.008)
Distance to river	-0.028***	-0.009	-0.000	-0.005***
	(0.006)	(0.018)	(0.000)	(0.002)
Elevation	-0.142***	-0.522***	-0.002	-0.017*
	(0.041)	(0.136)	(0.001)	(0.009)
Crop-Suitability-Index	Yes	Yes	Yes	Yes
NDVI	Yes	Yes	Yes	Yes
Latitude and longitude	No	No	No	No
Country FE	Yes	Yes	Yes	Yes
N	17,125	17,125	17,125	17,125

Notes: This table reports results from the OLS regression following equation (2.1) with alternative development measures as dependent variable. Distance to Silk Road site is standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

site is larger in magnitude when we use the DMSP night lights measure instead of the VIIRS night lights data. This is somewhat expected since the DMSP data has limited capability to detect lights at low radiance levels; therefore, the DMSP may have underestimated the level of night lights in areas with a lower level of economic activity, leading to an upward bias in the estimate. The second column shows the result when we use modern level of population to measure modern outcome. On average, a one standard deviation increase in distance to the Silk Road sites corresponds to a 49.8% decrease in population density.[7] The third and the fourth columns show the results when using data from the land cover data set as our dependent variable. In column 3, modern development is measured by the percentage of land that is covered by artificial surfaces, which are defined as areas that have artificial cover resulting from human activities such as construction. In column 4, the dependent variable is the percentage of land classified as artificial surfaces or cultivated land. On average, as the distance to the nearest Silk Road site increases by one standard deviation, the percentage of area covered by artificial surfaces decreases by 0.4%, and the percentage of area covered by artificial surfaces or cultivated land decreases by 6.2%.

2.4.3 Instrumental Variable Approach

Although regression (2.1) shows that proximity to Silk Road site corresponds to a higher level of modern development, one can still be concerned about other unobserved heterogeneity. To establish a causal relationship between the ancient Silk Road trade and modern development, we adopt an instrumental variable strategy. A valid instrument is a variable that, after controlling for other controls in the regression, i) strongly correlates with the proximity to the Silk Road sites, and ii) only affects modern development level through the Silk Road.

[7]We find a stronger effect when we proxy for development using population density, which is in line with what has been found in the literature. For example, Bleakley and Lin (2012) found that a 10% increase in distance away from a portage predicts a 6% lower population density and 2% lower night lights intensity.

2.4.3.1 Instrument: Seasonal Mobility Patterns of Nomadic Herders

To provide exogenous variation for the location of the Silk Road sites, we use two unique features of the Inner Asia Mountain Corridor and its inhabitants. First, ancient Silk Road traders wishing to travel between East Asia and the Mediterranean were obligated to use the overland routes in the IAMC. In the lowland regions, which were flatter and fertile with strong agricultural traditions, trading routes were generally selected based on ease of travel for traders. In the highland regions, because of the sparsely populated and often barren terrain in the highland area, the group of people who had generally traversed the area was the nomadic herders; subsequently, traders crossing the highland region followed the paths taken by the nomadic herders.

Second, unlike the traders, the nomadic herders' objective for moving overland was not to find the least-cost path of traversing the highland region to get to the lowland destination, but to exploit variation in the pasture quality as they traversed the highlands. The quality of pastures in the lowland and highland regions changes substantially over the seasons. In the summer months, lowland pastures become too arid while highland pastures become more productive. For over 4,500 years, nomadic herders have exploited these seasonal variations in pasture quality, moving from the lowland areas to the highland areas in the summer and returning to the lowland regions in the winter. Frachetti et al. (2017) find that the seasonal mobility patterns of nomadic herders shaped the highland Silk Road networks. By following the footsteps of the nomadic herders, the traders were not taking the optimal path they should have taken if they simply wanted to cross the highland region to get to the lowland destinations.

Given that the highland Silk Road networks are shaped by the nomadic herders' seasonal mobility patterns and that the nomadic herders' movements are not dictated by the least-cost path of traversing the IAMC, the seasonal mobility patterns of the nomadic pastoralist provide variations in the locations of the highland Silk Road that

are not dictated by the ease of travel. We therefore instrument for the locations of the Silk Road sites using the seasonal herding patterns. The exclusion restriction is that after controlling for the suitability of the area for growing agricultural crops, the seasonal herding patterns of the nomadic pastoralist only affect modern development through the Silk Road trade.

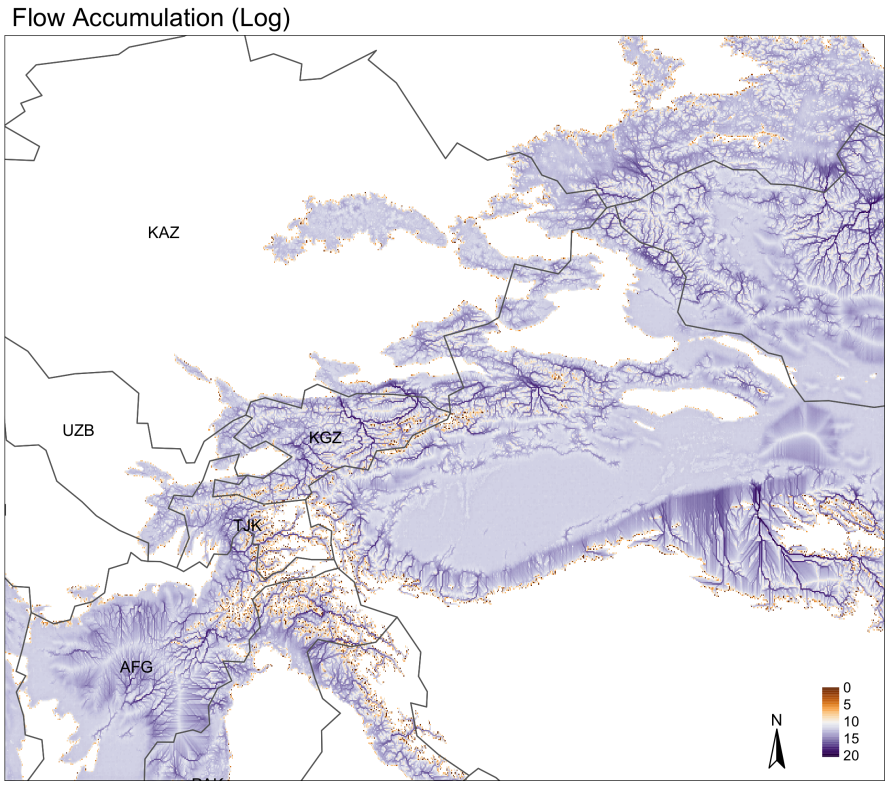
Actual historical herding patterns of nomadic pastoralists are unrecorded. We therefore proxy for the nomadic herders' herding paths using the simulated seasonal herding patterns from Frachetti et al. (2017). Frachetti et al. (2017) uses flow accumulation modelling[8] to simulate the aggregated seasonal movements of herd animals across the IAMC using classified grass fodder quality as the input. The model simulates a figurative count of 'animals' flowing from one cell to another across the IAMC based on the pasture quality, which is categorized based on the NDVI values over 20 human generations. The result is a flow accumulation measure that represents the simulated herding patterns. For more detail about the construction of the flow accumulation measure, please refer to Appendix B.1.1 or Frachetti et al. (2017) directly. Figure 2.5 provides a visual representation of the distribution of simulated flow accumulation values across the IAMC. The resulting paths from these simulated flow accumulation values can be thought of as the herding pattern of the nomads.

The map of flow accumulation values is continuous. For computational feasibility, we must choose a cutoff for flow accumulation to isolate a pathway before constructing a distance to the flow accumulation path. We choose 1.5 million as a cutoff. Figure 2.6 shows the path generated. We later vary this cutoff to check the robustness of our results. After implementing this cutoff, the method to calculate the distance to the flow accumulation path is identical to the previous calculation of the distance to Silk Road site.

We argue that the simulated seasonal nomadic herding patterns satisfy the exclu-

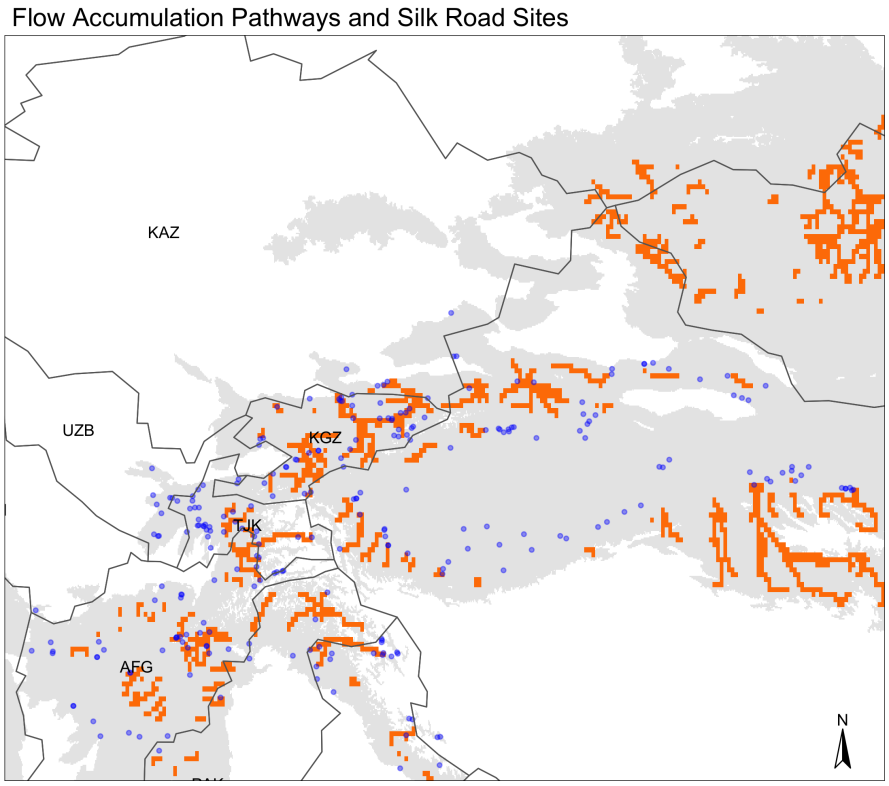
[8]Named the 'pastoralist participation' model.

Figure 2.5: Simulated Herding Path



Notes: This figure presents the log flow accumulation values from the simulated season mobility pattern of nomadic herders in the Inner Asia Mountain Corridor as initially introduced and calculated by Frachetti et al. (2017).

Figure 2.6: Simulated Herding Path and Silk Road Sites



Notes: This figure presents the simulated herding paths generated using flow accumulation values of 1.5 million as a cutoff. The gray area represents the area of interest, i.e., landmass of the IAMC that falls between 750m and 4,000m in elevation.

sion for the following reasons. First of all, nomadic mobility in the highland regions is primarily dictated by seasonal pasture quality rather than the 'ease of travel' (Frachetti et al., 2017). In fact, by construction, the simulated seasonal nomadic mobility is dictated by pasture quality and is not affected by typical factors that affect the ease of travel, such as slope. Secondly, the seasonal quality of pasture for grazing, which matters for the herders, does not translate to the productivity of land for farming, which is what matters for settlements. The highland region we are interested in is generally very rugged, making the entire area unsuitable for cultivation. As noted by Nunn and Puga (2012), cultivation becomes impossible when slopes are greater than 6; out of 36,000 cells in our region of interest, only 14 cells have a slope measure of less than 6. Therefore, the seasonal pasture quality that drives the nomadic herding patterns does not dictate the modern level of development of a certain location. Likewise, to further address concerns that pasture quality may correlate with the suitability of land for cultivation, we directly control for the suitability of land for various staple crops in our regressions using the FAO crop-suitability-indices. We also further directly control for the NDVI values in each grid cell. Thirdly, although nomadic herders move between lowland to highland areas over the year, they do not 'settle' permanently. Evidence of the impermanence and seasonality of herding patterns comes from rock art in the area. As Frachetti (2008) writes, "There is little rock art in the upland areas, even though usable rock faces exist there as well. Since the upland pastures are treacherously cold and uninhabited during the winter, it is unlikely that there would be significant human traffic there except during the summer. Logically, there would be no need to protect or mark settlement areas in the highlands." For the pastoral nomads, permanent developments in the highlands were and continue to be impractical due to the seasonal weather; because of this, their activities should not be directly affecting development levels across the highlands. Settlements, which matter for modern level of development, should only take place

as a result of trade through the Silk Road.

2.4.3.2 Instrumental Variable Results

Using the constructed distance to the flow pathways measure as an instrument, we run the following first-stage regression:

$$SilkRoad_i = \alpha_0 + \alpha_1 Flow_i + \alpha_2' Z_i + \lambda_i + \eta_i \tag{2.2}$$

where $Flow_i$ is the distance to the flow pathways measure that we constructed. We then use the predicted value $\widehat{SilkRoad}_i$ in the second-stage regression:

$$Y_i = \beta_0 + \beta_1 \widehat{SilkRoad}_i + \beta_2' Z_i + \gamma_i + \varepsilon_i \tag{2.3}$$

Table 2.5 shows the results from our IV specification. The first column shows the first stage result. The distance to the herding path positively correlates with the distance to the Silk Road sites, which is what we expect given Frachetti et al. (2017)'s finding that the seasonal mobility patterns of nomadic herders shape the highland Silk Road networks. The F-statistic from our first stage is around 31.6, which is in the acceptable range for a strong instrument.

Columns (2)-(6) in Table 2.5 present the second stage results. The IV estimates are larger in magnitude than the OLS estimates, which can be explained by measurement errors in the locations of the Silk Road sites. From the IV regression, a one standard deviation increase in distance from the Silk Road site results in a 10.0% decrease in night lights intensity using the VIIRS-DNB data and 38.4% decrease in night lights intensity using the DMSP-OLS data.[9] Columns (4)-(6) present the IV results when we

[9]Our result is smaller than that in Ahmad and Chicoine (2021), who found that night lights intensity is about 80% higher in cells within 50 km of the Silk Roads relative to cells that are 200-500 km away from the Silk Roads using DMSP-OLS data. This difference is somewhat expected since the area of study in Ahmad and Chicoine (2021) covers the lowland region of the Silk Road, where the routes are believed to be dictated by ease of travel and pass through big cities. Therefore, one

Table 2.5: IV Results

	(1)	(2)	(3)	(4)	(5)	(6)
	First stage	Second stage				
	Distance to Silk Road site	Night lights (VIIRS)	Night lights (DMPS)	Population density	Land cover (artificial surfaces)	Land cover (both)
Distance to Silk Road site		-0.105***	-0.485***	-1.930***	-0.014***	-0.131***
		(0.029)	(0.107)	(0.365)	(0.004)	(0.025)
Distance to herding path	0.190***					
	(0.034)					
Caloric suitability index	-0.041***	-0.004**	-0.028***	-0.003	-0.001**	-0.010***
	(0.014)	(0.002)	(0.009)	(0.034)	(0.000)	(0.003)
Ruggedness	-0.048*	-0.018***	-0.083***	-0.103*	-0.003***	-0.046***
	(0.026)	(0.004)	(0.015)	(0.060)	(0.001)	(0.005)
Precipitation	0.255***	0.029**	0.125***	0.737***	0.003**	0.036***
	(0.047)	(0.011)	(0.040)	(0.138)	(0.001)	(0.010)
Distance to river	-0.026***	-0.005***	-0.031***	-0.020	-0.000*	-0.006***
	(0.007)	(0.002)	(0.007)	(0.020)	(0.000)	(0.002)
Elevation	0.021	-0.042***	-0.199***	-0.737***	-0.004*	-0.028**
	(0.106)	(0.016)	(0.060)	(0.213)	(0.002)	(0.012)
Crop-Suitability-Indices	Yes	Yes	Yes	Yes	Yes	Yes
NDVI	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	31.590					
N	17,125	17,125	17,125	17,125	17,125	17,125

Notes: This table reports results from the IV regressions. Distance to Silk Road site and distance to herding path are standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

use alternative measures of modern development. The population density decreases by 85.5% with a one standard deviation increase in distance to the Silk Road site, while the percentage of area covered by artificial surfaces and the percentage of area covered by both artificial surfaces and cultivated land decrease by 1.4% and 12.3% respectively.

2.4.3.3 Robustness Checks

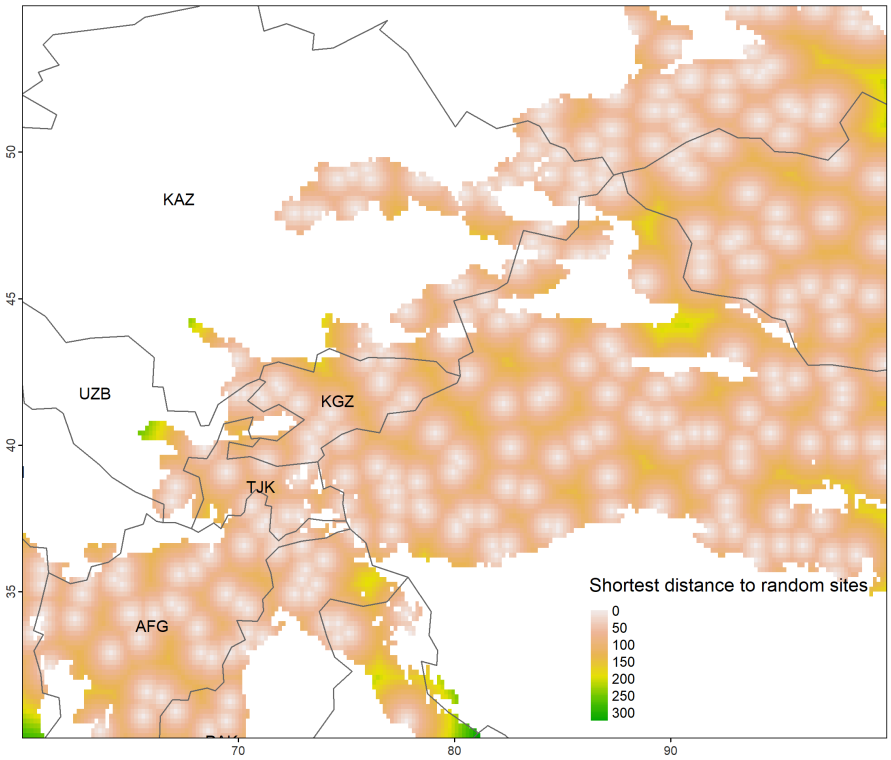
Random Sites: One could be concerned that the results in our analyses are driven by spatial noises. To ensure that this is not the case, we repeat our analyses using random sites instead of the Silk Road sites. Specifically, we generate the same number of random points in our region of interest as the number of Silk Road cells to represent random sites. We then calculate the shortest distance between the centroid of each grid cell to the centroids of the cells containing those random points, as illustrated in Figure 2.7.

We then re-estimate the OLS equation (2.1) using the shortest distance to a random site instead of the shortest distance to a Silk Road site. The results are shown in the first column of Table 2.6. We do not observe a statistically significant relationship between the shortest distance to a randomly generated site and the intensity of night lights.

Least-cost mobility patterns of nomadic herders: Although our simulated herding patterns use pasture quality as the cost raster in the flow accumulation modeling, one could still be concerned that pasture quality may correlate with elevation, which would violate the exclusion restriction. To address this concern, we simulate the least-cost paths of the nomadic herders. Specifically, we generate new mobility patterns of the nomadic herders using the same flow accumulation modeling algorithm, but with the slope of the terrain as a cost raster in the model instead of pasture qual-

would expect to see larger variations in night lights intensity in their paper.

Figure 2.7: Shortest Distance to Random Points



Notes: Shortest distance between the centroid of each grid cell and the centroid of the nearest cell with a randomly generated site. The number of randomly generated sites is the same as the number of the Silk Road cells present in the dataset.

Table 2.6: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Night lights (VIIRS)	Distance to herding path	Night lights (VIIRS)	Distance to herding path	Night lights (VIIRS)	Night lights (VIIRS)
Distance to Silk Road site			-0.087**		-0.131***	
			(0.036)		(0.033)	
Distance to herding path		0.204***		0.199***		-0.001
		(0.054)		(0.034)		(0.001)
Distance to random site	-0.011					
	(0.013)					
Caloric suitability index	0.001	-0.041***	-0.003*	-0.023**	-0.003*	-0.000
	(0.002)	(0.014)	(0.002)	(0.011)	(0.002)	(0.000)
Ruggedness	-0.013***	-0.045*	-0.017***	-0.045**	-0.014***	-0.001
	(0.003)	(0.027)	(0.004)	(0.022)	(0.005)	(0.001)
Precipitation	0.002	0.266***	0.025**	0.294***	0.064***	-0.003
	(0.008)	(0.051)	(0.012)	(0.052)	(0.012)	(0.003)
Distance to river	-0.004**	-0.022***	-0.004**	-0.032***	-0.008***	-0.001*
	(0.002)	(0.007)	(0.002)	(0.008)	(0.002)	(0.001)
Elevation	-0.023**	-0.029	-0.047***	0.075	-0.060***	0.002
	(0.011)	(0.116)	(0.015)	(0.085)	(0.016)	(0.004)
Distance to herder's		-0.000	-0.000			
least-cost path		(0.000)	(0.000)			
Crop-Suitability-Indices	Yes	Yes	Yes	No	No	Yes
NDVI	Yes	Yes	Yes	No	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
F-stat		14.278		34.271		
N	17,125	17,125	17,125	17,675	17,676	7,467

Notes: This table reports results from the robustness check exercises. Distance to Silk Road site, distance to herding path, and distance to random site are standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. Column 1 shows the OLS results when we replace distance to Silk Road site with distance to randomly generated sites. Columns 2 and 3 show the IV results when we control for distance to herder's least-cost path. Columns 4 and 5 show the IV results when we exclude crop-suitability-indices and NDVI from the control variables. Column 6 shows the relationship between distance to herding path and night lights intensity when we only consider locations that are at least 150km away from the nearest Silk Road site. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ity. Intuitively, this new, generated path represents the hypothetical path we would expect 'animals' to take with a preference for a flatter slope. In the same manner as the construction of our instrument, we choose a cutoff of 1.5 million flow values for ease of computation. We then compute the shortest distance between each grid cell and the simulated least-cost paths of the herders.

Column 3 of Table 2.6 present the second-stage IV results after controlling for the shortest distance to the simulated least-cost paths of the herders. We can see that controlling for the least-cost paths of the herders does not significantly impact our IV estimate. Once we take into account the herders' least-cost paths, a one standard deviation increase in distance to Silk Road site corresponds to a 8.3% decrease in the night lights intensity, which is relatively similar to our baseline estimate of 10.0%.

Excluding crop-suitability-indices and NDVI from control variables: Given that nomadic herders' mobility pattern is simulated using pasture quality as the input, one may be concerned that the variation in our instrument is all driven by agriculture-related factors. To verify that this is not the case, we exclude crop-suitability-indices and NDVI from our IV regressions. The fourth and fifth columns of Table 2.6 present the results. When we do not control for crop-suitability-indices and NDVI, a one standard deviation increase in distance to the nearest Silk Road site is associated with a 12.3% decrease in night lights, which is relatively close to our baseline estimate of 10.0%. The exclusion of crop-suitability-indices and NDVI does not significantly affect our results. Although the simulated herding path uses pasture quality as the input data, the actual simulation of the flow of animal stocks depends on the direction and the relative pasture quality. Nomadic herders move from locations with low pasture quality to locations with relatively higher pasture quality. The simulated herding path tracks their mobility pattern, crossing both places with low and high pasture quality. Our instrument is therefore not solely driven by

agriculture-related factors.

Testing correlation between proximity herding path and development: We conduct an additional exercise to verify that proximity to herding path only affects modern development through the Silk Road. If the herding path of the nomads only affects modern development through the Silk Road, we would expect to see no correlation between proximity to herding path and modern development in places where the Silk Roads are absent. We test this by only considering observations that are at least 150 km away from the nearest Silk Road site and run the OLS regression (2.1), replacing distance to Silk Road site with distance to herding path. The last column of Table 2.6 present the results. We can see that there is no significant relationship between proximity to herding path and night lights intensity in areas that are not in close proximity to the Silk Road. The result supports that proximity to herding path only affects modern development through the Silk Road.

Other robustness checks: We perform additional robustness checks, including using an alternative measure for proximity to Silk Road sites, excluding outliers, using different grid sizes, and using different cutoffs to construct our instrument. The results are reported and discussed in Appendix B.2.

2.5 Mechanisms

The results in the previous section indicate that the persistent impact of the ancient Silk Roads on modern development four centuries after the decline of the overland Silk Roads. However, what are the mechanisms through which the ancient trade routes impact the level of development in modern day? Night lights intensity itself is an indicator of economic activity. It does not answer why the ancient Silk Road's effects are persistent over time. In this section, we discuss two possible mechanisms: connectivity and capital investment.

Connectivity refers to the "connections" made between different places, which have endured over time. The ancient Silk Road both fostered new connections and strengthened existing connections across Eurasia. These connections may have persisted over time, even after the decline of the overland Silk Roads. Therefore, places that developed these connections as a result could remain more connected to markets today, and subsequently may be more developed and able to share new technologies more quickly. As time passes, this connectivity could manifest itself through a stronger transportation network.

To explore the connectivity mechanism, we substitute distance to major roads in place of the nighttime lights intensity as a dependent variable. The first column of Table 2.7 presents the second stage results from the IV regression, using the shortest distance to a major road as a dependent variable. On average, as the distance to Silk Road increases by one standard deviation, the distance to a major road increases by 1.1 standard deviations. The result provides some evidence of the persistent effect of the Silk Roads arising from the connectivity mechanism.

Additionally, the persistent effect of the ancient Silk Roads may be a result of capital investment. Capital investment may refer to physical capital such as irrigation and urban development. As trade activity grew, investment in infrastructure in areas along the ancient Silk Road arose to service the traders. Even after the decline of the overland Silk Road trade, such infrastructure continued to be a source of agglomeration for subsequent economic activity and infrastructure investment. As a result, places along the ancient Silk Road could benefit from better irrigation and a higher density of health services.

To explore the investment mechanism, we substitute irrigation and building stock measures as a dependent variable in place of the nighttime intensity. Columns 2-6 of Table 2.7 report the results from the second stage IV regressions. The second column uses the (log) percentage of land equipped for irrigation as the dependent variable.

Table 2.7: Mechanisms

	Distance to road	Irrigation	Building stock used for			
			Housing	Health	Education	All purposes
Distance to Silk Road site	1.087***	-1.252***	-1.656***	-0.192***	-1.236***	-2.141***
	(0.359)	(0.257)	(0.313)	(0.045)	(0.250)	(0.416)
Caloric suitability index	-0.008	-0.041	-0.072**	-0.012***	-0.039*	-0.084**
	(0.038)	(0.025)	(0.030)	(0.003)	(0.023)	(0.039)
Ruggedness	0.191***	-0.202***	-0.269***	-0.019***	-0.215***	-0.378***
	(0.053)	(0.038)	(0.047)	(0.006)	(0.037)	(0.062)
Precipitation	-0.594***	0.256***	0.484***	0.044***	0.361***	0.713***
	(0.168)	(0.093)	(0.115)	(0.014)	(0.090)	(0.156)
Distance to river	0.019	-0.105***	-0.146***	-0.011***	-0.112***	-0.197***
	(0.018)	(0.018)	(0.018)	(0.002)	(0.014)	(0.023)
Elevation	-0.217	-0.314**	-0.287*	-0.053**	-0.214*	-0.273
	(0.164)	(0.130)	(0.165)	(0.021)	(0.127)	(0.213)
Crop-Suitability-Indices	Yes	Yes	Yes	Yes	Yes	Yes
NDVI	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
N	17,125	17,125	17,125	17,125	17,125	17,125

Notes: This table reports results from the second-stage IV regressions. Distance to Silk Road site and distance to road are standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Columns 3-6 use the (log) value of building stock used for housing, healthcare, education, and any purposes in that order. We consistently see a negative and significant relationship between distance to Silk Road site and measures of capital stock, indicating that capital investment and agglomeration could be one of the mechanisms through which the ancient Silk Roads affect modern development.

Our results in this section have shown a higher level of connectivity and capital investment in areas closer to the Silk Road. Note that both types of mechanisms may be acting in complementarity with each other, i.e., capital investment enhances connectivity, and connectivity enhances investment. Because of this, due to our current empirical strategy, we do not aim to disentangle one category of mechanism from each other, but provide suggestive evidence of their functioning.

2.6 Conclusion

The paper studies the long-term effect of trade routes on development. Using a novel instrument constructed from herding flows in the Inner Asia Mountain Corridor, this paper finds that locations along the highland Silk Road continue to be more developed centuries after the decline in the importance of ancient Silk Road overland trade corridors. The results are robust across different measures of modern development and different specifications.

Future work in this area may include studying further which types of Silk Road sites continue to have persistent effects on modern development over time and what mechanisms seem to be the most important. Another direction for work includes developing a theoretical model, including mechanisms, to explain the empirical link between the ancient Silk Road and modern development.

Simply, this study seeks to understand the effects of the past on the present and demonstrates that even choices made in antiquity, which dwindled in importance by

the sixteenth century, still have enduring persistent effects on modern development today. Policymakers may consider that not only does the choice of a trade route or improvement in transportation have effects a decade later, but centuries later. Today, the ancient Silk Road is being revitalized in China's Belt and Road Initiative, which was introduced in 2013. Along with concentrating on maritime shipping lanes ("roads"), the Chinese government has identified several overland corridors ("belts") that they will target for development. One of these overland corridors is called the China-Central Asia–West Asia Economic Corridor (CCWAEC), a geographic area that encompasses the Inner Asia Mountain Corridor. Our study could have important implications for such an initiative.

Chapter III

Corporate Influence in Trade Agreements: Approaches Using Machine Learning

Joint with Dyanne Vaught

3.1 Introduction

Corporate influence in trade agreements has long been theorized informally (Rodrik, 2018; Stiglitz, 2015), but only recently with newer sources of trade and lobbying data that social scientists been able to more formally document this relationship. (Rodrik, 2018) sketches out two political economy perspectives; the first views import-competing interests as the dominant force in trade policy, and so trade agreements serve to counteract domestic protectionism. Indeed, prior to the 1980s, most trade agreements focused on reducing tariffs and quotas. The second political economy perspective posits that export-side interests are the dominant force in trade policy, and because their concerns are self-interested and rent-seeking, trade agreements may not necessarily result in welfare-increasing trade. Proponents of this perspective frequently point to the rise of non-trade issues, non-tariff barriers, and international arbitration in trade agreements.

In particular, the current lobbying landscape for trade agreements is dominated by large multinational corporations (MNCs) who overwhelmingly support free trade agreements with little to no opposition from opposing domestic firms (Kim and Mil-

ner, 2019; Blanga-Gubbay et al., 2023). The differences between MNCs and primarily domestic firms have been well-documented: MNCs tend to be larger and more productive (Bernard et al., 2009); MNCs tend to be the largest exporters, more integrated in global value chains, and employ the most skilled workers (Autor et al., 2017). Thus, domestic firms and MNCs are likely to have differing preferences in trade agreement implementations.

Various studies have found that MNCs have been the main proponents of provisions that address investment protection, service liberalization, and intellectual property rights (Manger, 2009; Kim, 2017; Baccini, 2019). Moreover, MNCs seem to prefer supranational venues for enforcement and decision-making, usually through established institutions such as the WTO or IMF, or through special arbitration tribunals outlined in trade agreements (Levy and Prakash, 2003; Alford, 2013).

Modal studies on the connection between corporations and trade agreements draw from the reports filed under the United States' Lobbying Disclosure Act of 1995, which provides information on which firms lobby for trade agreements, their position, and the lobbying effort in terms of expenditure and number of reports filed (Kim, 2017). However, the dataset does not include the reports themselves, which would allow for a more direct evaluation of the effect of the lobbying influence on the final trade agreement. Instead, we contribute to this literature by studying the similarity between standardized corporate language through SEC 10-K filings and trade agreements. Corporate lobbyists have been well documented to directly write drafts of laws and regulations, and receiving lawmakers have been documented to directly pull from such drafts (Chang, 2013; Lipton and Yaffe-Bellany, 2022). We further segment the corpus to build a subset of only multinational companies to examine whether there is a differential effect of multinational corporate language and whether this can provide suggestive evidence to the different motivations of MNCs in the trade policy space.

Our second primary contribution to the literature is a thorough examination of

the trends in corporate influence across alternative natural language processing techniques, up to the state-of-the-art in 2022. We find that while the broad trend of increasing corporate influence over time holds up across these different methods, initial assumptions and specifications are highly important and can lead to significantly different results. Ultimately, further work must be done to validate the findings of these methods.

We start with traditional hand-coded methods. First, we use an indicator variable dataset, Design of Trade Agreements (DESTA), and build an index of corporate influence, and find that an increase of one in the corporate influence index implies an increase of 1.1% in total trade flows and 0.89% for manufacturing trade flows. The second hand-coded method is a dictionary-based sentiment analysis, in which we find a one standard deviation increase in the Constraining proportion of words is associated with a 6.5% decrease in trade. We then instantiate two topic models which qualitatively overlap in four out of five topics; trade flows respond positively to an increase in the Services/Investment topic proportion from 0.19% to 0.92% for total trade flows and from 0.37% to 0.92% for manufacturing trade flows. Next, we deploy a doc2vec model to examine cosine score similarities between 10-K language and trade agreements, finding that both similarity with corporate language and similarity with multinational-specific corporate language result in significant, negative impacts on trade flows, which lends some credence to Rodrik (2018)'s conjecture that corporations may lobby for self-interested policies that may not necessarily increase trade flows and/or result in lower barrier trade. A 10% increase in corporate language similarity and a 10% increase in MNC language similarity tends to decrease total trade flows, on average, by -0.94%. Last, we repeat the similarity exercise using sentence transformers models and continue to find suggestive results that language associated with multinational corporations in particular may have differential, negative results on trade flows.

Section 3.2 briefly explicates the data used in the project; Section 3.3 reviews the basics of each model and its application in a structural gravity model; Section 3.4 provides a general review of results in comparison with each other; and Section 3.5 concludes.

3.2 Data

3.2.1 Texts of Trade Agreements

Text of Trade Agreements (TOTA) is a dataset of trade agreement texts[1] from on the WTO Regional Trade Agreement Database (Alschner et al., 2017) with agreements spanning from 1949 to 2017. In addition to making the text of the agreements machine-readable, TOTA also includes some basic metadata such as a list of signatories, date in force, force inactive. As of March 2023, we use version 0.5.0, which includes 448 agreement texts notified to the WTO, and two texts for the Trans-Pacific Partnership.

3.2.2 Supplement to TOTA

In addition to the texts supplied by TOTA, we collect 153 additional post-1940 trade agreements that are easily accessed and machine-readable. The new agreements span from 1954 to 2020. Adding this supplement to TOTA, we have a total of 601 agreements.[2] We further subset to only agreements in English, resulting in a final dataset of 532 trade agreement texts.

[1]TOTA does not include trade agreement appendices or riders.

[2]Interested parties may download the additional raw texts with metadata at the following LINK.

3.2.3 Trade Flows

This paper uses two datasets for trade flows. First, the paper uses CEPII's Gravity dataset for baseline regressions (Conte et al., 2022). The 2022 release provides data from 1948 to 2020 and includes 252 countries. The paper also uses the U.S. International Trade Commission's International Trade and Production Database for Estimation (ITPD-E) (Borchert et al., 2022). As of March 2023, we are using Release 2 which provides data from 1986 to 2019, includes 265 countries, and 170 industries. The ITPD-E has industry and sector coverage and includes both domestic and international trade flows (Borchert et al., 2021).

3.2.4 Design of Trade Agreements

The Design of Trade Agreements (DESTA) Database is a database of 710 preferential trade agreements signed between 1948 to 2019 for which individual readers hand-coded over 100 characteristics. DESTA documents "design features" of trade agreements, and includes basic information on the type of agreement (bilateral vs. multilateral), geographic location of the countries/parties involved in the agreement, and language. DESTA also provides detailed information on the topics covered in the agreement, such as services liberalization, dispute settlement, movement of people in addition to goods and services, and competition.

3.2.5 Corporate 10-K Filings

Form 10-K is a report filed annually with the United States Securities and Exchange Commission (SEC) by every publicly traded company in the U.S. These forms provide detailed information on the company's financial performance in the prior fiscal year, and they must include an overview of the company's business and operations, financial data, an outline of risk factors faced by the company, a discussion and anal-

ysis of the company and its current standing by management, and complete financial statements. We use 10-K filings from 1996 to 2020 as a source of "corporate" language to create a corporate language dictionary.

3.2.6 Loughran-McDonald Sentiment Word Lists

Introduced in Loughran and McDonald (2011), the Loughran-McDonald Sentiment Word Lists were developed to classify words in a financial context, specifically 10-X documents from 1993 onwards. The sentiment categories are: negative, positive, uncertainty, litigious, strong modal, weak modal, and constraining. Words may appear in more than one category, and may be removed or added as sentiment word lists are updated to reflect contemporary usage. We use the January 2022 version of the negative, litigious and constraining word lists to classify words in the trade agreements and Form 10-Ks.

3.3 Models

In this section, we briefly explicate and implement five methods to examine corporate and multinational influence in trade agreements over time and apply their outputs to a structural gravity equation. The specification for the structural gravity equation, estimated via Poisson Pseudo Maximum Likelihood, is as follows:

$$X_{ij,t} = \exp[\pi_{i,t} + \mu_t + \chi_{i,j} + \eta_1 \text{PTA}_{ij,t}] \times \varepsilon_{ij,t} \tag{3.1}$$

where $X_{ij,t}$ is nominal trade flows at non-consecutive year t, $\pi_{i,t}$ is the set of time-country fixed effects, μ_t are the time fixed effects, $\chi_{i,j}$ are bilateral characteristics (log distance, borders, common language), and $\text{PTA}_{ij,t}$ is the set of variables that identify whether the country pair has a trade agreement at time t and if relevant, interaction

variables with topic and similarity measures. $\varepsilon_{ij,t}$ is the error term. Each gravity equation is estimated using data every five years from 1960 to 2015.

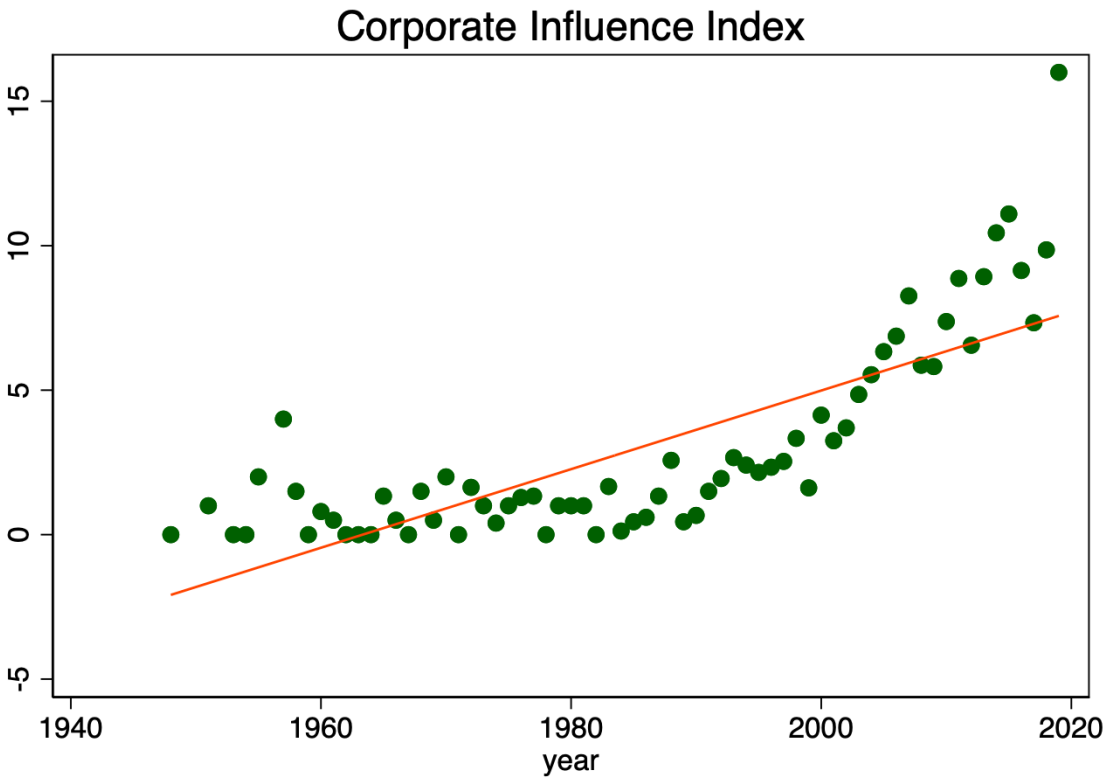
3.3.1 DESTA

3.3.1.1 Implementation

Following Rodrik (2018)'s observations that (1) provisions about services often require changes in domestic regulations and are driven by corporate interest and expertise (2) corporations tend to make more investments in partner countries when there are provisions for trade defense, investment protections, and intellectual property protections, we build a Corporate Influence Index. Simply, this index is a summation of all of the DESTA-counted provisions that would be particularly beneficial for corporations. The construction of the index follows that of DESTA's own procedure when building its own Depth Index (Dür et al., 2014), which adds the number of substantive provisions. For the specific provisions included, see Appendix C.1.1.

As seen in Figure 3.1, the Corporate Influence Index has been growing over time, noticeably accelerating in the 1990s. The individual components of the index also have a similar pattern (Figure C.1).

Figure 3.1: Corporate Influence Index (DESTA)



Notes: Data from DESTA. For the specific provisions included in the index, see Appendix C.1.1.

3.3.1.2 Application: Gravity Equation

Table 3.1 shows the regression results when including the Corporate Influence Index. Throughout this paper, columns (1) and (3) are the baseline gravity equations without any corporate influence or similarity scores with the dependent variables as total trade flows and manufacturing trade flows, respectively.

With the inclusion of the Corporate Influence Index, the PTA coefficient does decrease, and the interaction variables are positive and around similar magnitudes for both total trade flows and manufacturing trade flows. An increase of one in the corporate influence index implies an increase of 1.1% in total trade flows and 0.89% in manufacturing trade flows.

Table 3.1: Gravity Equations with DESTA Country Influence Index

	Baseline Trade Flows (1)	DESTA (2)	Baseline Manuf. Flows (3)	DESTA (4)
PTA	0.2488	0.1649	0.3183	0.2483
	(0.034)***	(0.051)***	(0.034)***	(0.047)***
PTA x Corp. Infl. Index		0.0111		0.0089
		(0.005)***		(0.005)**
Log Distance	-0.7165	-0.7102	-0.6931	-0.6885
	(0.018)***	(0.018)***	(0.017)***	(0.018)***
N	83196	83196	83196	83196
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] The Corporate Influence Index is created by summing together DESTA indicators for provisions that are more friendly to corporations in the services, investments, property rights, and trade defense categories. See C.1.1 for the complete list.

3.3.2 Sentiment Analysis

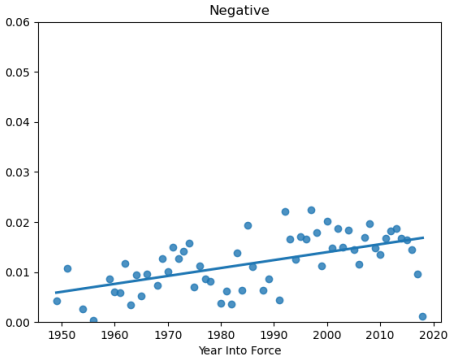
3.3.2.1 Implementation

Sentiment word lists are taken from Loughran and McDonald (2011), which is based on the 2of12inf list from the SCOWL (Spell Checker Oriented Word Lists) and Friends database. They further refine this word list into several categories for sentiment analysis in financial contexts. In this paper, we use the Negative, Litigious, and Constraining word lists. Figure 3.2 shows the mean proportions of each category for trade agreements by year; Figure 3.3 shows the mean proportions by year for 10-Ks.

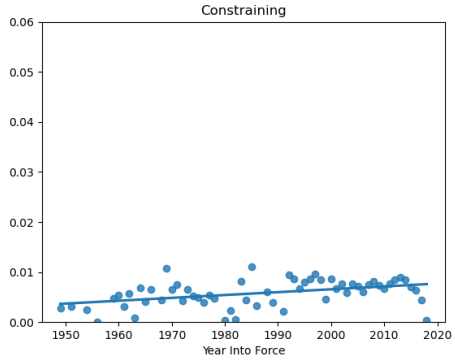
While the other categories are flat, the Negative category increases over time for both trade agreements and 10-Ks. The Negative category includes words strongly associated with arbitration and goods and services such as: "anticompetitive," "violation," "illegal," "retaliation," "investigate," and "embargo." Given that corporations who lobby for trade agreements tend to focus on arbitration and goods and services, the Negative category may reflect an increasing amount of corporate influence in trade agreements. However, patterns over time may be influenced by the dynamic nature of the sentiment word lists themselves, in which words are added and removed over time to reflect contemporary usage. Thus, the January 2022 vintage that is used in this paper may reflect more contemporary usage which by dint of recency, may be more present in the years closer to 2022.

Figure 3.2: Trade Agreements: Sentiment Word Lists Mean Proportions Over Time

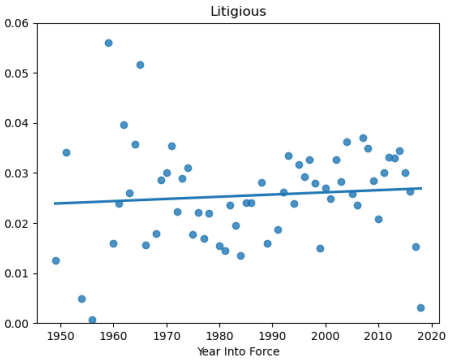
(a) Negative



(b) Constraining



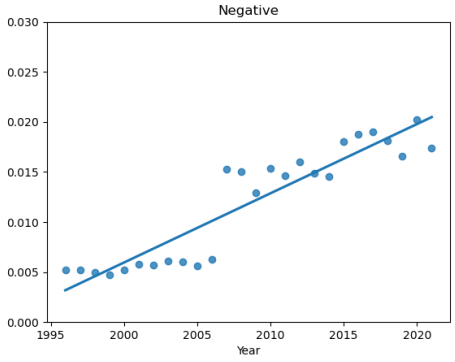
(c) Litigious



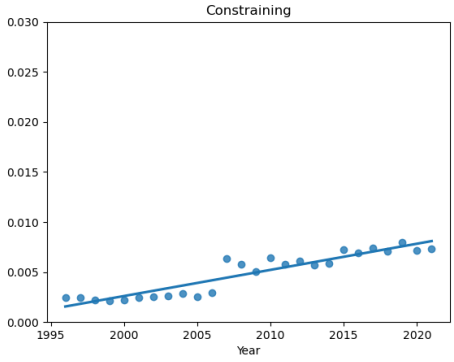
Notes: Proportions are calculated by taking the word list, counting the number of appearances in each document, and then dividing the count by the total word count of the document. Agreements are aggregated by year using a simple mean.

Figure 3.3: 10-Ks: Sentiment Word Lists Mean Proportions Over Time

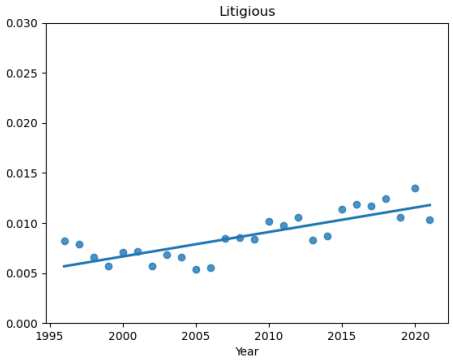
(a) Negative



(b) Constraining



(c) Litigious



Notes: Proportions are calculated by taking the word list, counting the number of appearances in each document, and then dividing the count by the total word count of the document. 10-Ks are aggregated by year using a simple mean.

3.3.2.2 Application: Gravity Equation

Table 3.2 and 3.3 show regression results including the sentiment categories. Each category is contained in its own regression as the word lists are not exclusive, e.g. words in Litigious may also appear in Constraining. The only significant result is for total trade flows with respect to the Constraining category: a one standard deviation increase in the Constraining proportion (0.005) is associated with a 6.5% decrease in trade flows. These overall insignificant results echo those of Loughran and McDonald (2011), who find that these word lists are not significantly associated with stock returns over the following twelve months.

Table 3.2: Gravity Equations with Loughran-McDonald Sentiment Word Lists

	Baseline Trade Flows (1)	(2)	(3)	(4)
PTA	0.2488	0.2705	0.337	0.3294
	(0.034)***	(0.051)***	(0.059)***	(0.043)***
PTA x Log Pro. Negative		-1.7723		
		(2.672)		
PTA x Log Pro. Litigious			-3.6804	
			(2.041)	
PTA x Log Pro. Constraining				-13.3688
				(5.009)***
Log Distance	-0.7109	-0.7207	-0.7216	-0.7232
	(0.018)***	(0.016)***	(0.016)***	(0.016)***
N	83196	83473	83473	83473
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3.3: Gravity Equations with Loughran-McDonald Sentiment Word Lists (Manufacturing)

	Baseline Manuf. Flows (5)	(6)	(7)	(8)
PTA	0.3183	0.2923	0.3301	0.3535
	(0.034)***	(0.051)***	(0.056)***	(0.041)*
PTA x Log Pro. Negative		0.5800		
		(2.598)		
PTA x Log Pro. Litigious			-1.1874	
			(1.915)	
PTA x Log Pro. Constraining				-8.3966
				(4.713)
Log Distance	-0.6931	-0.6977	-0.6984	-0.6998
	(0.017)***	(0.016)***	(0.016)***	(0.016)***
N	83196	83473	83473	83473
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.3.3 Topic Modeling

3.3.3.1 Implementation

Generally, a topic model is a semi-unsupervised machine learning method that categorizes the topics that occur in a collection of documents. In this class of models, we focus on Latent Dirichlet Allocation (LDA) (Blei et al., 2003) and Structural Topic Model (STM) (Roberts et al., 2019). Both models assume ordering of words is unimportant (bag of words) and that document sequencing is unimportant.

LDA starts with the assumption that each document $i \in \{1, ..., M\}$ in the corpus D is generated as a random mixture over independent latent topics, with each topic $k \in \{1, ..., K\}$ represented as a distribution over all the words in the corpus. Working backwards using the follow iterative process, one may infer these K latent topics:

1. Guess the distribution of topics in each document i, $\theta_i \sim \text{Dir}(\alpha)$ where $\text{Dir}(\alpha)$ is the Dirichlet distribution with parameter α.

2. Guess the distribution of words in topic k, $\varphi_k \sim \text{Dir}(\beta)$ where $\text{Dir}(\beta)$ is the Dirichlet distribution with parameter β.
3. For each of the words in document i and numerical location j in document, choose a topic $z_{i,j} \sim \text{Multinomial}(\theta_i)$ and word $w_{i,j} \sim \text{Multinomial}(\varphi_{z_{i,j}})$.
4. Compute proportion of words in document i that are assigned to topic k, and the proportion of assignments to topic k that come from word w.
5. Update probability for word w belonging topic k by multiplying both proportions together in the previous step.
6. Repeat until the distribution of topics converges to an acceptable tolerance.

STM has a similar framework to LDA, but extends LDA in that STM is able to use metadata (i.e. characteristics such as author, source, year), allows for topics to be correlated with each other, and distributions are specific to each document instead of one distribution for the whole corpus.

In both models, the user must choose the number of topics, K, to be evaluated. Quality of the topic model is often evaluated using semantic coherence score, which broadly measures the extent of word co-occurrence within topics. There are several methods of measuring coherence; in this paper, we use UMass and UCI methods. In general, the higher the score, the greater the coherence and supposedly better separation of topics. As observed in 3.4, UCI and UMass coherence scores dip at $K = 5$, and subsequently $K = 5$ is a natural inflection point to observe with human comprehension whether $K = 4$, $K = 5$, and $K = 6$ yields the most understandable set of topics. $K = 5$ is chosen, and subsequently $K = 5$ for STM is also chosen to mirror LDA.

Figure 3.5 and Figure 3.6 show the mean topic proportion of time as calculated by LDA and STM, respectively. LDA and STM overlap for four topics: Protocol, Goods, Community, Services and Investment. For the final two topics, LDA seems to favor Arbitration, and STM seems to weight towards words dealing with Contracts.

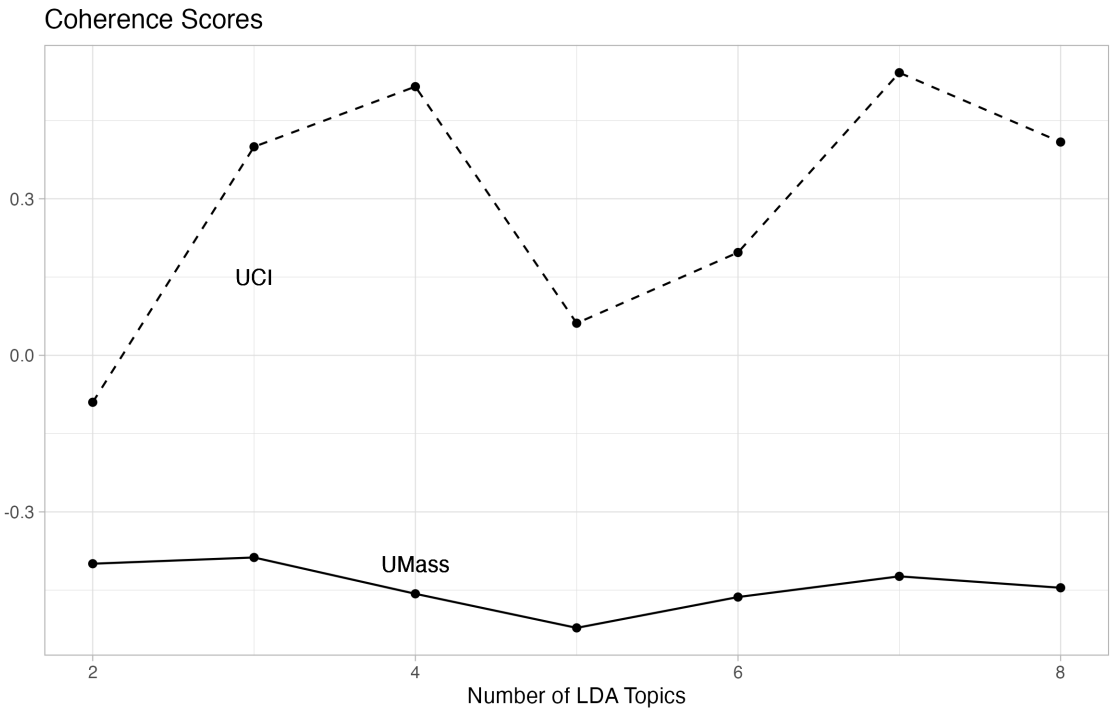
Despite the overlap in topics, LDA and STM differ substantially in general patterns over time; for example, Services and Investment has two humps of activity in

LDA whereas in the STM estimation, Services and Investment is generally increasing over time (Figure 3.5(d), 3.6(e)). For Protocol, there is a steady increase in the LDA topic proportion whereas STM is static (Figure 3.5(a), 3.6(a)). Goods topic proportions for both are largely static over time with a slight increase in LDA and a slight decrease in STM. The Community topic proportion in both LDA and STM are decreasing over time (Figure 3.5(b), 3.6(d)).

In contrast with the traditional sentiment analysis using a manual dictionary, LDA and STM approach topics differently. Table C.1 and C.2 report correlations between the sentiment word lists, LDA topics, and STM topics. The sentiment categories all have a strong correlation with each other (greater than 0.8). When comparing across the sentiment categories and LDA topics, the strongest correlation is between the Negative category and LDA topic Services and Investment at 0.539. The relation between the Litigious category and LDA topic Arbitration is weak at 0.178. Between the sentiment categories and STM, all correlations are weak. While manual word lists may provide an easily understandable way to classify sentiments, they lack the ability to use context unless researchers update the word list themselves to fit their use case. For example, the Loughran-McDonald list categorizes the word "regulations" and its stem words as "litigious"; however, a trade agreement is by nature an agreed upon set of regulations, so including this family of words does not necessarily indicate the agreement's focus on arbitration.

There are two topics that may reflect the rise of corporate influence; for LDA, the increase in Arbitration (Figure 3.5(e)); for STM, the increase in Services and Investment (Figure 3.6(e)).

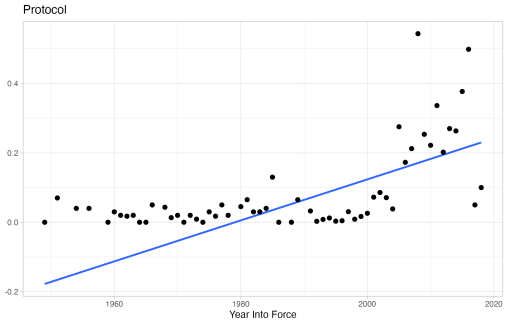
Figure 3.4: LDA Coherence Scores



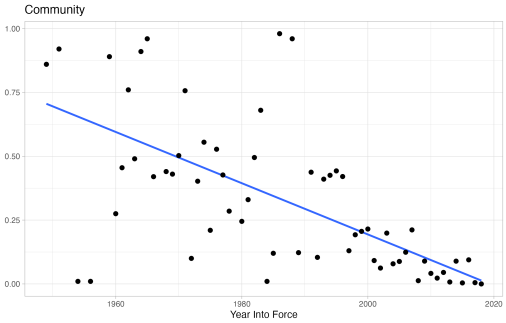
Notes: The UMass coherence score is calculated using the following formula for each pair of words, w_i and w_j: $C_{\text{UMass}}(w_i, w_j) = \log \frac{D(w_i,w_j)+1}{D(w_i)}$ where $D(w_i, w_j)$ indicates how many times w_i and w_j appear in documents together, and $D(w_i)$ is overall frequency of w_i. The UCI coherence score is calculated based on a sliding window (110 words default for Python's `gensim`) around the word in question: $C_{\text{UCI}}(w_i, w_j) = \log \frac{P(w_i,w_j)+1}{P(w_i)\cdot P(w_j)}$ where $P(w)$ is the probability of word w being in the sliding window, and $P(w_i, w_j)$ is the probability of words w_i and w_j appearing in the same sliding window.

Figure 3.5: LDA Topics' Mean Proportions Over Time

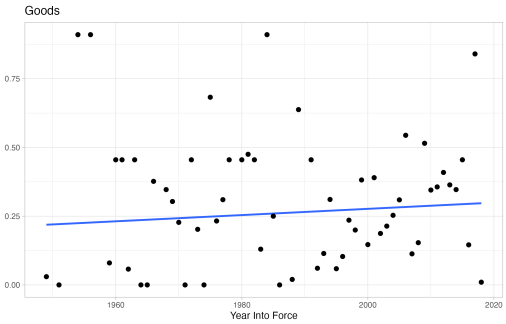
(a) Protocol



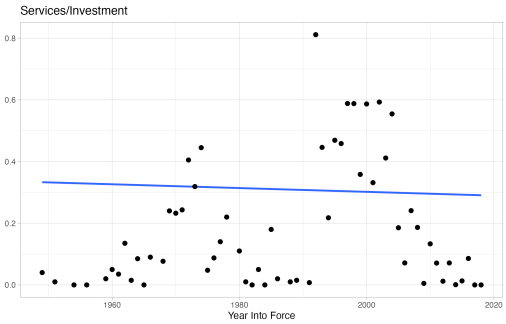
(b) Community



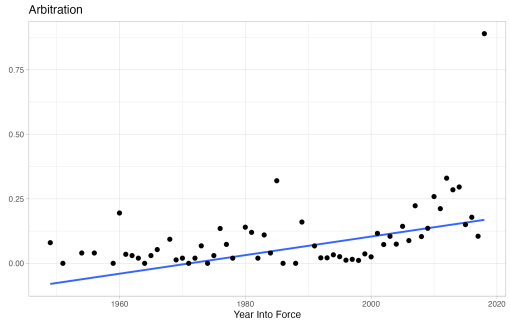
(c) Goods



(d) Services/Investment



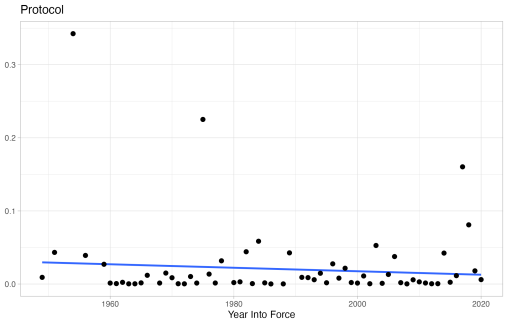
(e) Arbitration



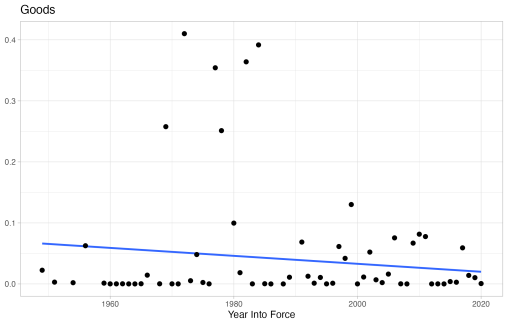
Notes: Agreements are aggregated by year using a simple mean.

Figure 3.6: STM Topics' Mean Proportions Over Time

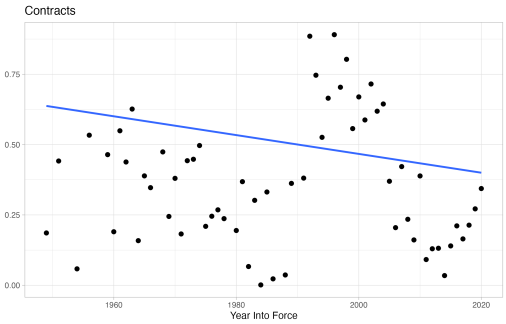
(a) Protocol



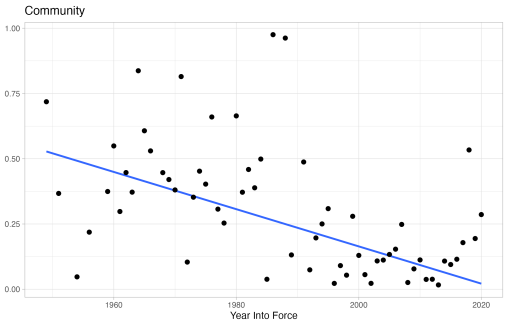
(b) Goods



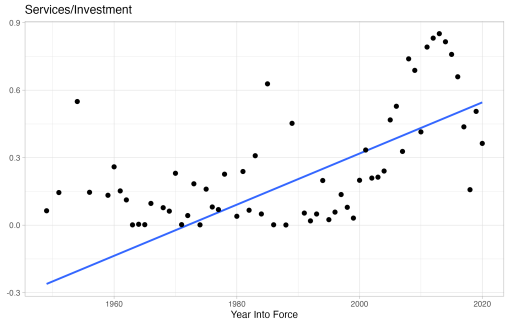
(c) Contracts



(d) Community



(e) Services/Investment



Notes: Agreements are aggregated by year using a simple mean.

3.3.3.2 Application: Gravity Equation

Table 3.4 shows the results from the regressions including LDA topic proportions, and Table 3.5 shows the results from the regressions including STM topic proportions. Within LDA, Services/Investment and Goods had a significant effect for both total trade flows and manufacturing flows, with a 10% increase in topic proportion resulting in around 0.92% increase and 0.77% increase in total trade flows. For STM, the Services/Investment is only significant for the manufacturing flows, but Goods is significant for both total flows and manufacturing flows. The STM coefficients are smaller in magnitude than the LDA coefficients and are the same sign.

Overall, as there is more focus in the trade agreement on the topics of Goods, Services, and Investments, there is an increase in trade flows, and is reflected across two topic modeling strategies.

Table 3.4: Gravity Equations with LDA Topics

	Baseline Trade Flows (1)	LDA (2)	Baseline Manuf. Flows (3)	LDA (4)
PTA	0.2488	0.3831	0.3183	0.4475
	(0.034)***	(0.060)***	(0.034)***	(0.051)***
PTA x Log Protocol		-0.0243		-0.0157
		(0.016)		(0.016)
PTA x Log Community		0.0211		0.0106
		(0.021)		(0.020)
PTA x Log Goods		0.0744		0.0717
		(0.029)**		(0.030)**
PTA x Log Services/Inv.		0.0877		0.0881
		(0.018)***		(0.017)***
Log Distance	-0.7109	-0.7130	-0.6931	-0.6910
	(0.018)***	(0.017)***	(0.017)***	(0.017)***
N	83196	83196	83196	83196
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] Refer to Appendix C.2.1 for words most associated with each LDA topic.

Table 3.5: Gravity Equations with STM Topics

	Baseline Trade Flows (1)	STM (2)	Baseline Manuf. Flows (3)	STM (4)
PTA	0.2488	0.6144	0.3183	0.6998
	(0.034)***	(0.097)***	(0.034)***	(0.017)***
PTA x Log Services/Inv.		0.0190		0.0366
		(0.012)		(0.013)**
PTA x Log Goods		0.0525		0.0476
		(0.009)***		(0.009)***
PTA x Log Contracts		-0.0370		-0.0164
		(0.016)*		(0.014)
PTA x Log Community		-0.0065		-0.0151
		(0.013)		(0.011)
Log Distance	-0.7109	-0.7076	-0.6931	-0.6838
	(0.018)***	(0.017)***	(0.017)***	(0.017)***
N	83196	83196	83196	83196
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] Refer to Appendix C.2.2 for words most associated with each STM topic.

3.3.4 doc2vec

3.3.4.1 Implementation

Doc2vec is an extension of word2vec; instead of estimating distributed vector representations of words, doc2vec estimates distributed vector representations of documents (Le and Mikolov, 2014). Unlike topic models, doc2vec is able to capture semantic meaning and context; in our implementation, we use the continuous bag-of-words (CBOW) architecture in which the model predicts the current word from the surrounding context word window.

Briefly, doc2vec uses a shallow three-layer neural network architecture based on word2vec:

1. For the input layer, each word in the vocabulary is one-hot encoded[3] to a

[3]Each word is associated with a vector where the length of the vector is the number of words in the n vocabulary, and the position of the word is encoded as 1 and all other positions are 0.

vector of size $[1 \times V]$. The number of neurons is V in this layer. In doc2vec, an additional input neuron is added to represent the document.

2. The model assigns or updates weights between the input layer and hidden layer, and between the hidden layer and output layer.

3. The input layer is multiplied by the weights between the input layer and hidden layer, resulting in the hidden layer matrix of size $[V \times N]$ where N is an arbitrary size matrix that determines the size of the embedding space that is used to represent a word. N is also the number of neurons in the hidden layer.

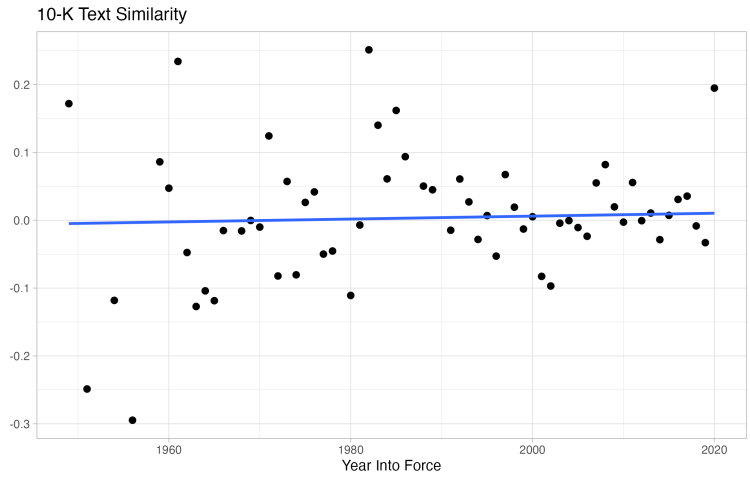
4. The hidden layer is multiplied by the weights between the hidden layer and output layer, which results in the number of neurons V, each of size $[1 \times V]$, where each entry is the probability that if you choose a word near v_i, that will be word v_j.

5. The error between output and true target word is computed, and weights are re-adjusted in the algorithm until convergence tolerance is reached.

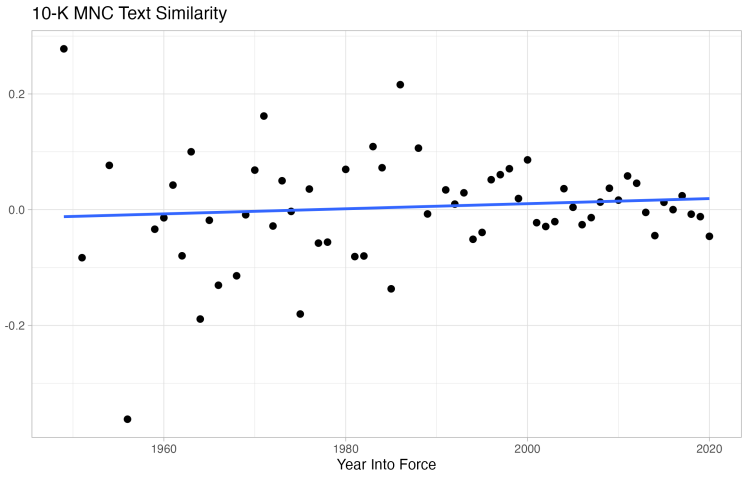
In our application, we trained the doc2vec model on the entirety of Wikipedia from October 2017, and then applied the word embeddings to compute the cosine similarities between trade agreements and corporate 10-K filings. Overall, similarity between 10-K Forms and trade agreements have remained static mostly with a slight trend up for both the 10-K corpus and the 10-K MNC corpus (Figure 3.7).

Figure 3.7: doc2vec Similarity Cosine Scores Over Time

(a) Total



(b) MNCs



Notes: doc2vec is trained on English Wikipedia, October 2017. 10-K Forms are from 1996. Similarities for each year are averages of each of the pairwise similarities between 10-K Forms and the individual trade agreement.

3.3.4.2 Application: Gravity Equation

Table 3.6: Gravity Equations with Doc2vec Similarity Cosine Scores

	Baseline Trade Flows (1)	Doc2vec (2)	Baseline Manuf. Flows (3)	Doc2vec (4)
PTA	0.2488	0.1488	0.3183	0.2324
	(0.034)***	(0.042)***	(0.034)***	(0.039)***
PTA x Log Similarity		-0.0626		-0.0440
		(0.018)***		(0.019)***
PTA x Log Similarity MNC		-0.0348		-0.0390
		(0.013)***		(0.012)***
Log Distance	-0.7109	-0.7258	-0.6931	-0.6968
	(0.018)***	(0.018)***	(0.017)***	(0.018)***
N	83196	83198	83196	83198
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] Similarity scores are based on 2018 10-K forms.

Table 3.6 shows the results for the regressions that include doc2vec similarity scores. In column (2) and (4), the coefficients for the similarity terms are uniformly negative and significant, suggesting that with a 10% increase in corporate language similarity and a 10% increase in MNC language similarity, a trade agreement will tend to decrease total trade flows by -0.94% percent. This is the first evidence provided in this paper of Rodrik (2018)'s postulation that trade agreements influenced by multinationals may in fact curtail and introduce welfare-reducing trade policies.

3.3.5 Transformers

3.3.5.1 Implementation

Transformers belong to subset of machine learning called "deep learning", which leverages deep neural networks with more layers and neurons than a typical neural network (e.g., word2vec, doc2vec). Transformers use the mechanism of self-attention

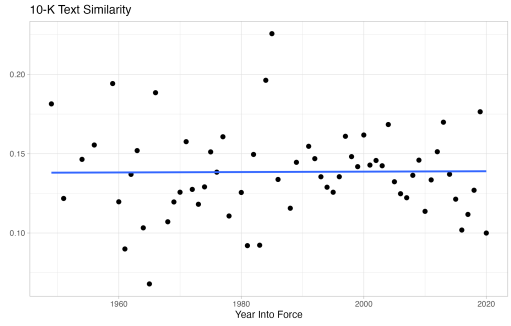
which allows the model to look within other elements of the input sequence as it encodes the input (Vaswani et al., 2017). For this paper, we use a language model called Bidirectional Encoder Representations from Transformers (BERT) in order to compute cosine similarity between documents (Devlin et al., 2018). BERT models improve upon word2vec/doc2vec in the following ways: they are able to handle different meanings of the same word (e.g., bank, a financial entity, or land along the river), they are able to take word order into account, and representations are learned at subword rather than at the word level, allowing BERT models to classify out-of-vocabulary words. Generally, BERT outperforms older methods such as word2vec in sentence completion tasks (von der Mosel et al., 2022).

At a high level, masked language models such as BERT are trained by masking a word in a sentence and forcing itself to bidirectionally use words on either side to predict the masked word. For an in-depth explanation of BERT architecture, please refer to (Devlin et al., 2018).

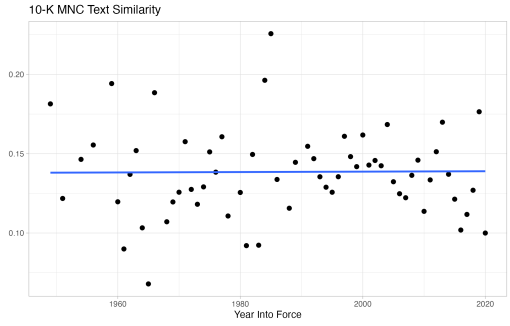
In our implementation, we use two SBERT-trained models, `paraphrase-albert-small-v2` (albert) and `paraphrase-multilingual-MiniLM-L12-v2` (MiniLM), which are trained on datasets that are well-suited to similarity, such as on citation pairs and duplicate questions (Reimers and Gurevych, 2019). In Figure 3.8, similarity scores generated with albert show a slight upward trend of similarity over time where MiniLM does not seem to have a discernible upward or downward trend.

Figure 3.8: BERT Similarity Cosine Scores Over Time

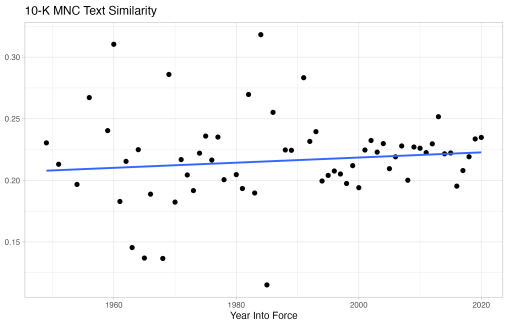
(a) MiniLM - Total



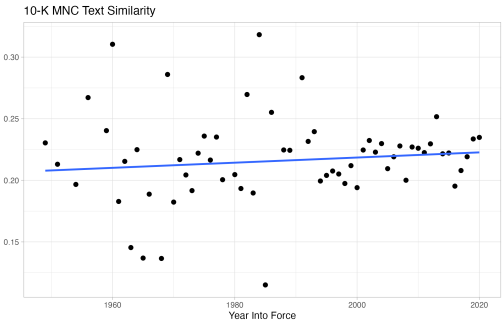
(b) MiniLM - MNCs



(c) albert - MNCs



(d) albert - MNCs



Notes: MiniLM refers to pre-trained model, `all-MiniLM-L12-v2`, trained by Reimers and Gurevych (2019). albert refers to pre-trained model, `paraphrase-albert-small-v2`, trained by Reimers and Gurevych (2019). 10-K Forms are from 1996.

Table 3.7: Gravity Equations with BERT Similarity Cosine Scores

	Baseline Trade Flows	MiniLM	albert	Baseline Manuf. Flows	MiniLM	albert
	(1)	(2)	(3)	(4)	(5)	(6)
PTA	0.2488	0.6933	0.2319	0.3183	0.8322	0.3261
	(0.034)***	(0.144)***	(0.133)***	(0.034)***	(0.135)***	(0.121)**
PTA x Log Similarity		0.1530	0.1219		0.1707	0.1786
		(0.039)***	(0.127)		(0.037)***	(0.123)
PTA x Log Similarity MNC		0.0077	-0.0950		0.0167	-0.1275
		(0.049)	(0.061)		(0.045)	(0.061)*
Log Distance	-0.7109	-0.7182	-0.7060	-0.6931	-0.6966	-0.6809
	(0.018)***	(0.017)***	(0.018)***	(0.017)***	(0.017)***	(0.018)***
N	83196	83198	83198	83198	83198	83198
Common Characteristics	X	X	X	X	X	X
Year Fixed Effects	X	X	X	X	X	X
Country Fixed Effects	X	X	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] MiniLM refers to pre-trained model, `all-MiniLM-L12-v2`, trained by Reimers and Gurevych (2019).
[4] albert refers to pre-trained model, `paraphrase-albert-small-v2`, trained by Reimers and Gurevych (2019).

3.3.5.2 Application: Gravity Equation

Table 3.7 shows the results when including the BERT similarity scores in a structural gravity equation. In all implementations and for both total trade flows and manufacturing flows, the coefficient on similarity is positive; a 10% increase in similarity results in 1.7% (MiniLM) or 1.3% (albert) increase in trade flows.

However, MiniLM and albert differ on the coefficient for MNC similarity where MiniLM returns an imprecise 0 and albert returns a negative coefficient, significant at the 5% level for manufacturing flows, in which a 10% increase in MNC similarity results in a -1.2% decrease in manufacturing trade flows. These results suggest that in contrast to the total corpus, the MNC corpus may have different impact on trade flows and further work would benefit from more precise estimation or fine-tuning of the machine learning models used to determine similarities.

3.4 Comparing Methods

The paper has followed a technological progression of quantitative techniques in order to see whether there is pattern of increasing corporate influence. We study four broad categories of techniques–hand labeling (DESTA, dictionary-based sentiment analysis), topic modeling (LDA, STM), word embeddings (doc2vec), deep learning (MiniLM, albert)–and measure prominence of certain topics that corporations and MNCs are most concerned about and the similarity scores to corporate 10-Ks over time. From this, we can see broadly that each method provides some evidence that corporate influence is increasing over time.

However, this evidence is highly influenced by the method used. Indeed, for most measures, DESTA indicators are increasing over time or becoming more frequent as trade agreements have naturally progressed and become deeper over time. Similarly, dictionary-based sentiment analysis requires human input on categorization tasks.

While DESTA and dictionary-based sentiment analysis are more easily interpreted, they may suffer from being sensitive to context. Machine learning methods have the ability to add dimensions using the text of trade agreements that is not easily quantifiable in an indicator statistic, such as similarity to corporate language and distribution of topics within the corpus.

Topic modeling is the next natural progression and is a self-contained exercise which only requires the text of trade agreements. LDA and STM pick up mostly the same topics, and those coefficients have similar signage in their respective structural gravity estimations. The results in this paper indicate that the prevalence of topics of more relevance to corporate concerns (i.e., Goods, Services/Investment) results in more trade flows. Yet, the difference in magnitudes outlines the difficulties in establishing ground truth for such topic models even when prudent scholars use training and test splits and multiple rounds of validation.

The following methods, doc2vec and sentence transformers, are able to train on outside datasets and transfer the learning to the corpus at hand. For both the doc2vec and sentence transformers models, similarity to trade agreements increase only modestly over time. In the doc2vec model, all the similarity coefficients were negative, whereas for the sentence transformers, only the albert MNC similarity coefficient was negative. Taken together with imprecise estimates, this suggests that there may be differential effects between the general corpus similarity and MNC similarity, but there needs to be more precise estimation or fine-tuning of models for trade agreement language. When considering the characteristics of doc2vec and sentence transformers, one would expect sentence transformers to be more accurate as they can take into account word ordering, grammatical structure, and are trained explicitly on pairwise data (duplicate questions, pairwise citations). Yet, even within the family of sentence transformers, the structural gravity equations results are not broadly consistent.

Overall, while the state of the art in natural language process has moved quickly

over the past twenty years, even with hand-coded data and our ability to discern a (slightly) increasing corporate influence, the relationship between corporate influence and trade agreements is not easily generalizable even within method families. However, there are promising hints of the welfare-reduction preferences of MNCs within the structural gravity results using the doc2vec and sentence transformer output.

3.5 Conclusion

The detection of corporate influence is an important issue as trade agreements are negotiated at country level, but because of the oversize lobbying power of large corporations, trade agreements may come to reflect their own aspirations and rent-seeking behavior, and not national interest. Studying how corporate influence has evolved over the years in trade policy continues to be a large issue to disentangle, but one that may be assisted by recent advances in natural language processing.

This paper examines the development of corporate influence over time with respect to trade agreements. Using a variety of techniques, from hand-coded indicators to deep learning methods, we find that the growth of corporate interests over time in trade agreements. Dictionary and topic modeling show an increase in topics that corporations historically lobby for. This observation is also coupled with deep learning methods that show an increase in similarity between trade agreements and 10-Ks over time.

Across methods, the impact of corporate influence on total trade flows leans positive. DESTA and topic models all agree on the positive impact of increased focus for goods and services, and investments on trade flows. Sentence transformers models trained explicitly on pairwise data also indicate the positive impact of increased 10-K similarity on total trade flows.

As a subquestion, this paper then examines whether trade agreement similarity to

multinational corporations' corporate language has a differential effect on trade flows. In both machine learning methods, doc2vec and sentence transformers, results from the structural gravity estimations suggest that multinational corporation language had a negative effect on trade flows. This negative association between similarity to multinational corporation language and trade flows lends credence to theories positing that multinational corporations may be associated with trade policies that enable rent-seeking behavior rather than freer trade.

While this paper constructs a starting point to examine the influence of corporations on trade agreements over time, there are clear next steps to follow. First, sentence transformers and to a lesser extent, doc2vec, are able to be more finely tuned to more accurately detect similarity between trade agreements and corporate documents. Second, with datasets such as LobbyView, researchers are better able to directly link firms to lobbying activities for or against trade agreements (Kim, 2018). However, as it is well-documented that lobbyists do write drafts of laws they wish to see enacted (Chang, 2013), the next lode of information to seek is the lobbying reports submitted by the lobbyists to the parties at hand. This will provide a clearer and more direct linkage to the corporate influence on trade agreements. Further Developing our understanding of unstructured language data will be key to researching these topics.

APPENDIX A

Appendix for Chapter I

A.1 Tables

Table A.1: LODES in New York City

Name	Statistic
Years	2002-2019
Number of Tracts	4899
Number of Block Groups	15417
Mean Fraction of High Income Per Block Group	0.42
Mean Distance Between Block Groups	3.78mi
Mean Duration Between Block Groups (Car)	22min
Mean Duration Between Block Groups (Walking)	198min

Statistics are at the block group level.

Table A.2: NYC PLUTO Summary Statistics

Year: 2002					
Variable	Mean	Median	STD	Min	Max
Lot Area (SqFt)	7909	2520	356924	0	2.10E+08
Building Area (Sqft)	5997	2000	62766	0	2.40E+07
Number of Floors	2.2	2	1.9	0	300
Built FAR	1.1	0.74	12	0	9471
Residential Max FAR	1.5	0.9	1.4	0	15
Max FAR in 1961	1.4	0.75	1.4	0	15
Year Built	1938	1930	27	1661	2020
Average Unit Size	1308	1181	1464	0	780000
Price Per SqFt	784	84	210048	0	9.70E+07
Residential Use Indicator	0.81	1	0.39	0	1
Year: 2020					
Variable	Mean	Median	STD	Min	Max
Lot Area (SqFt)	8618	2521	496611	0	2.10E+08
Building Area (Sqft)	6888	2000	151529	0	8.60E+07
Number of Floors	2.3	2	2	0	104
Built FAR	1.1	0.74	5.5	0	2484
Residential Max FAR	1.3	0.75	1.4	0	12
Max FAR in 1961	1.4	0.75	1.4	0	15
Year Built	1941	1931	31	1540	2020
Average Unit Size	1199	1100	919	0	339320
Price Per SqFt	554	145	132042	0	1.20E+08
Residential Use Indicator	0.83	1	0.38	0	1

Note: Statistics are at the tax lot level.

Table A.3: Dependent Variable: ln(Price per Square Foot)

	(1) OLS	(2) OLS	(3) IV - First Stage	(4) IV - Second Stage
Panel A: All Block Groups				
ln(FAR/Max FAR)	-0.1255	-0.1919		0.2117
	(0.0363)***	(0.0396)***		(0.26)
ln(FAR)	0.1441	0.1736		-0.2329
	(0.0613)**	(0.0927)*		(0.2759)
ln(FAR 2000) - ln(Max FAR 1961)			0.1376	
			(0.09)	
ln(Max FAR 1961)			0.068	
			(0.0292)***	
F-Stat			6.693	
N	6100	6100	6100	6100
Controls		X	X	X

Notes: All variables are first differences from 2002–2010 except for the instruments. Bandwidth for spatial HAC is equivalent to half mile, or about 10 North-South blocks in New York City. Regressions are at the block group level.

Table A.4: Dependent Variable: ln(Price per Square Foot)

	(1) OLS	(2) OLS	(3) IV - First Stage	(4) IV - Second Stage
Panel B: Block Groups with Zoning Change Prior to 2005				
ln(FAR/Max FAR)	-0.0618	-0.2975		0.4586
	(0.0868)	(0.1156)***		(0.9133)
ln(FAR)	0.0578	0.3222		-0.4355
	(0.1089)	(0.1383)***		(0.9153)
ln(FAR 2000) - ln(Max FAR 1961)			0.0263	
			(0.0128)***	
ln(Max FAR 1961)			0.0241	
			(0.0073)***	
F-Stat			5.447	
N	2000	2000	2000	2000
Controls		X	X	X

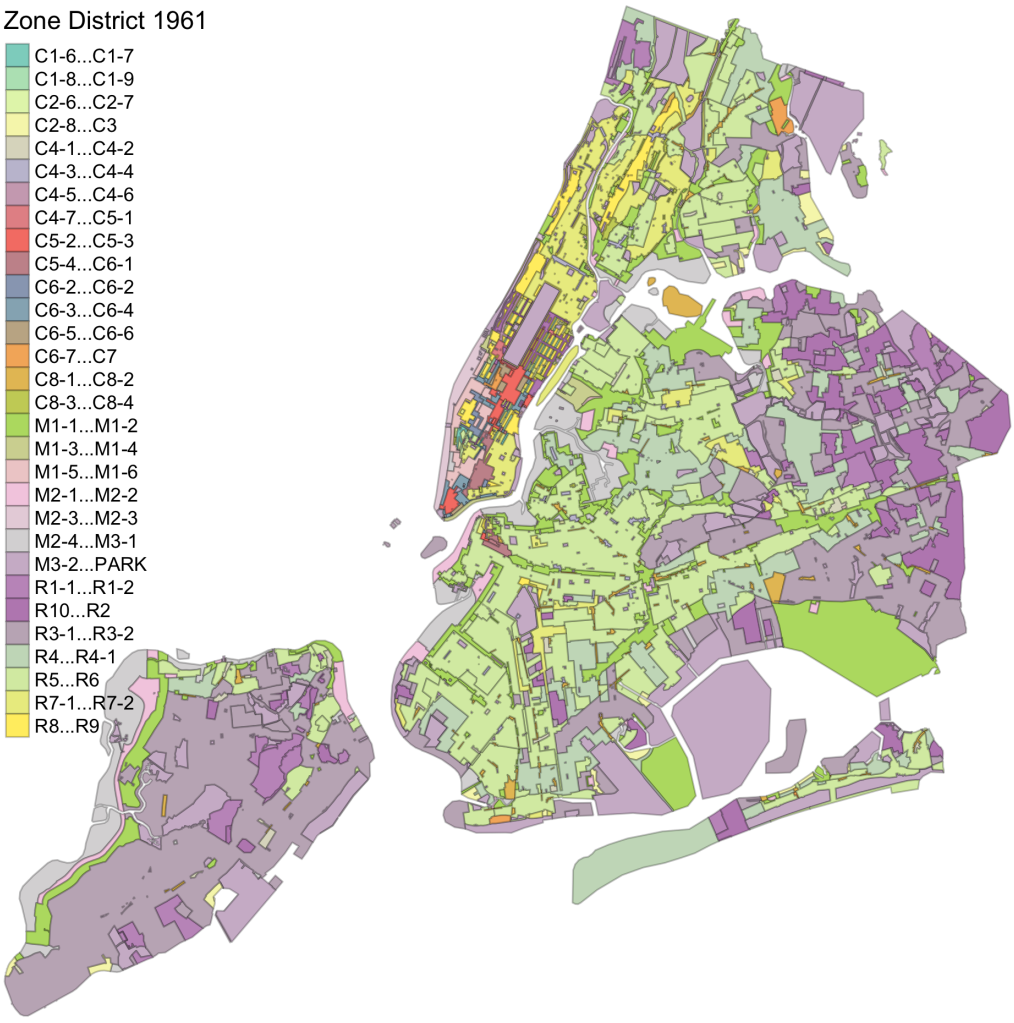
Notes: All variables are first differences from 2002–2010 except for the instruments. Bandwidth for spatial HAC is equivalent to half mile, or about 10 North-South blocks in New York City. Regressions are at the block group level.

Table A.5: Parameter Summary

Parameter	Description	Method	Data Source	Value
κ	Size of commuting costs	Literature	Ahlfeldt et al. (2015)	0.01
σ	Substitution elasticity between high and low-income workers	Literature	Card (2009)	1.429
β	Fraction of expenditure on housing	Calibration	Decennial Census 2000	0.29
T_h	Level of high-income worker productivity	Calibration	Decennial Census 2000	4.14
T_n	Level of low-income worker productivity	Calibration	Decennial Census 2000	1.00
$\bar{h}$	Minimum level of housing (sqft)	Calibration	NYC Zoning Regulations	80
ρ_h	Density spillover for high-income workers	Estimation	Decennial Census 2000	1.275
ρ_n	Density spillover for low-income workers	Estimation	Decennial Census 2000	0.897
η_h	Fréchet dispersion parameter for living preferences	Estimation	Decennial Census 2000, ACS 2016-2019, NYC PLUTO	1.082
η_n	Fréchet dispersion parameter for living preferences	Estimation	Decennial Census 2000, ACS 2016-2019, NYC PLUTO	0.581
θ_h	Sensitivity of commuting to time	Estimation	LODES 2002, OpenStreetMap	1.22
θ_n	Sensitivity of commuting to time	Estimation	LODES 2002, OpenStreetMap	1.42
ν	Elasticity of construction with respect to FAR	Estimation	Decennial Census 2000, ACS 2016-2019, NYC PLUTO	4.24

A.2 Figures

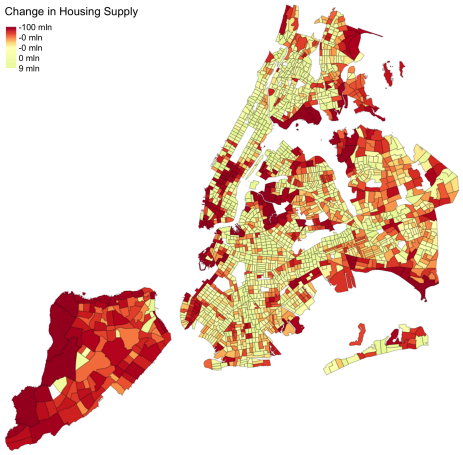
Figure A.1: Zoning Categories in 1961



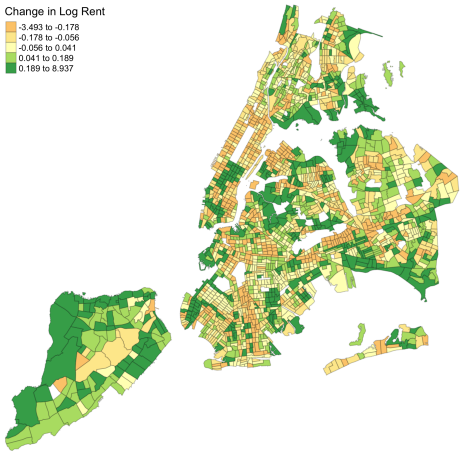
Notes: Data from the New York City Department of City Planning, Technical Review Division.

Figure A.2: Counterfactual 2

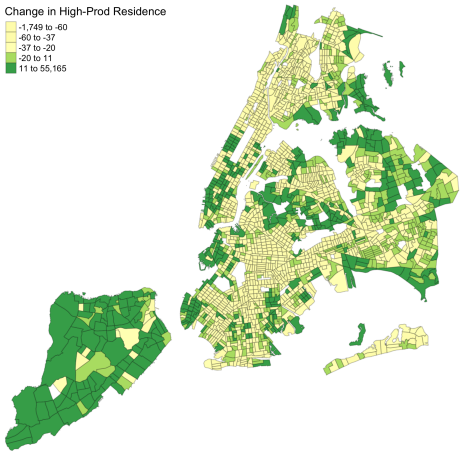
(a) Change in Housing Supply



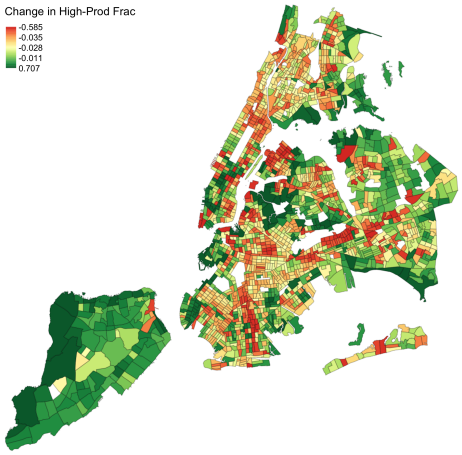
(b) Percent Change in Rent



(c) Change in High-Productivity Residents



(d) Change in High-Productivity Fraction



APPENDIX B

Appendix for Chapter II

B.1 Data

B.1.1 Constructing Flow Accumulation

The flow accumulation measure is the key simulated quantity that forms the basis of our instrumental variable strategy. It was introduced by Frachetti et al. (2017) and, in simple terms, is a sum of seasonal nomadic migrations over 20 human generations. We briefly describe its construction.

Frachetti et al. (2017) first construct their three base raster datasets. First, using Global Multi-Resolution Terrain Elevation Data (GMTED2010) with a pixel resolution of 30 arcseconds (about 1 km) within the inner Asian corridor, the researchers define the lowland and highland boundary at 750 m elevation. Next, they use a multispectral eMODIS image transformed into Normalized Difference Vegetation Index (NDVI) values from August 2008 at the peak of the grassland productivity to generate an NDVI raster averaged over seven days. Using this averaged NDVI raster and further information on fodder value and range productivity of biologically documented highland pasture types, the authors assign vegetation classes to each cell. Last, using

the two most fertile classes of vegetation, they also create an animal weight raster which corresponds to 16 animals per hectare, the average range capacity for inner Asian highland grasslands.

Botanical and archaeological evidence suggests that despite climate and geographical change over Central Asia in the past 4000 years, the grassland vegetation is mostly unchanged (Khotinskiy, 1984). Thus, modern imagery like eMODIS to calculate NDVI and modern measures of highland pasture productivities may be used to simulate historical flow accumulation measures.

Having constructed their base rasters, Frachetti et al. (2017) begin their recursive algorithm. The following steps are iterated 500 times to simulate 500 years of flow patterns. First, they randomly generate 5,000 lowland campsites. Using these campsites as sources and the vegetation classes generated previously as weights, the researchers apply the "Cost Distance" tool in ArcGIS to generate a cost distance raster. The cost distance raster gives the distance, for each grid cell, to the nearest campsite for each cell in the raster based on the least-accumulative cost. The cost distance raster, which was calculated using measures of pasture quality, and the animal weight raster are then used as inputs to the ArcGIS tool "Flow Accumulation", which produces a hypothetical count of animals flowing from the best pastures into each cell across the highlands region. Reverse flows are prohibited. This final raster is the flow accumulation measure (Figure 2.5) that we use in this paper.

B.2 Robustness Checks

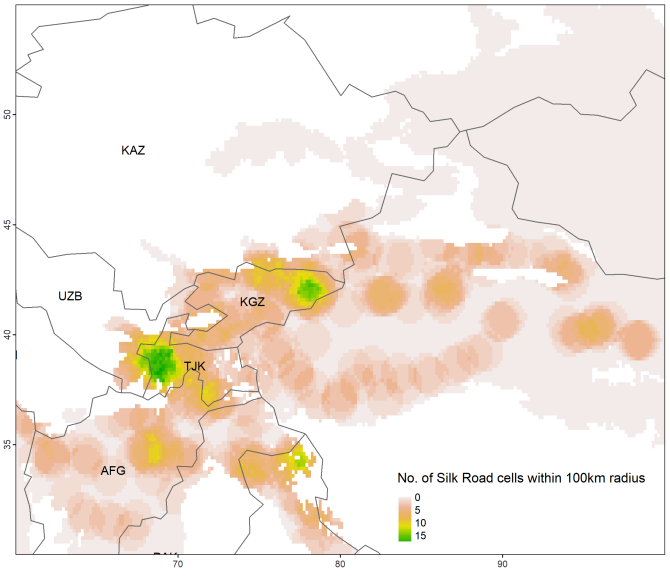
In this section, we present the additional robustness checks that we perform. All our robustness checks reported below use our baseline IV specification with the VIIRS night lights data as a measure of modern development.

Excluding outliers: We first exclude outliers from our analysis. To ensure that

the negative relationship between proximity to Silk Road site and modern development is not driven by the extreme values in the data, we exclude observations with VIIRS values in the top and bottom 2 percentiles from our analysis. The first two columns of Table B.1 present results. We still observe a negative and significant relationship between proximity to Silk Road site and night lights intensity after excluding the outliers.

Alternative measure of proximity to the Silk Road: In our analyses throughout the paper, we use the shortest distance to the nearest Silk Road site as a measure of proximity to the Silk Road. As a robustness check, we construct an alternative measure by constructing the variable c_{di}, which measures the number of Silk Road cells that can be reached within distance d from cell i. We call this the connectedness of cell i to the Silk Road.[1] We show this measure in Figure B.1, in which we use 100km as a distance cutoff.

Figure B.1: No. of Silk Road cells within 100km



Notes: This figure presents the number of Silk Road cells within 100km radius.

[1]This measure is similar to the one used in Bakker et al. (2019).

Table B.1: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Excluding outlier		Using connectivity measure		Using 0.083×0.083 degree cells		Using 4.167×4.167 degree cells	
	Distance to Silk Road site	Night lights (VIIRS)	Distance to Silk Road site	Night lights (VIIRS)	Distance to Silk Road site	Night lights (VIIRS)	Distance to Silk Road site	Night lights (VIIRS)
Distance to Silk Road site		-0.028***				-0.088***		-0.117***
		(0.007)				(0.025)		(0.028)
No. of Silk Road cells within 100 km				0.106***				
				(0.026)				
Distance to herding path	0.182***		-0.189***		0.181***		0.230***	
	(0.036)		(0.041)		(0.034)		(0.030)	
Caloric suitability index	-0.046***	-0.002***	0.051**	-0.005*	-0.033**	-0.004**	-0.061***	-0.006**
	(0.016)	(0.001)	(0.023)	(0.003)	(0.013)	(0.002)	(0.018)	(0.003)
Ruggedness	-0.053*	-0.005***	0.043	-0.018***	-0.038	-0.015***	-0.081***	-0.018***
	(0.028)	(0.001)	(0.027)	(0.004)	(0.024)	(0.003)	(0.030)	(0.006)
Precipitation	0.259***	0.007***	-0.042	0.006	0.249***	0.024***	0.292***	0.032**
	(0.050)	(0.003)	(0.071)	(0.009)	(0.048)	(0.009)	(0.051)	(0.015)
Distance to river	-0.025***	-0.002***	0.015	-0.004**	-0.027***	-0.005***	-0.028***	-0.008**
	(0.007)	(0.001)	(0.011)	(0.002)	(0.007)	(0.001)	(0.008)	(0.003)
Elevation	0.016	-0.012***	-0.181*	-0.025*	0.006	-0.035***	0.059	-0.035*
	(0.116)	(0.004)	(0.107)	(0.013)	(0.107)	(0.013)	(0.109)	(0.019)
Crop-Suitability-Indices	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NDVI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	25.785		21.488		27.870		60.212	
N	15,732	15,733	17,125	17,126	65,305	65,306	2,984	2,984

Notes: This table presents results from additional robustness check exercises. Distance to Silk Road site, distance to herding path, and number of Silk Road cells within 100km are standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We then use this connectivity measure in place of our standard distance to Silk Road site measure. The results are reported in the third and fourth columns of Table B.1. As expected, there is a positive relationship between the connectedness to the Silk Road and the intensity of the night lights. As the number of Silk Road cells within 100km increases by one standard deviation, the night lights intensity increases by 11.2%.

Alternative grid sizes: Our baseline unit of observation is a 0.167 degrees $\times$0.167 degrees grid cell. We now check if our results are robust to the sizes of the cells. We first construct an alternative dataset using 0.083 degrees $\times$ 0.083 degrees grid cells, which is half the width and height of our baseline grid cells. Columns 5 and 6 of Table B.1 report the results. We then consider an alternative dataset with grid cells that are 25 times the width and height of our baseline grid cells. The results are shown in columns 7 and 8 of Table B.1. Overall, our results are robust to different cell sizes.

Constructing instrument with different cutoffs: Due to computational feasibility, we set a flow value of 1.5 million as a cutoff to construct the instrument in this paper. We now check the robustness of our results to different cutoff values. Table B.2 shows the results when we construct the instrument using 1 million and 2 million as our cutoff values. Our result is robust to different cutoff values.

Table B.2: IV Results Using Different Cutoffs to Construct an Instrument

	1 million cutoff		2million cutoff	
	First stage Distance to Silk Road site	Second stage Night lights (VIIRS)	First stage Distance to Silk Road site	Second stage Night lights (VIIRS)
Distance to Silk Road site		-0.093***		-0.111***
		(0.025)		(0.033)
Distance to herding path	0.204***		0.171***	
	(0.029)		(0.036)	
Caloric suitability index	-0.042***	-0.004*	-0.042***	-0.005**
	(0.014)	(0.002)	(0.014)	(0.002)
Ruggedness	-0.047*	-0.017***	-0.049*	-0.018***
	(0.026)	(0.004)	(0.026)	(0.004)
Precipitation	0.258***	0.026**	0.256***	0.030**
	(0.047)	(0.011)	(0.048)	(0.012)
Distance to river	-0.025***	-0.005***	-0.025***	-0.005***
	(0.007)	(0.002)	(0.007)	(0.002)
Elevation	0.045	-0.039***	0.002	-0.043***
	(0.103)	(0.015)	(0.106)	(0.016)
Crop-Suitability-Indices	Yes	Yes	Yes	Yes
NDVI	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
F-stat	49.635		23.072	
N	17,125	17,125	17,125	17,125

Notes: This table presents results from additional robustness check exercises. Distance to Silk Road site and distance to herding path are standardized. All other variables are log-transformed. Standard errors are adjusted to allow for spatial clustering as in Conley (1999), with a bandwidth of 2 degrees using Bartlett kernel. Crop-suitability-indices include indices for wheat, rice, barley, flax, and millet. All regressions include a constant. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX C

Appendix for Chapter III

C.1 DESTA

C.1.1 Variables Included in the Corporate Influence Index

Services often require changes in domestic regulations which are driven more by corporate interest/expertise:

1. ser_specific
2. ser_list_negative
3. ser_nonestablishment

Protecting corporate investments:

1. inv_pre_est_oper
2. inv_est_oper
3. inv_post_est_oper
4. inv_acquisition_merger
5. inv_stand_treat_fps
6. inv_stand_treat_fps_defined
7. inv_compensation_expropriation
8. inv_compensation_strife

Protecting intellectual property/property rights:

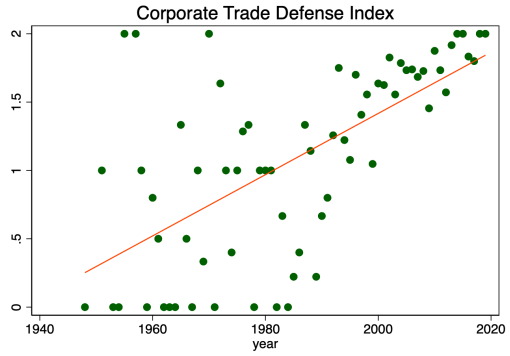
1. ipr_pharma
2. ipr_scope_substantial_dummy
3. ipr_specific_enforcement_dummy

Trade defense:

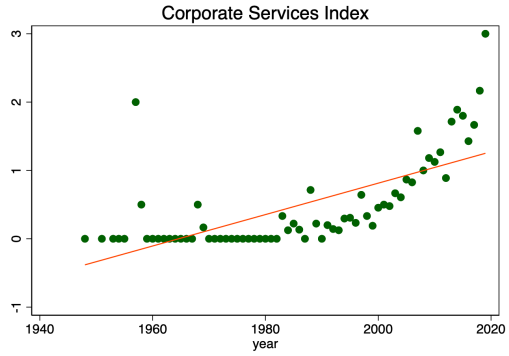
1. tr_ad_prov
2. tr_ad_allo

Figure C.1: Components of the Corporate Influence Index

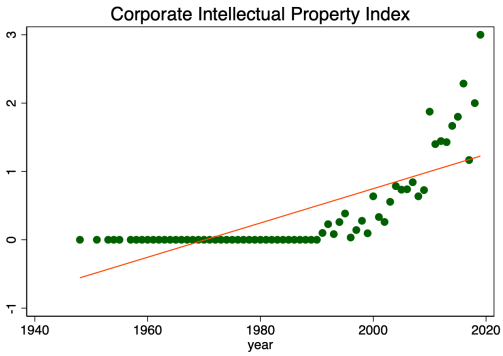
(a) Trade Defense



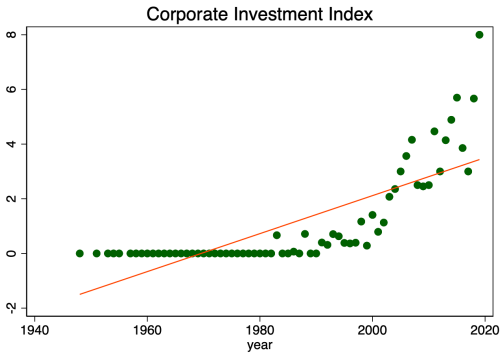
(b) Services



(c) Intellectual Property



(d) Investment



Notes: Data from DESTA.

C.2 Topic Modeling

C.2.1 LDA Topic Words

Protocol, Topic #0: contracting countries present agreed import quantitative protocol commodity foreign imported taxes transit signatory states similar non commission materials kingdom governments united control nomenclature central regulation concluded enterprises produced exportation industry regime concessions rates mentioned list european preferential reexport compliance raw based importation right activity use authorized organizations bilateral july injury

Community, Topic #1: states community member council association european common treaty financial technical convention regional transport companies programmes assistance training account union activities policies commission nationals resources energy social treatment legislation years sector projects management use institutions agricultural level ministers framework listed protocol arrangements specific members secretariat undertake provision laid research integration promotion

Goods, Topic #2: difficulties agricultural states laid practice undertakings exports introduced duty quantitative protocol intellectual examine safeguard abolished related notified cause grant adopt aid basic listed discrimination solution years conformity rise proper decide remedy duration markets incompatible schedule procurement policies january fiscal exporting essential european certain functioning progressively reduced technical restriction injury prejudice

Services/Investment, Topic #3: good service financial measure suppliers investment person treatment entity authority does procurement law non telecommunications enterprise use supplier days related include panel investor producer mer technical used laws tribunal pursuant available material claim persons right materials commission require notice written subparagraph authorities access reasonable duty extent agree maintain level section

Arbitration, Topic #4: measure service arbitral dispute good days laws tribunal panel conformity safeguard arbitration specific technical persons investment agree commitments authority pursuant natural treatment materials assessment authorities person standards bilateral sanitary law mer agreed mutually non related unless importing suppliers intellectual competent exporting review activities supply report facilitate term used include operation

C.2.2 STM Topic Words

Protocol, Topic 1 Top Words:

- Highest Prob: union, reservation, official, transport, type, service, journal
- FREX: transcription, residency, characters, provincial, httpwww, lawyers, estate
- Lift: districts, phat, abandoning, adici, adjudications, adores, affaire
- Score: unbound, journal, directive, reservation, provincial, cooked, inspectorate

Goods, Topic 2 Top Words:

- Highest Prob: exceeding, manufacture, parts, weight, machines, agriculture, fibers
- FREX: plates, rods, alloy, rolls, omer, rectangular, plated
- Lift: acct, adsorption, aerosols, agon, alanine, alkyl, amaranth
- Score: salts, thickness, fabrics, rolled, fibers, yarn, machines

Contracts, Topic 3 Top Words:

- Highest Prob: contracting, laid, undertakings, introduced, basic, abolished, proper
- FREX: reductions, abolished, incompatible, threatens, contracting, introduced, arab
- Lift: abolishes, abolishing, arenas, arithmetical, ashgabat, belorussian, bildt
- Score: contracting, abolished, csfr, objected, repayment, harmoniously, disturbances

Community, Topic 4 Top Words:

- Highest Prob: member, community, council, treaty, union, association, commission
- FREX: renumbered, acqu, landlocked, ambassadors, moved, parliaments, majesty

- Lift: abiding, absorbed, abstaining, affinities, ambassadors, apartheid, appraisals
- Score: community, renumbered, acqu, member, cotonou, ambassadors, council

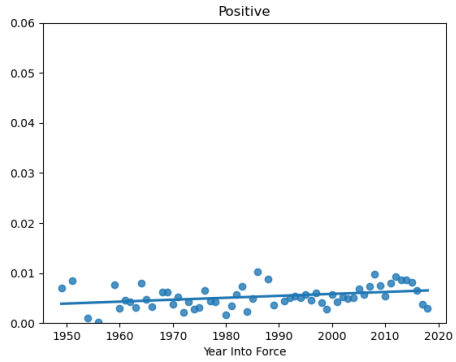
Services/Investment, Topic 5 Top Words:

- Highest Prob: service, investment, authority, person, suppliers, laws, otherwise
- FREX: disputing, producer, arbitral, importer, expropriation, qualifies, investor
- Lift: academies, accredits, accumulate, aconcagua, acquittals, admits, adoptive
- Score: disputing, investor, procuring, supplier, arbitral, claim, certainty

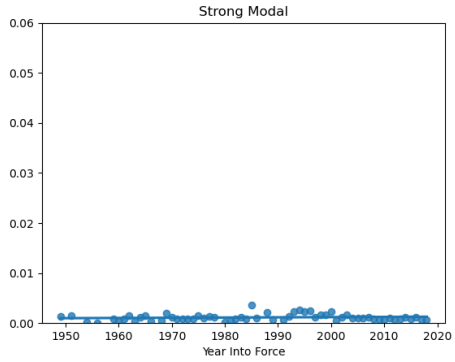
C.3 Sentiment Analysis

Figure C.2: Trade Agreements: Sentiment Word Lists Mean Proportions Over Time

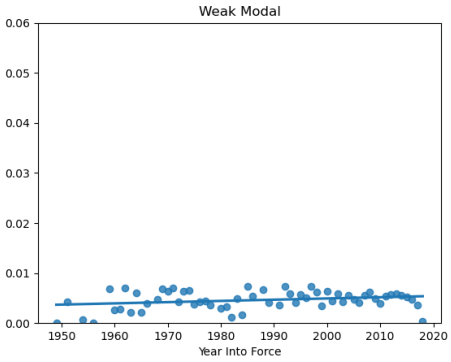
(a) Positive



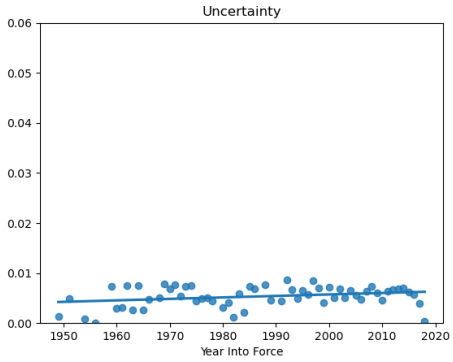
(b) Strong Modal



(c) Weak Modal



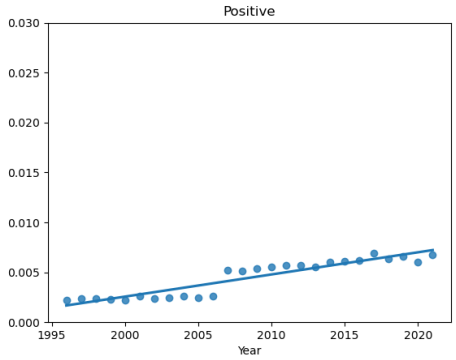
(d) Uncertain



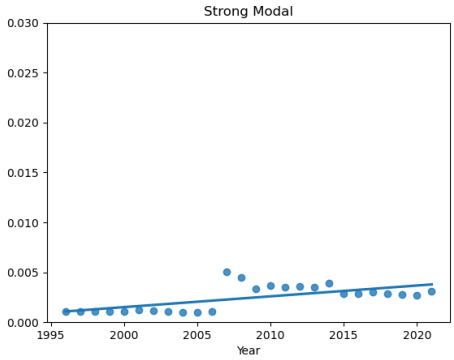
Notes: Proportions are calculated by taking the word list, counting the number of appearances in each document, and then dividing the count by the total word count of the document. Agreements are aggregated by year using a simple mean.

Figure C.3: 10-Ks: Sentiment Word Lists Mean Proportions Over Time

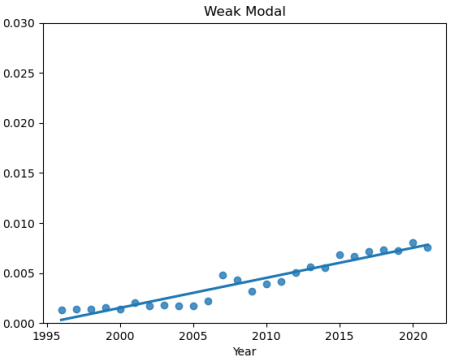
(a) Positive



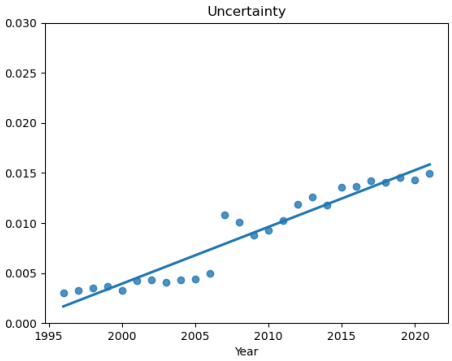
(b) Strong Modal



(c) Weak Modal



(d) Uncertain



Notes: Proportions are calculated by taking the word list, counting the number of appearances in each document, and then dividing the count by the total word count of the document. 10-Ks are aggregated by year using a simple mean.

Table C.1: Correlations Between Sentiment Word List and LDA Topic Proportions

	Pro. Litigious	Pro. Constraining	Pro. Negative	Protocol	Community	Goods	Serv/Inv	Arbitration
Pro. Litigious	1.000							
Pro. Constraining	0.897	1.000						
Pro. Negative	0.913	0.905	1.000					
Protocol	0.179	0.147	0.171	1.000				
Community	0.350	0.237	0.214	-0.053	1.000			
Goods	-0.563	-0.468	-0.593	-0.256	-0.620	1.000		
Serv/Inv	0.386	0.367	0.539	-0.098	0.120	-0.718	1.000	
Arbitration	0.178	0.141	0.176	0.091	0.046	-0.309	-0.023	1.000

Table C.2: Correlations Between Sentiment Word List and STM Topic Proportions

	Pro. Litigious	Pro. Constraining	Pro. Negative	Protocol	Goods	Contracts	Community	Serv/Inv
Pro. Litigious	1.000							
Pro. Constraining	0.897	1.000						
Pro. Negative	0.913	0.905	1.000					
Protocol	-0.034	-0.030	-0.040	1.000				
Goods	-0.065	-0.055	-0.068	-0.034	1.000			
Contracts	0.103	0.096	0.131	-0.162	-0.349	1.000		
Community	-0.031	-0.036	-0.049	-0.210	-0.321	-0.224	1.000	
Serv/Inv	0.010	0.010	0.008	0.000	-0.289	-0.314	-0.419	1.000

C.4 doc2vec

Table C.3: Gravity Equations with Doc2vec Similarity Cosine Scores

	Baseline Trade Flows (1)	Doc2vec (2)	Baseline Manuf. Flows (3)	Doc2vec (4)
PTA	0.2488	0.1342	0.3183	0.2166
	(0.034)***	(0.047)***	(0.034)***	
PTA x Log Similarity		-0.1042		-0.0974
		(0.020)***		(0.020)***
PTA x Log Similarity MNC		0.1013		0.1121
		(0.016)***		(0.016)***
Log Distance	-0.7109	-0.7163	-0.6931	-0.6942
	(0.018)***	(0.018)***	(0.017)***	(0.017)***
N	83196	83198	83196	83198
Common Characteristics	X	X	X	X
Year Fixed Effects	X	X	X	X
Country Fixed Effects	X	X	X	X

[1] Standard errors in parentheses.
[2] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[3] Similarity scores are based on 1996 10-K forms.

BIBLIOGRAPHY

Acemoglu, Daron and Pascual Restrepo, "Tasks, automation, and the rise in US wage inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

Acosta, Camilo, "The Incidence of Land Use Regulations," 2020.

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Ahmad, Zofia and Luke Chicoine, "Silk Roads to Riches: Persistence Along an Ancient Trade Network," *Available at SSRN 3760490*, 2021.

Alford, Roger P, "The convergence of international trade and investment arbitration," *Santa Clara J. Int'l L.*, 2013, *12*, 35.

Alschner, Wolfgang, Julia Seiermann, and Dmitriy Skougarevskiy, "Text-as-data analysis of preferential trade agreements: Mapping the PTA landscape," 2017.

Arnott, Richard J and James G MacKinnon, "Measuring the Costs of Height Restrictions with a General Equilibrium Model," *Regional Science and Urban Economics*, 1977, *7* (4), 359–375.

Aubekerov, B.Z., S.A. Nigmatova, and M.D. Frachetti, "Geomorphological Particulars in the Region of the Archaeological Monument Begash, Northern Dzhungar Alatau [in Russian]," *Aktual'nye Problemy Geostistem Aridnikh Territorii*, 2003, pp. 287–289.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "Concentrating on the Fall of the Labor Share," *American Economic Review*, 2017, *107* (5), 180–185.

Baccini, Leonardo, "The economics and politics of preferential trade agreements," *Annual Review of Political Science*, 2019, *22*, 75–92.

Bakker, Jan David, Stephan Maurer, Jörn-Steffen Pischke, and Ferdinand Rauch, "Of Mice and Merchants: Connectedness and the Location of Economic Activity in the Iron Age," *Review of Economics and Statistics*, 2019, pp. 1–44.

Baragwanath, Kathryn, Ran Goldblatt, Gordon Hanson, and Amit K Khandelwal, "Detecting Urban Markets with Satellite Imagery: an Application to India," *Journal of Urban Economics*, 2019, p. 103173.

Barjamovic, Gojko, Thomas Chaney, Kerem Coşar, and Ali Hortaçsu, "Trade, Merchants, and the Lost Cities of the Bronze Age," *The Quarterly Journal of Economics*, 2019, *134* (3), 1455–1503.

Baum-Snow, Nathaniel and Lu Han, "The Microgeography of Housing Supply," *University of Toronto*, 2019.

_ , Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang, "Roads, Railroads, and Decentralization of Chinese Cities," *Review of Economics and Statistics*, 2017, *99* (3), 435–448.

Been, Vicki, Josiah Madar, and Simon McDonnell, *Underused Lots in New York City*, Lincoln Institute of Land Policy., 2009.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "The margins of US trade," *American Economic Review*, 2009, *99* (2), 487–493.

Bertaud, Alain and Jan K Brueckner, "Analyzing building-height restrictions: predicted impacts and welfare costs," *Regional Science and Urban Economics*, 2005, *35* (2), 109–125.

Blanga-Gubbay, Michael, Paola Conconi, and Mathieu Parenti, "Lobbying For Globalization," *Unpublished manuscript*, 2023.

Bleakley, Hoyt and Jeffrey Lin, "Portage and Path Dependence," *The quarterly journal of economics*, 2012, *127* (2), 587–644.

Blei, David M, Andrew Y Ng, and Michael I Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, 2003, *3* (Jan), 993–1022.

Blomquist, Sören and Matz Dahlberg, "Small sample properties of LIML and jackknife IV estimators: experiments with weak instruments," *Journal of Applied Econometrics*, 1999, *14* (1), 69–88.

Borchert, Ingo, Mario Larch, Serge Shikher, and Yoto V Yotov, "The international trade and production database for estimation (ITPD-E)," *International Economics*, 2021, *166*, 140–166.

_ , _ , _ , and _ , "The International Trade and Production Database for Estimation-Release 2," 2022.

Brueckner, Jan K, Shihe Fu, Yizhen Gu, and Junfu Zhang, "Measuring the stringency of land use regulation: The case of China's building height limits," *Review of Economics and Statistics*, 2017, *99* (4), 663–677.

Bruederle, Anna and Roland Hodler, "Nighttime Lights as a Proxy for Human Development at the Local Level," *PloS one*, 2018, *13* (9), e0202231.

Bunten, Devin, "Is the rent too high? aggregate implications of local land-use regulation," 2017.

Cameron, A Colin and Pravin K Trivedi, *Microeconometrics: Methods and Applications*, Cambridge university press, 2005.

Card, David, "Immigration and inequality," *American Economic Review*, 2009, *99* (2), 1–21.

Chang, Alisa, "When Lobbyists Literally Write the Bill," *NPR*, 2013.

Chau, Kwong-Wing, Siu Kei Wong, Yung Yau, and AKC Yeung, "Determining Optimal Building Height," *Urban Studies*, 2007, *44* (3), 591–607.

Chor, Davin and Bingjing Li, "Illuminating the Effects of the US-China Tariff War on China's Economy," Technical Report, National Bureau of Economic Research 2021.

Ciolek, T.M., "Old World Trade Routes (OWTRAD) Project," 2014.

City, New York, "PLANYC: A Greener, Greater New York," 2007.

Conley, Timothy G, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

_ , "GMM Estimation with Cross Sectional Dependence," *Journal of econometrics*, 1999, *92* (1), 1–45.

Conte, Maddalena, Pierre Cotterlaz, and Thierry Mayer, "The CEPII Gravity Database Highlights," 2022.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income growth and the Distributional Effects of Urban Spatial Sorting," Technical Report, National Bureau of Economic Research 2019.

Cribb, Roger, *Nomads in Archaeology*, Cambridge University Press, 2004.

Dalgaard, Carl-Johan, Nicolai Kaarsen, Ola Olsson, and Pablo Selaya, "Roman Roads to Prosperity: Persistence and Non-Persistence of Public Goods Provision," 2018.

Davis, Jenna, "How do upzonings impact neighborhood demographic change? Examining the link between land use policy and gentrification in New York City," *Land Use Policy*, 2021, *103*, 105347.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

Dingel, Jonathan I, Antonio Miscio, and Donald R Davis, "Cities, Lights, and Skills in Developing Economies," *Journal of Urban Economics*, 2019, p. 103174.

Dür, Andreas, Leonardo Baccini, and Manfred Elsig, "The design of international trade agreements: Introducing a new dataset," *The Review of International Organizations*, 2014, *9*, 353–375.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner, "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 2014, *81* (2), 681–724.

Engstrom, Ryan, Jonathan Samuel Hersh, and David Locke Newhouse, "Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being," *World Bank Policy Research Working Paper*, 2017, (8284).

Frachetti, Michael D., "Variability and Dynamic Landscapes of Mobile Pastoralism in Ethnography and Prehistory," in H. Barnard and W. Wendrich, eds., *The Archaeology of Mobility: Nomads in the Old and in the New World*, Los Angeles: Cotsen Institute of Archaeology, UCLA, 2008, chapter 17, pp. 366–396.

Frachetti, Michael D, C Evan Smith, Cynthia M Traub, and Tim Williams, "Nomadic Ecology Shaped the Highland Geography of Asia's Silk Roads," *Nature*, 2017, *543* (7644), 193–198.

Freemark, Yonah, "Upzoning Chicago: Impacts of a zoning reform on property values and housing construction," *Urban Affairs Review*, 2020, *56* (3), 758–789.

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal, "Suburbanization and Highways in Spain When the Romans and the Bourbons Still Shape Its Cities," *Journal of Urban Economics*, 2015, *85*, 52–67.

Gibson, John, Susan Olivia, Geua Boe-Gibson, and Chao Li, "Which Night Lights Data Should We Use in Economics, and Where?," *Journal of Development Economics*, 2021, *149*, 102602.

Glaeser, Edward L, Joseph Gyourko, and Raven Saks, "Why is Manhattan so expensive? Regulation and the rise in housing prices," *The Journal of Law and Economics*, 2005, *48* (2), 331–369.

Gorbunova, Natalya, "Traditional Movements of Nomadic Pastoralists and the Role of Seasonal Migraitons in the Formation of Ancient Trade Routes in Central Asia," *Silk Road Art and Archaeology 3*, 1993, pp. 1–10.

Gyourko, Joseph and Raven Molloy, "Regulation and housing supply," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 1289–1337.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring Economic Growth from Outer Space," *American economic review*, 2012, *102* (2), 994–1028.

Hill, Daniel J, "Climate change and the rise of the Central Asian Silk Roads," *Socio-environmental dynamics along the historical Silk Road*, 2019, pp. 247–259.

Hoelzlein, Mattias, "Two-sided sorting and spatial inequality in cities," 2020.

Hoffman, Nicholas Von, "1980-1989: The Decline and Fall of the New York Empire," *Village Voice*, 1990.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The Allocation of Talent and US Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Hu, Yingyao and Jiaxiong Yao, "Illuminating Economic Growth," *Journal of Econometrics*, 2021.

IIASA/FAO, "Global Agro-ecological Zones (GAEZ v3.0)," 2012.

III, Robert N Spengler, *Fruit From the Sands: the Silk Road Origins of the Foods We Eat*, University of California Press, 2019.

Jedwab, Remi and Alexander Moradi, "The Permanent Economic Effects of Transportation Revolutions in Poor Countries: Evidence from Africa," *Review of Economics and Statistics*, 2015.

Kelejian, Harry H and Ingmar R Prucha, "HAC Estimation in a Spatial Framework," *Journal of Econometrics*, 2007, *140* (1), 131–154.

Kelly, Morgan, "The Standard Errors of Persistence," 2019.

_ , "Understanding Persistence," 2020.

Khotinskiy, N.A., "Holocene Vegetation History," *Late Quarternary Environments of the Soviet Union*, 1984, pp. 179–200.

Kim, In Song, "Political cleavages within industry: Firm-level lobbying for trade liberalization," *American Political Science Review*, 2017, *111* (1), 1–20.

_ , "Lobbyview: Firm-level lobbying & congressional bills database," *Unpublished manuscript, MIT, Cambridge, MA. http://web. mit. edu/insong/www/pdf/lobbyview. pdf Google Scholar Article Location*, 2018.

_ **and Helen V Milner**, "Multinational corporations and their influence through lobbying on foreign policy," *Multinational Corporations in a Changing Global Economy*, 2019, pp. 497–536.

Lambert, Dayton M, Raymond JGM Florax, and Seong-Hoon Cho, "Bandwidth Selection for Spatial HAC and Other Robust Covariance Estimators," Technical Report 2008.

Le, Quoc and Tomas Mikolov, "Distributed representations of sentences and documents," in "International conference on machine learning" PMLR 2014, pp. 1188–1196.

Levy, David L and Aseem Prakash, "Bargains old and new: Multinational corporations in global governance," *Business and politics*, 2003, *5* (2), 131–150.

Lipton, Eric and David Yaffe-Bellany, "Crypto Industry Helps Write, and Pass, Its Own Agenda in State Capitols," *New York Times*, 2022.

Loughran, Tim and Bill McDonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of finance*, 2011, *66* (1), 35–65.

Lowe, Matt, "Night Lights and ArcGis: A Brief Guide," *Avaliable online: http://economics. mit. edu/files/8945 (accessed on 22 August 2014)*, 2014.

Manger, Mark S, *Investing in protection: The politics of preferential trade agreements between north and south*, Cambridge University Press, 2009.

Michaels, Guy and Ferdinand Rauch, "Resetting the Urban Network: 117–2012," *The Economic Journal*, 2018, *128* (608), 378–412.

Michalopoulos, Stelios and Elias Papaioannou, "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 2013, *81* (1), 113–152.

Millward, James A, *The Silk Road: a Very Short Introduction*, Oxford University Press, 2013.

Mohamed, Rayman, "Why do residential developers prefer large exurban lots? Infrastructure costs and exurban development," *Environment and Planning B: Planning and Design*, 2009, *36* (1), 12–29.

Nunn, Nathan and Diego Puga, "Ruggedness: The blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

Rauch, James E, "Does History Matter Only When It Matters Little? The Case of City-Industry Location," *The Quarterly Journal of Economics*, 1993, *108* (3), 843–867.

Reimers, Nils and Iryna Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in "Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing" Association for Computational Linguistics 11 2019.

Roberts, Margaret E, Brandon M Stewart, and Dustin Tingley, "Stm: An R package for structural topic models," *Journal of Statistical Software*, 2019, *91*, 1–40.

Rodrik, Dani, "What do trade agreements really do?," *Journal of economic perspectives*, 2018, *32* (2), 73–90.

Stein, Rona B., "New York City's Economy in 1980," 1981.

Stiglitz, Joseph, "The secret corporate takeover of trade agreements," *The Guardian*, 2015.

Tan, Ya, Zhi Wang, and Qinghua Zhang, "Land-use regulation and the intensive margin of housing supply," *Journal of Urban Economics*, 2020, *115*, 103199.

Tobler, Waldo R, *Spectral Analysis of Spatial Series*, Library Photographic Service, U. of California, 1970.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's Transmilenio," 2019.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention Is All You Need," *CoRR*, 2017, *abs/1706.03762*.

von der Mosel, Julian, Alexander Trautsch, and Steffen Herbold, "On the validity of pre-trained transformers for natural language processing in the software engineering domain," 2022.

Wendrich, Willeke and Hans Barnard, "The Archaeology of Mobility: Definitions and Research Approaches," *The archaeology of mobility: Old World and New World nomadism*, 2008, pp. 1–21.

Williams, Tim, "The Silk Roads: An ICOMOS Thematic Study," 2014.

Wu, Yilin, "The Impact of Density Zoning on Neighborhood Segregation: Evidence from Mid-century New York City," 2019.