UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Regional Economic Development in the United States

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Youngjin Song

2018

ABSTRACT OF THE DISSERTATION

Essays on Regional Economic Development in the United States

by

Youngjin Song
Doctor of Philosophy in Management
University of California, Los Angeles, 2018
Professor Paola Giuliano, Chair

My dissertation consists of three essays on U.S. regional economics. The aim of the research is to understand the variations in economic outcomes across regions and how those persistent differences can be reduced. In particular, I study the geographic patterns and the determinants of internal migration, and examine the development of social capital.

In the first chapter, I document U.S. internal migration between the period from 1960-2000 using a newly collected data set. I find that the recent decline in migration is driven by lower migration across states, while within state migration has increased during the observed periods. To explain this, I use a gravity framework and estimate the effect of state borders on migration flows. I find that the border effect is strongly significant, and within state migration is 3.2 times higher than across state migration. Furthermore, the border effect has increased from 2.7 in 1960 to 3.6 in 2000. I show that the differences in social and economic characteristics between areas contribute to a larger state border effect, and the increase of the border effect over time is associated with the rising differences in house prices as states implement more restrictive land use regulations. I find that this is due to lower in-migration in states that are highly regulated in land uses. For high income destinations, the rise in regulations can explain all of the increase in the border effect.

In the second chapter, I examine the effect of political differences on migration decisions and provide empirical evidences of partisan geographic sorting among American migrants. Using presidential election returns and the same migration data for the period from 1960-2000, I show that political differences between areas decrease migration, as Americans prefer to live in areas with similar ideology and political views. I find that there are a lower number of migrants between places that supported different political parties, and migration flows decrease as differences in vote shares increase. In addition, I do not find evidence that Americans are increasingly sorting by partisanship over time, which is previously known in the literature as the "big sort" hypothesis.

The last chapter studies the level of social trust across regions in the United States, and how it arises using natural disasters as exogenous shocks between 1973 and 2010. Every year, the United States is hit by natural disasters that take away lives and cause property damage. In the event of a natural disaster, the victims are more likely to have increased interaction with others. If the experience of increased social interaction is positive, this can positively affect the level of social trust. To test this, I combine two U.S. survey data sets with the natural disaster data, and find that the individuals who have recently experienced natural disasters are more likely to have higher level of trust. One standard deviation increase in natural disasters is associated with an increase of 0.014 standard deviations in trust. I also provide evidence of positive social experiences by showing that natural disasters are associated with an increase in cooperation with neighbors.

The dissertation of Youngjin Song is approved.

Sarah J Reber

Romain T. Wacziarg

Christian Dippel

Paola Giuliano, Committee Chair

University of California, Los Angeles

2018

To My Family

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor Paola Giuliano. Her continuous support, guidance, and advice were instrumental in my growth. I am truly grateful to have an exceptional mentor, whose insight, patience, and encouragement helped me move forward throughout my studies.

I am also extremely grateful to my committee members Christian Dippel, Sarah Reber, and Romain Wacziarg for their invaluable guidance and insightful comments. Their commitment and willingness to provide support were inspirational. Completion of this thesis would not have been possible without their help.

I also wish to acknowledge other UCLA faculty members Ricardo Perez-Truglia, Nico Voigtländer, Melanie Wasserman, my colleagues Andrea Di Miceli, Mikhail Poyker, Yang Yang, all of my GEM office mates, and the seminar participants at UCLA GEM Brown Bag and PAA 2017, for helpful comments and suggestions. All remaining errors are mine.

A special thanks to the loved ones in my life. I would like to thank my parents, whose love, prayers, patience, and support made everything possible. I am grateful to John for being by my side through it all.

Lastly, thank you, God.

Vita

2009	Bachelor in Business Administration Korea University Seoul, South Korea
2011	M.Sc. in Economics London School of Economic and Political Science London, U.K.
2013-2018	Graduate Student Researcher UCLA Anderson School of Management

CHAPTER 1

Internal Migration in the United States 1960-2000: The Role of the State Border

1.1 Introduction

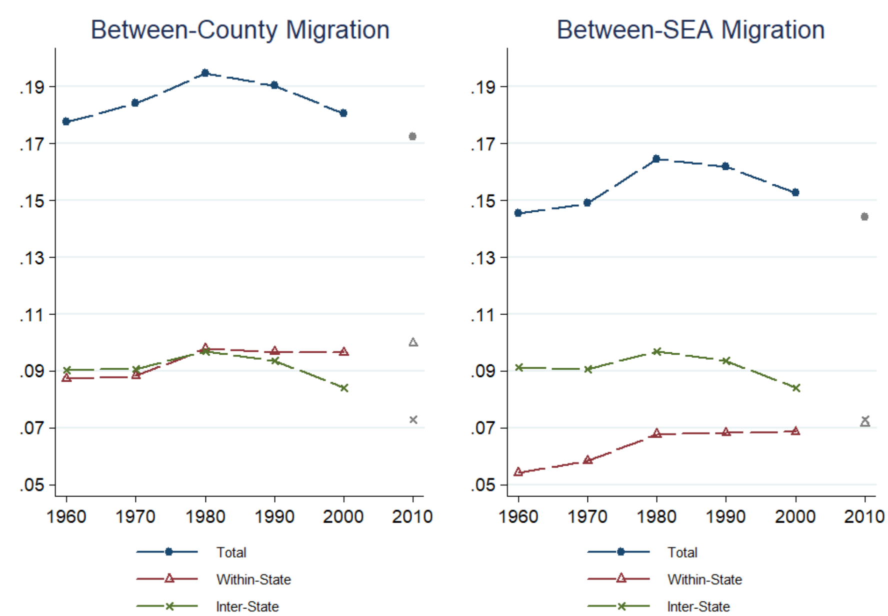
High internal migration is one of the distinctive features of the United States (Greenwood 1997), with 1.5 percent of the population moving across states annually. During recent decades, however, researchers have found a universal decline in migration across multiple demographic and socioeconomic groups (Molloy et al. 2011). They also find that cyclical factors, such as recessive housing market and economic downturn, fail to explain the decrease.

The dramatic slowdown in mobility is a puzzle and has triggered substantial research, but has yet to be adequately explained. One of the reasons this question remains unanswered is due to the limited availability of disaggregated migration data at the sub-state level prior to 1980, which has restricted researchers to observing cross-state moves or moves for recent decades only.[1] This trade-off between geography and period may lead to erroneous conclusions if, for example, the decline in migration is an extension of a previously existing trend.

[1]Directly observed migration data between substate destination and origin pairs were not available prior to 1980, but researchers have inferred lifetime migration from state of birth and the current state of residence, or 5-year migration using households with 5-year-old (Rosenbloom and Sundstrom 2004) to extend data.

I address this by using a newly constructed bilateral data of 5-year migration flows between State Economic Areas –a group of counties contained within states– starting from 1960 to 2000. The disaggregated data collected from the decennial Census Published Volumes allows for decomposing the declining trend in migration, and shows that this trend is absent in migration within states. By plotting the breakdown of the aggregate trend in migration, Figure 1 demonstrates the usefulness of sub-state migration data and provides the key motivation for my analysis of the border effect. Between-State Economic Area (SEA) moves on the right are a subset of between-county moves on the left, and for both graphs, the total migration rate is decomposed into movements within and across states.

Figure 1.1: Internal Migration Rates



Note: Author's calculation based on the 5-year migration data from the decennial Censuses and American Community Survey (ACS). Migrants are in shares of the total population above 5 years old. Data point in 2010 is an approximation of the 5-year migration rate, using the average annual migration rates from the ACS 2006-2011. The county migration data is only available at the aggregate level before 1980.

There are three notable patterns to highlight from Figure 1. First, consistent with the literature, it is evident that total migration rate has declined since 1980.

Using SEA data, the rate drops from 16.4 percent in 1980 to 15.3 percent in 2000, indicating that 1.1 percentage point of the total population moved less over time. The extended migration data shows that the aggregate migration follows a hump-shaped trend, and was increasing prior to 1980. The breakdown of total migration into within- and cross-state moves demonstrates the second fact that the recent decline is entirely driven by the fall in interstate migration. Compared to 1980, migration rate across states falls by 1.3 percentage point from 9.7 to 8.4 percent in 2000, whereas migration within states has increased by 0.1 percentage point. Thus, follows the third fact, that within-state migration and cross-state migration are progressively moving in opposite directions as more migrants move within states and increasingly less across states over time.[2]

These patterns suggest an increasing preference for moves within state borders, and this study contributes to the literature by providing answers to the questions: (1) how large is the state border effect for the U.S. internal migration? (2) how does the border effect vary across time? and (3) why did the border effect increase?

First, as intranational migrants are not subject to formal or informal border barriers, such as visa policies or language differences, and the states in the U.S. are highly integrated, the reason for the presence of the border effect at the state line is not apparent. In order to measure the significance of this "home state bias" in migration, I use the gravity framework. For estimation, I employ the Poisson Pseudo-maximum Likelihood (PPML) estimator following Silva and Tenreyro (2006), which performs well in the presence of heteroskedastic error terms and a large number of zero flows in the dependent variable. This is the first paper, to my

[2]Molloy, et al. (2013) use the Current Population Survey (CPS) annual migration data and finds decline in migration between 1980-2010 for all geographical levels, across state, within state, and also within county. Intra-county drops from 14 percent to 10 percent, and intra-state drops from 3.5 to 2 percent. There may be few possible reasons why my findings are different. The same authors do state in their 2011 paper that the CPS overstates migration decline compared to other data sources. The CPS is a much smaller sample compared to the decennial Census and observes migration in the previous year. As annual migration picks up more temporal moves, it may be that there is less repeat migration.

knowledge, to present a measure of the state border effect for the U.S. domestic migration starting from 1960, which is made possible due to the newly collected data that includes within-state migration flows. I find that a significant border effect exists for domestic migration and it is robust across different specifications. The size of the border effect implies that within state migration is 3.2 times higher than migration across states.

To answer the second question, I use the panel structure of the data and estimate state border effects over 5 decades. Given the lower direct and indirect costs of moving due to improved transporation and technology, the border effect is expected to have fallen over time. On the contrary, I show that consistent with the patterns in Figure 1, the state border effect has grown larger over time. The state border effect not only persists throughout the 1960-2000 period, but the effect of the border increases from 2.7 in 1960 to 3.6 in 2000.

Third, to explain the border effect, I evaluate how the border effect changes as destination and origin areas differ across various socioeconomic characteristics. I find that the dissimilarities between pairs contribute to increasing the border 'barrier', hindering migrants from moving across states. For example, between destination and origin areas that have the same median family income, the border effect falls by 37 percent from 3.2 to 2. Over time, I find that for a given pair of SEAs, the border effect grows larger as the dissimilarities increase. In particular, I show that the increase in the border effect rises with the house price differences over time.

The change of regulatory climate toward constraining housing supply has increased house price dispersion in the last few decades (Glaeser et al. 2005). I use the land use regulation data constructed by Ganong and Shoag (2017), who show that the land use regulations reduce net migration and regional income convergence. This study tests the effect of land use regulations on the state borders, providing empirical support for the link between regulations and migration.

Specifically, I find that the regulations negatively affect in-migration to states. I show that the states with restrictive land uses have higher border effects, and particularly for high income destinations, the border trend is completely explained by the rise in regulations. Thus, there are increasingly less migrants moving across states to high income areas in land use restricted states, as limited housing supply reduces the housing affordability in the area.

A comparison of two areas from the data, LA/Orange County in California and Dallas County in Texas illustrates this idea well. Between 1960 and 2000, the total 5-year migration inflow to LA/Orange County dropped by almost 30 percent from 1,044,545 to 756,845 as its median house value increased by 1400 percent from $15,900 to $239,650. During the same period, the influx of migrants to Dallas County more than doubled from 164,134 to 312,593 while its median house values increased only by half as much from $11,200 to $92,700. California is one of the most highly land use regulated states, while Texas has the lowest regulations.

This paper builds on the "border puzzle" in the trade literature and the internal migration literature. Researchers find that there are significant barriers to trade at intranational borders (Agnosteva et al. 2014). This paper is the first to provide a measure of the state border effect for the U.S. internal migration, and to evaluate how socioeconomic differences affect the border effect. There is also a growing body of research on documenting and explaining the decline of internal migration in the United States (Molloy et al. 2011), through labor market changes (Kaplan and Schulhofer-Wohl 2017; Molloy et al. 2017), demographic shifts (Rhee and Karahan 2015), or state regulation changes (Johnson and Kleiner 2015; Ganong and Shoag 2017). I contribute to this literature by constructing a novel data set of bilateral migration flows. I show that the border effect has increased over time, and this increase is correlated with large differences in housing prices, as state governments increasingly implement land use regulations.

1.2 Related Literature

This paper is broadly related to two strands of literature. First, it is related to the trade literature on the "border puzzle" and especially on domestic borders. Since the seminal finding of McCallum (1995)'s border effect between Canada and the United States, researchers have continued to find significant border effects in trade flows despite increasingly integrated global economy. Studies show that the estimated effects of trade frictions at the border can be explained in part by the tariffs and trade barriers, currency, home bias in preferences, historical colony experiences, regional trade agreements, as well as technical issues in estimating gravity model specifications (Anderson and van Wincoop 2003; Bergstrand et al. 2015; Helpman et al. 2008; Silva and Tenreyro 2006).

However, significant home bias is also found for domestic trade flows across subnational borders, which are not subject to the aforementioned trade barriers and have more comparable distance measures.[3] Wolf (2000) finds that the U.S. trade flows within states are three times higher than across states, even in the absence of formal and informal trade barriers. Researchers have shown that the presence of domestic border effects can be explained in part through factors such as information networks or wholesaling activity (Combes et al. 2005; Hillberry and Hummels 2003; Millimet and Osang 2007). This study uses the same methods to measure the state border effect for internal migration flows.

Second, it is related to literature examining the aggregate trend and the determinants of internal migration in the United States.[4] The decline of the U.S. migration from 1980, using multiple sources of migration data, is well-documented in Molloy et al. (2011). My data extends the 5-year migration flows at sub-state levels of geography and confirms the decrease in migration. Migration is deter-

[3] See Agnosteva et al. (2014) for literature on intranational border barrier for trade.

[4] For the history of internal migration in the U.S. and the overview of literature, see Greenwood 1997; Molloy et al. 2011.

mined by a combination of multiple factors, including but not limited to, demographic characteristics, job opportunities, amenities, family reasons, government policies, and natural disasters. This paper is closely related to literature that studies the drivers of the recent decline in mobility, and researchers are able to explain the trend in part through channels such as changes in labor market (Kaplan and Schulhofer-Wohl 2017; Molloy, Smith, and Wozniak 2017) or demographic shifts (Rhee and Karahan 2015). Most of the research focus on the decline post-1990 as better data is available, and the decrease is more pronounced, but my data extends the period of observation, making it possible to document and observe earlier trends.

State regulations and interstate agreements are also possible explanations for migration slowdown, as state policies such as occupational licenses, land use regulations, or other interstate compacts can have direct implications for mobility.[5] This study explains the state border trend through the relationship between land use regulations, housing supply, and migration (Glaeser et al. 2005;Glaeser and Ward 2009; Quigley and Raphael 2005;Ganong and Shoag 2017). Ganong and Shoag (2017) build a panel measure of land use regulations at the state level between 1940-2010, and explain the lower income convergence after 1980 through skill sorting driven by lower net migration in high income places with strict land use regulations. Their idea is that for land use restricted states with high income, limited housing supply increases the house prices, making it less affordable for low-skilled workers in particular.

Research that lie at the intersection of the two strands of the literature are most closely related to this study. Kone et al. (2016) uses the same empirical strategy as this paper and measures the border effect at the state line for different

[5]Studies on the impact of occupational licensing on migration has been limited due to lack of a comprehensive data, and existing works show mixed results for select occupation groups (Johnson and Kleiner 2015). Feng (2014) finds that interstate banking deregulation frees capital flows, reducing labor mobility as wage differences decline. This leads to a decline in interstate migration during 1990-2005.

subgroups of populations in India. They also suggest that state level policies contribute to the state border and provide some preliminary evidence. While their data is more disaggregated and offer a good comparative measure for the measure of domestic administrative border effects, it is for a single cross-section. I observe the U.S. state border effect for multiple periods, and show that it has increased over time. Ganong and Shoag (2017) is also closely related to this paper. My results complement their findings and support the claim that land use regulations negatively affect net migration through higher house prices. I show that this effect is strong for in-migration, and it increases over time.

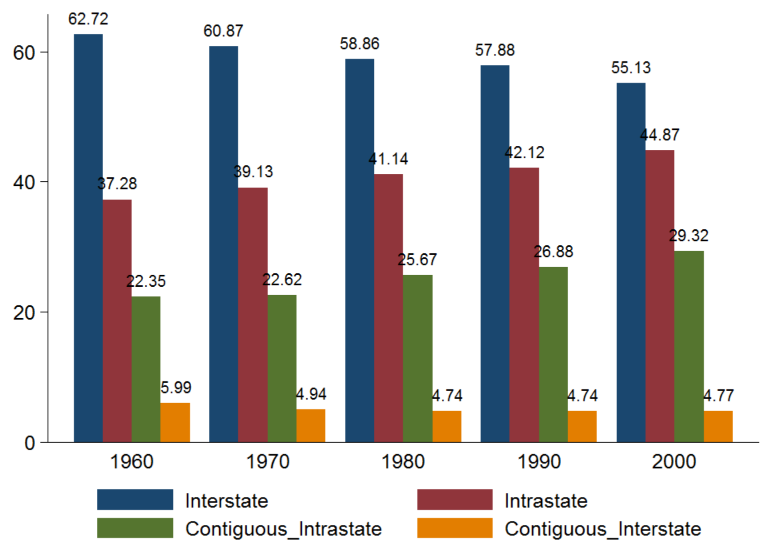
1.3 Data and Empirical Framework

1.3.1 Data

The data consists of bilateral migration flows, destination and origin characteristics, and the bilateral controls that include distance, contiguity, the dissimilarity measures of socioeconomic variables, and the land use regulation measure.

All of the variables are defined at the State Economic Area (SEA) level, except the land use regulation data which is only available at the state level. The SEAs were defined by the Census Bureau in 1950 as single counties or groups of contiguous counties within the same state that had similar characteristics. The 1960 set of SEAs that are used in this study were revised to reflect minor changes and also include Alaska and Hawaii. The size and the number of SEAs per state vary widely, between 1 and 31 per state. For empirical analysis, the time-invariant differences of SEAs will be controlled by the destination and origin fixed effects. There are 509 SEAs in total that cover the entire United States. They constitute 258,572 pairs of destination and origin SEAs in total for each census year, out of which 6,666 (2.58%) are the pairs of within state moves and 2,794 (1.08%) pairs are contiguous SEAs. Figure 2 shows the migrants for each category in shares of total inter-SEA migrants over time. The contiguous intrastate moves alone comprises 20 to 30 percent of the total migration, and on average, more than 80 percent of all intrastate moves. Thus, contiguity is a strong indicator of high migration, and the baseline estimates are also reported for the contiguous sample.

Figure 1.2: Share of Migrants by Category



The migration data is collected from the Decennial Published Census Volumes for every decade between 1960 and 2000. Starting from 1940, the Decennial Census includes the question on the respondent's migrant status and the previous residence five years ago. The decennial census publishes SEA-to-SEA bilateral 5-year migration flows for 1960 and 1970, and county-to-county bilateral 5-year migration flows from 1980 to 2000.[6] By combining the two datasets at the SEA level, a 509-by-508 matrix of migration flows for five periods is constructed. The 5-year migration flows between SEAs published by the Census has some limitations. Repeated migration, within-SEA moves, or any migration outside the five-year period will not be counted. Despite the shortcomings of the data, the Census bilateral migration flow data best serves the purpose of this study as it

[6]The data starts from 1940, but in the 1950 decennial census, only annual migration flows are available. Thus, data between 1960 and 2000 is used in this study.

covers the entire United States and is representative of the whole population above 5 years old for the longest period of time.[7]

For destination and origin characteristics, county level data on population, median family income, unemployment rate, education, number of manufacturing plants, urbanization, median rent and house values, percentage of blacks, and the vote shares for the Republican Party are collected and aggregated at the SEA level for use.[8] As most of the variables are collected in the census year, lagged variables are used for the possibility of reverse causality.

The data also include dyadic variables that are fixed over time, such as distances and contiguity, and time-varying bilateral variables that proxy for how alike the SEAs are. The bilateral distances of SEAs are calculated by averaging the distances between all the possible combinations of the county pairs in two SEAs. The contiguity variable also uses the contiguity of the consisting counties, and takes the value of one if a pair of SEAs have counties that are adjacent to each other.[9] As explained in previous section, the dissimilarity measures are Euclidean distances of destination and origin on socioeconomic variables. The cross-sectional summary statistics of all variables are included in Table 1.

The land use regulations data is from Ganong and Shoag (2017). The authors use the state level counts of state supreme and appellate court cases with string "land use" as a proxy for the strictness of the regulations. The measure is con-

---

[7]The three main sources for U.S. migration data are the Current Population Survey (CPS), Internal Revenue Service (IRS), and the decennial census (hereafter, the Census). Both the CPS and the IRS data provides annual migration data. The CPS data is at individual level and starts from 1947 but the sample is much smaller and the previous residence is only available at the state after 1985. The IRS data is more disaggregated county-to-county annual migration data, but it starts from 1978 and only includes tax-payers. The Census microdata also provides individual level migration but geographical information lower than SEA, Countygroup or PUMA is suppressed for different years, making it difficult to build data at a consistent level over time.

[8]ICPSR 2896. http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/2896

[9]County bilateral distance data is from NBER County Distance Database (http://www.nber.org/data/county-distance-database.html). County contiguity data is from the Census County Adjacency File
(https://www.census.gov/geo/reference/county-adjacency.html).

structed as a rank of per capita cases for each state every year between 1940-2011, and takes a value between [0,1].

1.3.2 Descriptive Facts

1.3.2.1 Push and Pull Factors of Migration

Before discussing the state border effects for migration, this section examines the economic factors at origin and destination that attract migrants during the period from 1960 to 2000. For the main analysis, all of the time-varying observables and unobservables at destination and origin will be absorbed by the fixed effects.

Following the traditional gravity equation, the OLS estimates are reported in Table 2. All specifications include the bilateral controls of distance and contiguity to account for the cost of migration between each pair. As destination, origin, and year fixed effects capture all of the time-invariant destination and origin factors that may affect migration, such as climate or areas size, and any decade-specific migration shocks, the identification comes from variation in the control variables over time. To prevent possible endogeneity and reverse causality issues, the control variables are lagged and the values of previous decades are used for each census year instead.

The first column shows the OLS outputs from specification without any push or pull factors, and then each factor is added one by one in the following columns. The estimates for distance and contiguity are very stable and strongly significant across regressions, and the migration flows are decreasing in distance and non-contiguity of the pairs. Consistent with the findings from previous works, the coefficients for population at both destination and origin are positive and significant. The SEAs with growing populations both send and receive more migrants. Similar to population, median family income, education, and median house value also have positive coefficients for both destination and origin. The three variables

Table 1.1: Summary Statistics

Census Year	1960		1970		1980		1990		2000	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Migrants	88.12	612.55	102	716.79	133	989.79	144	1,183.33	154	1,210.44
Population	297,092	588,421.10	352,260	706,423.60	399,007	800,024.50	444,792.30	806,004.40	488,246	880,601.20
Median Family Income	2,630	769.07	4,836	1,284.36	8,418	1,865.14	15,596	3,013.74	31,542	7,094.46
Median Rent	35	9.05	58	15.44	90	22.84	210	40.95	367	104.05
Median House Value	5,809	2,200.90	9,324	3,083.54	13,604	4,695.81	40,446	14,589.73	67,058	40,295.26
High School Graduates (%)	30.69	10.10	37.82	9.77	48.44	10.95	62.87	10.42	72.53	8.50
Unemployed (%)	4.27	1.96	5.32	1.88	4.59	1.68	6.93	2.50	6.62	2.18
Urbanization (%)	44.67	26.09	48.78	26.82	51.75	27.35	53.18	27.57	53.82	28.01
Black (%)	9.85	14.18	9.60	13.43	9.37	12.51	9.43	12.42	9.75	12.53
Republican Votes (%)	55.36	12.60	43.53	12.49	47.92	8.06	55.16	8.35	43.32	8.56
Land Use Regulations (raw)	0.72	1.07	1.67	2.43	3.49	3.96	7.35	8.62	10.60	14.32
Land Use Regulations (centile)	0.22	0.18	0.33	0.23	0.52	0.24	0.71	0.20	0.73	0.19

[a] All variables are at the SEA level, except migrants and the land use regulations, which are at the SEA-pair and the state level. Migrants are defined for the previous 5 years to the census year. The control variables and the land use regulations are lagged, and the periods nearest to the census year are used.

Table 1.2: Push and Pull Factors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Logdistance	-1.538***	-1.538***	-1.539***	-1.539***	-1.538***	-1.538***	-1.538***	-1.538***	-0.931***	-1.538***
	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0303)	(0.0235)
=1 if SEA Contiguous	1.683***	1.683***	1.683***	1.683***	1.683***	1.683***	1.683***	1.683***	2.971***	1.683***
	(0.0481)	(0.0481)	(0.0481)	(0.0481)	(0.0481)	(0.0481)	(0.0481)	(0.0481)	(0.0659)	(0.0481)
Destination: Log(Population)		0.289***							0.950***	0.410***
		(0.0660)							(0.0387)	(0.0734)
Origin: Log(Population)		1.043***							0.981***	1.023***
		(0.0425)							(0.0280)	(0.0455)
Destinatin: Log(Family Income)			0.766***						-1.132***	0.708***
			(0.106)						(0.211)	(0.112)
Origin: Log(Family Income)			0.463***						-0.910***	0.262***
			(0.103)						(0.148)	(0.0697)
Destination: %High school grad				3.004***					4.651***	2.410***
				(0.427)					(0.375)	(0.412)
Origin: %High school grad				1.744***					4.041***	1.047***
				(0.414)					(0.280)	(0.260)
Destination: Log(House Value)					0.268***				0.388***	-0.154**
					(0.0695)				(0.126)	(0.0700)
Origin: Log(House Value)					0.574***				0.344***	0.000206
					(0.0507)				(0.101)	(0.0421)
Destination: % Unemployed						1.412**			3.359***	1.834***
						(0.589)			(1.223)	(0.515)
Origin: % Unemployed						-5.697***			2.339**	-2.739***
						(0.556)			(0.962)	(0.400)
Destination: %Urban							0.193		0.0736	-0.794***
							(0.234)		(0.137)	(0.197)
Origin: %Urban							1.654***		0.0600	-0.429***
							(0.237)		(0.104)	(0.140)
Destination: Manufact. plants								-0.0704*	-0.369***	0.0687
								(0.0408)	(0.0664)	(0.0632)
Origin: Manufact. plants								-0.318**	-0.384***	-0.0268
								(0.123)	(0.0505)	(0.0288)
Observations	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.627	0.635	0.629	0.629	0.629	0.628	0.628	0.628	0.476	0.637
Year	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Destination-FE	Y	Y	Y	Y	Y	Y	Y	Y		Y
Origin-FE	Y	Y	Y	Y	Y	Y	Y	Y		Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.

are highly correlated (0.83-0.96) and this affects the outcomes as areas with high house values also have high family income and larger shares of educated populations. The growing income at destinations attract migrants, whereas the rising house values at origin causes out-migration. Areas with increasing shares of educated populations also have positive coefficients, as educated populations are more likely to migrate. When all of the controls are included in column 10, the coefficient for house value at destination turns negative as the positive effect of income is controlled for.

The coefficients of unemployment rates have the opposite of the expected signs. An SEA with increasing unemployment will send less migrants, but receive more. One possible explanation is that the unemployment rate is inversely correlated with the share of rural populations, and areas that are increasingly rural are unattractive destinations but at the same time, there are less migration activities. (It may also be the lagged problem, as contemporaneous values have the expected signal.) The coefficient of urban populations at destination is insignificant, while origin with growing urban populations sends more migrants. This is due to population effect, however, as both coefficients turn negative when population is included. Urban areas are more populated, and after controlling for the positive effect of population, increasingly urban areas have less migration activities.

1.3.3 Empirical Framework

My empirical strategy is based on Bertoli and Moraga (2015)'s gravity model for migration, which is theoretically micro-founded by the Random utility maximization (RUM) model and yields the migration flow in the gravity-like form (Grogger and Hanson 2011; Beine and Oezden 2011; Beine and Parsons 2015).[10]

[10]The variable m_{ijt}, the bilateral migration flow from origin i to destination j at time t, is a function of the sending ability of origin (s_{it}), the attractiveness of destination (y_{jt}), the accessibility of destination from origin (ϕ_{ijt}), the multilateral resistance to migration (Ω_{it}), and the stochastic term (η_{ijt}).

Traditionally, this gravity equation was transformed into a log-linear form and estimated using OLS. However, researchers find that the OLS estimates of log-linear regressions are inconsistent, and suffers from selection bias as zero migration flows are dropped from sample.[11] Silva and Tenreyro (2006) show that the Poisson pseudo-maximum likelihood (PPML) estimator performs well in the presence of heteroskedastic error term and accommodates zeros in the dependent variable. As migration data is highly correlated and likely to be heteroskedastic, and 35% of pairs have zero migration flows in the data, PPML should be used for consistent estimate.

Thus, follows the baseline specification,

$$m_{ijt} = \exp[\beta_0 + \beta_1 \ln(d_{ij}) + \beta_2 contig_{ij} + \beta_3 border_{ij} + o_{it} + d_{jt}] + \epsilon_{ijt} \tag{1.1}$$

where $m_{ijt}$is the 5-year bilateral migration from origin SEA i to destination SEA j at census year t; $d_{ij}$is the bilateral distance between the SEAs i and j; $contig_{ij}$is a dummy variable that equals one if the SEAs i and j are contiguous to each other.[12] The variable $border_{ij}$ is a dummy that equals one if origin i and destination j are in different states. Following Wolf (2000), it indicates whether the migration is an intrastate or an interstate movement. The specification also includes time-varying origin and destination dummies to control for SEA-specific unobservables, such as

---

$$m_{ijt} = \phi_{ijt} \frac{y_{jt}}{\Omega_{it}} s_{it} \eta_{ijt}$$

$$\Omega_{it} = \sum_{k \in D} \phi_{ikt} y_{kt}$$

$$\ln(m_{ijt}) = \beta_0 + \beta_1 \ln(\phi_{ijt}) + \beta_2 \ln y_{jt} - \beta_3 \ln \Omega_{ijt} + \beta_4 \ln s_{it} + \ln \eta_{ijt}$$

[11] Refer to Anderson and Wincoop (2003), Silva and Tenreyro (2006), and Helpman, Melitz, and Rubinstein (2008) for more details.

[12] Unlike previous studies on bilateral intranational trade or interstate migration where intrastate moves are always coded as contiguous or as non-contiguous, the variable $contig_{ij}$ measures the contiguity of the SEA-pairs separately due to availability of data at disaggregated level.

population, unemployment rate, income, area size, climate, and so on.

The coefficient of interest is β_3 that measures the effect of crossing the border at the state line. Similar to trade, the significance and the magnitude of the border effect indicate the home bias for domestic migrants. Because the state border can be a discontinuous function of distance, I include as many distance controls as possible, such as distance-squared and distance-cubed, and a dummy for state contiguity in addition to the baseline regression. For all specifications to follow, the additional distance and contiguity measures are included on top of log-linear distance and SEA contiguity, and will be jointly denoted as X_{ij}. The size of the border effect is the antilog of coefficient β_3.

To understand what drives the border effect, the heterogeneity of the border effect can be explored over different socioeconomic characteristics of destination and origin. I provide measure of the state borders for different subsamples of the data. Also, the baseline specification is extended to include the interaction of the border dummy with bilateral dissimilarity measures of destination and origin SEAs as follows: $dissimilarity_{ijt}$ is a vector of how similar the origin and destination SEAs are on socioeconomic characteristics including race, urbanization, party vote shares, unemployment, and median rent. That is, all $dissimilarity_{ijt}$ variables are defined as $\mid (Variable_{jt} - Variable_{it}) \mid$, the absolute difference in values between destination and origin. The interaction term accommodates for any differential effects of the state border on changes of $dissimilarity_{ijt}$ measures.

$$
\begin{aligned}
m_{ijt} \quad &= \quad \exp[\beta_0 + \beta_1 X_{ij} + \beta_2 border_{ij} + \beta_3 dissimilarity_{ijt} \\
&\quad + \beta_4 border_{ij} * dissimilarity_{ijt} + o_{it} + d_{jt}] + \epsilon_{ijt} \qquad (1.2)
\end{aligned}
$$

By exploiting the panel structure of the data, I provide estimates for the temporal pattern of the border. The cross-section of baseline regressions estimate the border for each decade. The data also allows me to use the most rigorous specification possible and destination-origin pair fixed effects are included to capture all time-invariant pair-specific effects. For this regression, only time-interacted border survives.

$$m_{ijt} = \exp[\beta_0 + \beta_{2t} border_{ij} * Year_t + o_{it} + d_{jt} + pair_{ij}] + \epsilon_{ijt} \tag{1.3}$$

All regressions are estimated using PPML and the standard errors are two-way clustered at the destination and origin SEA level.[13]

Lastly, the land use regulation at destination state is added to the specification (3) and interacted with time-varying borders.

$$m_{ijt} = \exp[\beta_0 + \beta_{2t} border_{ij} * Year_t + \beta_{3t} border_{ij} * Year_t * landuse_{jt} + o_{it} + d_{jt} + pair_{ij}] + \epsilon_{ijt} \tag{1.4}$$

The estimates indicate whether regulations at destination affects the border effect over time.

1.4 Empirical Results

1.4.1 Baseline Results

Table 3 presents the measure of the state border effect estimated from the equation (1). The regressions include time-varying destination and origin fixed effects, which will capture all of the unobserved push and pull factors shown in the previous section. The fixed effects are also necessary to control for the effect of

[13]Estimation results with clustered standard errors at SEA-pair level are also available upon request. The t-statistics are inflated by almost an order of magnitude. Table A2 presents OLS results with different level of clustering.

Table 1.3: Baseline Regressions

		All		Contiguous
	(1)	(2)	(3)	(4)
Logdistance	-0.855***	2.958***	1.588**	2.644**
	(0.0412)	(0.790)	(0.753)	(1.236)
Logdistance2		-0.791***	-0.508***	-0.809***
		(0.144)	(0.140)	(0.299)
Logdistance3		0.0512***	0.0346***	0.0505**
		(0.00864)	(0.00846)	(0.0238)
=1 if SEA Contiguous	0.797***	0.651***	0.746***	
	(0.0406)	(0.0727)	(0.0631)	
=1 if State Contiguous			0.490***	
			(0.0529)	
=1 if State Border	-1.342***	-1.235***	-1.171***	-0.989***
	(0.0436)	(0.0418)	(0.0419)	(0.0360)
Observations	1,292,860	1,292,860	1,292,860	13,970
R-squared	0.692	0.691	0.704	0.961
Destination*Year-FE	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y
Border Effect	3.827	3.439	3.224	2.689
Border(distance)	3,589	4,983	8,801	82

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.

alternative destinations, which Bertoli and Fernandez-Huertas Moraga (2013) refer to as the multilateral resistance to migration. In column 1, the traditional gravity variables, such as the log of distance and the contiguity of the SEAs, are included with the state border dummy. In column 2 and 3, the distance polynomial terms are added to control for the potentially non-linear relationship between migration and distance (Davies et al. 2001). Column 3 also includes a state contiguity dummy in addition to the SEA contiguity. The last column shows results for contiguous SEAs only.

The first column shows that the elasticity of distance to migration is -0.855. One percent increase in distance decreases -0.86 percent of the migration flow. Also consistent with the previous studies that include the quadratic distance term (Davies et al. 2001;Arzaghi and Rupasingha 2013), column 2 and 3 show that distance is indeed non-linear, indicating a declining negative elasticity of distance on migration. This implies that migration is decreasing in distance at a diminishing rate, and once a fixed cost of a long distance move has incurred, be it economical or psychological, the distance elasticity is reduced. Intuitively, an additional increase of one mile in distance between places that are 10 miles apart and those that are 1,000 miles apart will not have the same effect.

The contiguity variable is defined at the SEA level and the coefficient of SEAs being adjacent to each other is 0.797 in column 1. This indicates that given all else equal, on average the migration flows between contiguous SEAs are $exp(0.797) = 2.22$ times higher than between non-contiguous SEAs. In column 3, another variable is introduced to control for the effect of states sharing the same border. While all contiguous SEAs are in adjacent states, there are also some non-contiguous SEA pairs between contiguous states. Even after controlling for the non-linearity of distances, the result in column 3 implies that the migration flows across non-contiguous states are on average $exp(0.490) = 1.63$ times lower.

The size and the significance of the state border effect, the variable of interest for this study, is highly significant and robust across all specifications. The size of the border effect ranges between $exp(1.342) = 3.83$ and $exp(1.171) = 3.22$ depending on specifications. Column 1 shows that there are on average 3.83 times less migration flows across states than within. After including distance polynomial and contiguity terms in column 3, the border effect is reduced to 3.22, but still remains strongly significant.[14]

[14]Comparison of border effects for intranational migration: Kone et al. (2016) uses 2001 Census of India and finds that between neighboring districts, migration across states is 1.56 times lower. Between non-neighboring districts, the border effect is close to 2. Compared

The deterring effect of border on migration can be expressed in distance (miles) by $\overline{D} \times [\exp(\beta_{StateBorder}/\beta_{Distance}) - 1]$, where $\overline{D}$ is the sample mean distance (Parsley and Wei 2001). That is, the border "width" is the distance from mean which produces the equivalent negative effect of crossing the border. As reported in the bottom of column 1, the border width is 3,589 miles, which means crossing the state border has the same negative effect as being 3,589 miles apart.[15] As this measure is sensitive to the coefficient estimates, it changes across specifications. In column 2 and 3, the border width increases up to 4,983 and 8,801 miles.[16] One of the reason why the number is so large is because this is the effect of border at the sample mean distance, 943 miles, which is not a short-distance migration. As the relationship between distance and migration is nonlinear, the negative effect of distance diminishes for long-distance moves and hence, the border is "wider." The distance effect is even smaller as additional control for contiguity is included in column 3, and this further increases the border width despite the smaller border size.

In column 4, the sample is restricted to the contiguous SEAs only. Although distance and contiguity are controlled for, a large share of total migration takes place between contiguous areas, and given the proximity, the contiguous pairs are likely to share common natural amenities or labor markets, and may be more comparable. The migrants are also likely to be better informed. Thus, between more comparable pairs of contiguous SEAs, the border effect drops by $exp(0.989-$

to non-neighboring districts across states, migration flows are each 5.6 and 8.8 times larger if moving between different but neighboring districts and between neighboring districts in same state. Poncet (2006) uses inter-provincial migration data in China between 1985-1990 and 1990-1995, and the size of the estimated province border effect is between 21 and 25. Comparison of border effects for trade in U.S.: Wolf (2000)'s estimated state border effect is 4.39 using U.S. intranational trade data for 1993. The size ranges from 4.39 to 3.15 depending on the specifications and all are estimated without the fixed effects. Millimet and Osang (2007)'s estimates also range between 4.9 and 7.14 in 1993, and between 5.91 and 8.45 in 1997.

[15] Comparison of the state border width for trade: Millimet and Osang (2007)'s estimate ranges between 6,450 to 7,174 miles in 1993, and over 10,000 miles in 1997.

[16] The border width is calculated by solving for d^* from $\beta_{StateBorder} = [\ln(\overline{D} + d^*) - \ln(\overline{D})](\beta_{Distance} + 2\ln\overline{D} \times \beta_{Distance^2} + 3\ln\bar{D}^2 \times \beta_{Distnace^3})$.

$1.171) - 1 = 16.6\%$. The size of the border effect is 2.689, and this is equivalent of being 82 miles apart at the mean distance for contiguous SEAs.[17] Even for the SEAs that are adjacent to each other, the state boundaries inhibit migration flows substantially.

In Appendix Table 1, the OLS estimates of the same specifications are also reported for each column. Although OLS estimates of log-linear equation is known to be biased, it allows more flexibility in adding fixed effects or using different levels of clustering. The PPML estimates are mostly smaller than the OLS estimates, which is the typical result of PPML (Silva and Tenreyro 2006).[18] The border effect is highly significant and robust, ranging from 4.6 to 3.76. The border width from the OLS estimates, however, is much lower than PPML estimates at 1,874 miles, and this is due to a higher OLS distance elasticity.

1.4.2 Understanding the Border

To understand why the state border effect is so significant, I now calculate the border effects for different sub-groups of the data. The idea is to observe border heterogeneity, as the barrier that migrants face at the state border may not be the same between places that have higher income, for instance. I divide the sample by the distributions of income, education, and urbanization at destination and origin SEAs.

The first column in Table 4 is the benchmark regression from Table 3 column 3. In column 2, I limit the sample to SEAs whose per capita income are in the top 25 percentile. I find that the border effect falls by $exp(0.998 - 1.171) - 1 = -15.88\%$ compared to the benchmark estimate. The SEAs that have more

[17]The sample mean distance for the contiguous SEAs is 95 miles. For the pooled sample, the sample mean is 943 miles. If 943 is used instead, the border 'width' would be 1,039 miles.

[18]Silva and Tenreyro (2006) state "OLS greatly exaggerates the roles of colonial ties and geographical proximity. Using the Anderson–van Wincoop (2003) gravity equation, we find that OLS yields significantly larger effects for geographical distance. The estimated elasticity obtained from the log-linearized equation is almost twice as large as that predicted by PPML."

educated population also have lower state borders. High education SEAs are defined as those in the top 25 percentile when ranked by the share of population who attained high school education or more. Column 3 shows that between high education SEAs , the size of the border effect falls by $exp(0.919 - 1.171) - 1 = -22.28\%$. For highly urban SEAs, the border coefficient is the lowest at -0.692, dropping by close to 40 percent ($exp(0.692 - 1.171) - 1 = -38.05\%$). On the contrary, low income, low education or low urbanized SEAs (bottom 25 percentile) in column 5, 6, and 7 have similar or larger border sizes. Thus, there is more free mobility across states between SEAs with high income, high share of educated population, and especially between urban SEAs. This is also consistent with well-known findings that individuals who are educated and have high incomes are more likely to migrate.

While this exercise alone does not help explain the driver of the state border effect, it shows differential border effects. The border heterogeneity can further be examined by using the $dissimilarity_{ijt}$ vector following specification (2). The interacted border terms estimate to what extent the level of the state borders is affected by the social and economic differences between the areas and are reported in Table 5. All regressions include distance polynomial terms, contiguity measures, and time-varying destination, origin fixed effects as before, and while not included in the table, the elasticities of the traditional gravity variables are robust and do not differ largely from the estimates of the baseline regression. As explained earlier, the dissimilarity measures are defined as the absolute differences in socioeconomic factors such as population, income, rent, house prices, unemployment, urbanization, race, and party vote shares are reported in each column. The effect of the state borders now equals the coefficient of the border plus the interaction term, and a negative coefficient of the interaction indicates that the deterring effect of the border is increasing in the corresponding dissimilarity measure. Almost all of the coefficients of interacted terms are negative, implying

Table 1.4: Border for Different Groups

	(1) All	(2) High Income	(3) High Education	(4) High Urban	(5) Low Income	(6) Low Education	(7) Low Urban
Logdistance	1.588**	-4.389***	1.718	-3.468***	-3.245**	-4.341**	-0.0169
	(0.753)	(1.058)	(1.119)	(0.950)	(1.325)	(1.970)	(1.653)
Logdistance2	-0.508***	0.579***	-0.442**	0.406**	0.127	0.293	-0.389
	(0.140)	(0.192)	(0.198)	(0.180)	(0.245)	(0.353)	(0.288)
Logdistance3	0.0346***	-0.0281**	0.0267**	-0.0185*	0.00309	-0.00236	0.0288*
	(0.00846)	(0.0115)	(0.0115)	(0.0110)	(0.0149)	(0.0207)	(0.0164)
=1 if SEA Contiguous	0.746***	0.136	0.629***	-0.0685	0.550***	0.565***	0.554***
	(0.0631)	(0.125)	(0.127)	(0.144)	(0.0472)	(0.0522)	(0.0582)
=1 if State Contiguous	0.490***	0.285***	0.471***	0.252***	0.378***	0.434***	0.424***
	(0.0529)	(0.0683)	(0.0539)	(0.0625)	(0.0453)	(0.0480)	(0.0424)
=1 if State Border	-1.171***	-0.998***	-0.919***	-0.692***	-1.145***	-1.180***	-1.303***
	(0.0419)	(0.0936)	(0.0798)	(0.0865)	(0.0404)	(0.0420)	(0.0412)
Observations	1,292,860	80,010	80,010	79,758	80,010	80,518	80,518
R-squared	0.704	0.807	0.811	0.832	0.887	0.926	0.933
Destination*Year-FE	Y	Y	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y	Y	Y
Border Effect	3.224	2.714	2.506	1.998	3.142	3.253	3.680

[a] Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Two-way clustering is used at destination-origin SEA level.

Table 1.5: Heterogeneity of the Border Effect

	(1)	(2)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if State Border	-3.676***	-0.698***	-1.065***	-1.377***	-1.047***	-0.780***	-1.022***	-1.129***
	(0.305)	(0.121)	(0.0634)	(0.108)	(0.0479)	(0.0685)	(0.0543)	(0.0550)
|Population(Dest. - Origin)|	-0.187***							
	(0.0259)							
Border X |Population(Dest. - Origin)|	0.195***							
	(0.0244)							
|Family Income Med.(Dest. - Origin)|		0.0174						
		(0.0145)						
Border X |Family Income Med.(Dest. - Origin)|		-0.0651***						
		(0.0155)						
|Rent(Dest. - Origin)|			-0.0256					
			(0.0165)					
Border X |Rent(Dest. - Origin)|			-0.0393**					
			(0.0156)					
|House Value(Dest. - Origin)|				-0.0421**				
				(0.0170)				
Border X |House Value(Dest. - Origin)|				0.0237*				
				(0.0121)				
|%Unemployed(Dest. - Origin)|					2.269			
					(1.461)			
Border X |%Unemployed(Dest. - Origin)|					-7.831***			
					(1.667)			
|%Urban(Dest. - Origin)|						0.627***		
						(0.124)		
Border X |%Urban(Dest. - Origin)|						-1.326***		
						(0.173)		
|%Black(Dest. - Origin)|							0.348	
							(0.335)	
Border X |%Black(Dest. - Origin)|							-2.167***	
							(0.402)	
|%Republican Vote(Dest. - Origin)|								-0.981***
								(0.335)
Border X |%Republican Vote(Dest. - Origin)|								-0.494
								(0.347)
Observations	1,288,808	1,288,738	1,286,462	1,288,008	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.715	0.709	0.707	0.700	0.708	0.732	0.713	0.708
Destination*Year-FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y	Y	Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.
[b] All specifications include distance polynomial, contiguity measures, and time-varying destination and origin fixed effects.

that crossing state borders is more difficult between areas that are dissimilar.

Compared to the benchmark result of in Table 4 column 3, the border coefficient ranges from -0.698 to -3.676, depending on how similar the destination and origin are to each other. Consistent with the results in Table 4, income and urbanization have the largest effects on border. If a pair of SEAs are perfectly similar in median family income and urbanization rate, the border effects will each drop by 37 percent and 32 percent given all else equal. This indicates that the economic disparities between urban and non-urban areas explain a significant part of the border. The result in column 7 shows that urbanization has a different effect within and across states. Positive coefficient for within state moves implies that there are more moves between rural to urban, or urban to rural areas while for across states, moves between urban areas or rural areas dominate.

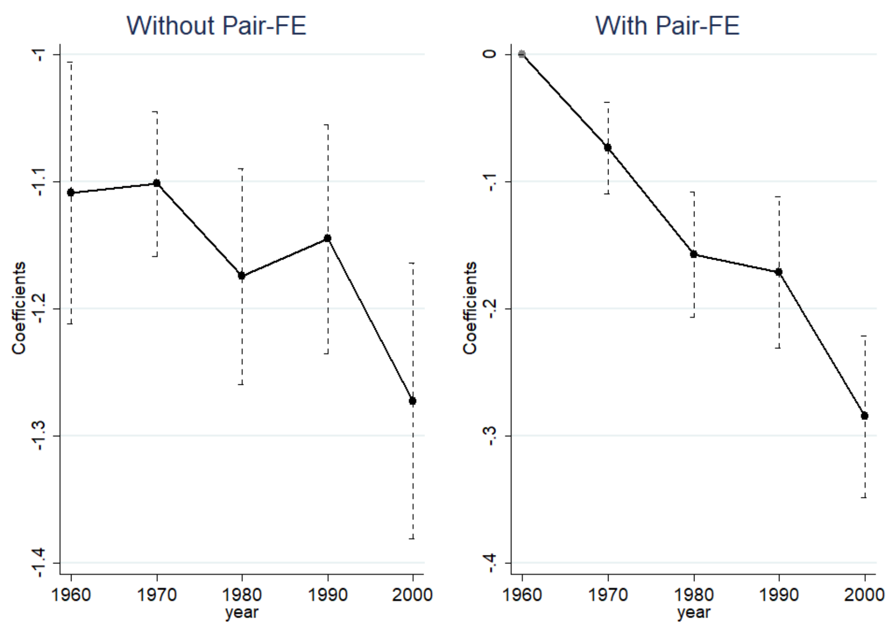
Population differences have an opposite effect on border, and the border is smaller for areas that are more different in population. This may be driven by the fact that as the variance for population is large, and small differences are mostly between less populated areas, this systematically lowers migration flows in between. Once pair-specific factors are controlled for, like other control variables, migration is decreasing in population differences. The coefficient for interacted border with house value differences is also positive but weakly significant. I will discuss this in the later section.

Non-economic controls, such as share of blacks and the votes for the Republican Party, are also included in the last two columns. Similar share of black population also lowers the border effect by 13.8 percent. The border is higher between areas that are more different in racial compositions, and at the maximum difference of 73.9 percent, the border coefficient will increase up to -2.62. Interaction with vote share differences is not significant, but interestingly, migration flows are smaller for SEAs with different party preferences within states, suggesting evidence of political sorting.

1.4.3 Border Trend

I have established a substantial level of the border effect at the state line. In this section, the panel structure of the data is utilized to examine the temporal pattern of the state border effect. Given lower transportation costs and improved accessibility over time, barrier of crossing the state is expected to decrease over time.[19]

Figure 1.3: Border Trend



Note: Border coefficients are from regression results in Table 8 Columns 1 and 2.

Figure 5 depicts the state border trend. The left panel plots the antilog coefficients of time-interacted border from column 2 with pair fixed effects. It is possible that this strong increase in border trend shown in the left panel is because the effects of other gravity variables, such as distance and contiguity variables, are fixed over time. The right panel shows that even after interacting the distance

[19]The fall in transportation and communication cost is well documented in Rhode and Strumpf (2003).

and contiguity variables with year, the growing border trend is significant and the size of the increase is even larger. Thus, I find that the size of the state border effect has increased over the period parallel with the growing difference between aggregate migration trend for within and across states, shown in Figure 1.

In the first two columns in Table 6, the interaction term between year and the border dummy are reported, and the cross-section PPML outputs for each census year are included in the following columns. In both columns 1 and 2, I find that contrary to expectations, the border has actually increased over time. The state lines act as higher barriers on migration over time, and there are less migrants crossing state lines. Column 1 shows that between 1960 and 2000, the border effect has increased by $exp(1.273 - 1.109) - 1 = 17.82\%$. In column 2, the pair fixed effects are introduced following specification (3). This is the most rigorous specification demanded of my data controlling for all unobservables specific to an origin-destination pair, and the only available variability for identifying coefficients is within-pair across time. Thus, all other variables are collinear with the fixed effects, and only the interaction terms between border and year dummies survive. The coefficients of time-interacted border are relative to the base level in 1960, which is omitted, and the increasing trend is more obvious. The magnitude of the border effect increases by $exp(0.285) - 1 = 32.97\%$ in 2000 compared to the border in 1960.[20]

For the cross-section regressions in the following columns, the pair fixed effects can no longer be included, as there are no time variation within each sample, and the destination and origin characteristics are static and absorbed by the destination and origin fixed effects. The results show that the size of the border effect ranges from 2.646 in 1960 to 3.583 in 2000, and has increased by $exp(1.276 - 0.973) - 1 = 35.39\%$, which is consistent with the increase found in

[20] The sample size is smaller because with pair fixed effect, the origin-destination pairs that have zero migrant flows for all five decades are dropped for PPML regression. This is in total 119,845 observations, 23,969 origin-destination pairs. (But this does not affect the results.)

Table 1.6: Border over Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	All	1955-60	1965-70	1975-80	1985-90	1995-00
Logdistance	1.573**		2.208***	2.833***	1.068	1.067	1.231*
	(0.753)		(0.819)	(0.849)	(0.877)	(0.762)	(0.717)
Logdistance2	-0.505***		-0.641***	-0.725***	-0.410**	-0.405***	-0.449***
	(0.140)		(0.153)	(0.156)	(0.164)	(0.142)	(0.133)
Logdistance3	0.0344***		0.0427***	0.0469***	0.0289***	0.0282***	0.0314***
	(0.00845)		(0.00937)	(0.00934)	(0.00997)	(0.00856)	(0.00800)
=1 if SEA Contiguous	0.744***		0.719***	0.666***	0.754***	0.762***	0.784***
	(0.0634)		(0.0803)	(0.0717)	(0.0691)	(0.0637)	(0.0599)
=1 if State Contiguous	0.491***		0.512***	0.538***	0.504***	0.465***	0.453***
	(0.0529)		(0.0601)	(0.0572)	(0.0580)	(0.0501)	(0.0544)
=1 if State Border			-0.973***	-1.135***	-1.214***	-1.174***	-1.276***
			(0.0461)	(0.0420)	(0.0466)	(0.0453)	(0.0424)
Border1960	-1.109***						
	(0.0524)						
Border1970	-1.102***	-0.0738***					
	(0.0448)	(0.0184)					
Border1980	-1.175***	-0.158***					
	(0.0458)	(0.0251)					
Border1990	-1.145***	-0.171***					
	(0.0458)	(0.0303)					
Border2000	-1.273***	-0.285***					
	(0.0469)	(0.0325)					
Observations	1,292,860	1,173,015	258,572	258,572	258,572	258,572	258,572
R-squared	0.706	0.981	0.578	0.670	0.661	0.745	0.762
Destination*Year-FE	Y	Y	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y	Y	Y
Pair-FE		SEA					
Border effect			2.646	3.110	3.365	3.234	3.583
Border(distance)			4,360	8,270	10,813	8,389	11,123

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.

column 2. The border width also increases greatly from 4,360 miles in 1960 to 11,123 miles in 2000. This big jump in width is due to increasingly discounted long-distance moves relative to rising border effect over time.

One concern is the zero flows. Out of all possible SEA pairs, close to half of the pairs have zero migration flows and are mostly interstate pairs excepting few. Large number of zero flows may have an upward bias on the border effect as zero migration will imply high border barrier, but there were more no-flow pairs in the earlier periods. To address this concern, I also separately estimate the pooled and cross-sectional baseline regression for SEA pairs with positive migration flow only. If it is the zero migration flows that drive the border barrier, limiting sample will significantly affect both the level and the trend of the border estimate. I find that the results are similar. I also find the trend is even stronger when limited to contiguous SEAs only.

1.4.4 Understanding Border Trend

In Table 7, as with the level of the border, I interact the border trend with the dissimilarity in control variables to see whether border across time is affected by differences between destination and origin characteristics. The results show that the divergence in population, income, and housing costs can explain the increase in border effects. The changes in border effect over time are canceled out when interacted with differences in population and median house values. This implies that the border barrier increases with population sizes and house values for a given pair of destinations and origins over time. For example, in the 1990 census, the maximum difference in median house value was $267,700 between Santa Clara county in California and a SEA in Kansas that includes Smith, Jewell, Norton, Phillips, Republic, Marshall, and Washington counties.

While the border effect was decreasing in house value differences in the previ-

Table 1.7: Border Trend Interactions

Control Variable	(1) Population	(2) Income	(3) Rent	(4) House Value	(5) Urban	(6) Black
Border1970	-0.134*	-0.127**	-0.0939***	-0.286***	-0.0914***	-0.0578***
	(0.0787)	(0.0498)	(0.0240)	(0.0571)	(0.0235)	(0.0217)
Border1980	0.141	-0.0367	-0.0975***	-0.0718	-0.105***	-0.148***
	(0.101)	(0.0421)	(0.0214)	(0.0437)	(0.0280)	(0.0286)
Border1990	0.377***	-0.0810*	-0.0692***	0.0783	-0.135***	-0.188***
	(0.132)	(0.0467)	(0.0253)	(0.0759)	(0.0354)	(0.0360)
Border2000	0.153	-0.130**	-0.127***	-0.0789	-0.250***	-0.309***
	(0.158)	(0.0638)	(0.0380)	(0.103)	(0.0387)	(0.0379)
|Control(Dest. - Origin)|	-0.00266	0.0201***	0.00770	0.0183***	-0.104	-0.960***
	(0.00501)	(0.00619)	(0.00545)	(0.00637)	(0.0647)	(0.170)
Border1970X|Control(Dest. - Origin)|	0.00451	0.00806	0.00911	0.0274***	0.0496	-0.542***
	(0.00666)	(0.00699)	(0.00774)	(0.00741)	(0.0546)	(0.161)
Border1980X|Control(Dest. - Origin)|	-0.0234***	-0.0167***	-0.0210***	-0.0105*	-0.175***	-0.569***
	(0.00881)	(0.00627)	(0.00609)	(0.00623)	(0.0580)	(0.161)
Border1990X|Control(Dest. - Origin)|	-0.0423***	-0.0115*	-0.0293***	-0.0267***	-0.125**	-0.212
	(0.0113)	(0.00673)	(0.00797)	(0.00976)	(0.0635)	(0.200)
Border2000X|Control(Dest. - Origin)|	-0.0334**	-0.0177**	-0.0346***	-0.0193*	-0.126*	-0.125
	(0.0131)	(0.00858)	(0.0106)	(0.0114)	(0.0737)	(0.222)
Observations	1,168,927	1,168,864	1,166,613	1,168,153	1,168,927	1,168,927
R-squared	0.981	0.981	0.981	0.981	0.981	0.981
Destination*Year-FE	Y	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.

ous table, this effect has reversed when controlled for pair-specific time-invariant factors. This may be driven by the fact that while large house price differences induce less migration for all, this negating effect is smaller across states because within state moves are more sensitive to house prices.Once the pair-specific unobservables are controlled for, within-pair increase in house values will further increase border. The border effect also increases based on income differences. Between destination and origin with similar median family incomes, the border will only increase by half as much. Thus, I find the increasing disparities in population, income and housing costs can explain the border trend.

1.4.5 Land Use Regulations

1.4.5.1 Land Use Regulations

Ganong and Shoag (2017) find that the effect of income on outcomes such as housing constructions, house prices, population growth differs as land use regulations increase. This section provides some descriptive facts in line with the literature that the land use regulations are associated with increasing house prices and discouraging migration.

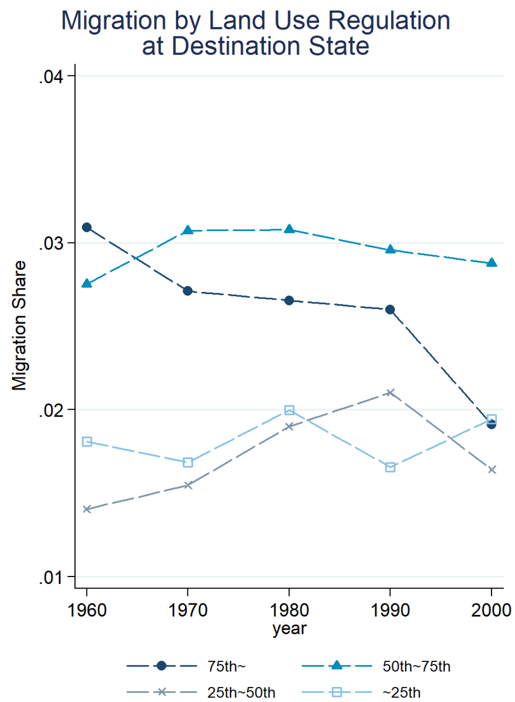
In order to see this relationship between house prices and regulations, Figure 3 plots the log of house value on the land use regulation measures for two groups, the high income and low income SEAs, defined by areas with income in the top and the bottom quartiles. Income and house values are observed at the SEA level, and the regulation data is defined at the state level. High income areas will always have larger house prices, but with the implementation of land use regulations, this relationship is further strengthened as higher income feeds into house prices for tightly regulated states due to limited housing supply. A positive slope indicates that the more regulated a state is, the larger the increase in house prices. The steeper slope in 2000 for high income group suggests this correlation

Figure 1.4: House Value and Land Use Regulation

Note: The SEAs with median family income above 75th percentile and below 25th percentile are defined as the high income and the low income groups. Author's calculation using data from Census and Ganong and Shoag (2017).

has grown over time as the number of land use regulations increase. For the low income group, the land use regulations are positively associated with house prices in 2000, but the relationship is much weaker as the housing demand will be lower. This implies that within state, the land use regulations will have different effects on areas as there are stronger effects for higher income areas consistent with Ganong and Shoag (2017).

Figure 1.5: Between State Migration and Land Use Regulation



Note: Author's calculation using data from the U.S. Decennial Census and Ganong and Shoag (2017).

Figure 4 displays the aggregate cross-state migration in population shares by the land use regulation measures at destination states. The states are grouped in quartiles and the sum of all four lines will be identical to the hump-shaped interstate migration line shown in the right panel of Figure 1. This demonstrates that there is a clear drop in cross-state migration flows to the most highly regulated states. The total decrease in the top quartile group accounts for 1.18 percentage point decline in the total migrating population. Between 1980-2000, there is a 0.7 percentage point decline in the population moving to highly regulated states, and this accounts for more than half of the drop (1.26 percentage point) in total interstate migration as shown earlier in Figure 1. While the migration from other states are on a declining trend by 1990 for most states, the bottom quartile group of states with low land use regulation displays no such decline.

1.4.5.2 Land Use Regulations and the Border Effect

The proliferation of land use restrictions constrains the housing supply, reducing housing affordability, and consequently, migration. Figure 3 and 4 have shown the effect of land use regulations on house prices and migration. Using the land use regulation data from Ganong and Shoag (2017), this section examines the effects of the regulations on the actual migration over the periods 1960-2000. Consistent with their findings, I expect that the more land use regulations are adopted by states, the more discouraged the incoming migration will be. This effect is also expected to be stronger for high income areas where the housing demand is not met due to the limited housing supply. In short, I test for the following claims: 1) highly regulated states have lower in-migration; 2) the effect of state border is increasing in land use regulations; and 3) this effect is stronger for high income areas.

The regression outputs of specification (4) are reported in Table 8.[21] As with other controls, I use the average of the regulation measures over the nearest 5 years that does not overlap with migration years for each census year.[22] The interaction terms between the border and the land use regulation measure at destination states indicate whether the state border is "wider" for the more regulated states. Negative coefficients of the interaction terms imply that regulations increase the border effect, and it is on average more difficult to move to a tightly regulated state.

Column 1 shows the results for total sample. The interaction coefficient is negative and significant in 1990, and the border increase is weaker post-1980 if there are no regulations at the destination state. In the following columns 2 and

[21] The regression outputs with the measure of zoning restrictions from Ganong and Shoag (2017) are also reported in the Appendix Table 2. The results are similar. I also repeat their placebo exercise using using total number of cases and find no effect in the Appendix Table 3. Border effect is not increasing in the general litigious environment.

[22] Ganong and Shoag (2017) use the average over the last ten years for decennial data. I find the effects are stronger if the average over the previous decade is used.

Table 1.8: Land Use Regulations

	(1)	(2)	(3)	(4)	(5)
Destination SEAs by Income	All	Above Median	Above 75th	Below Median	Below 25th
Border1970	-0.182***	-0.112**	-0.121***	-0.133**	-0.157**
	(0.0424)	(0.0531)	(0.0453)	(0.0541)	(0.0629)
Border1980	-0.178***	-0.0561	-0.0223	-0.203***	-0.279***
	(0.0552)	(0.0775)	(0.0826)	(0.0591)	(0.0594)
Border1990	-0.0343	0.0798	0.215	-0.0826	-0.0702
	(0.0734)	(0.109)	(0.149)	(0.0718)	(0.0791)
Border2000	-0.201**	-0.0297	-0.0123	-0.312***	-0.285***
	(0.0857)	(0.121)	(0.155)	(0.0941)	(0.110)
LanduseXBorder1970	0.264***	0.208**	0.268***	0.0485	0.225
	(0.0706)	(0.0855)	(0.0849)	(0.141)	(0.201)
LanduseXBorder1980	-0.00294	-0.156**	-0.213*	0.132*	0.396***
	(0.0549)	(0.0777)	(0.129)	(0.0736)	(0.111)
LanduseXBorder1990	-0.241***	-0.383***	-0.559***	-0.0394	0.00302
	(0.0775)	(0.116)	(0.188)	(0.0846)	(0.115)
LanduseXBorder2000	-0.153	-0.345***	-0.336*	0.0976	0.0700
	(0.0960)	(0.130)	(0.182)	(0.118)	(0.162)
Observations	1,166,923	582,570	291,524	555,136	270,369
R-squared	0.981	0.985	0.988	0.982	0.981
Destination*Year-FE	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level. Data is from Ganong and Shoag (2017).

3, the sample is limited to all migration flows to high income SEAs whose median family income is above the 50th and the 25th quantiles, and the negative effects of high regulations are even larger and increasing as expected. What is more, the growing effects of regulations completely absorb all of the increase in border trend. Land use regulations, however, will have no impact on migration inflows if the housing supply is not constrained due to low housing demand, and this is what the results show in columns 4 and 5. For low income SEAs, the land use regulations have no effect on border or if any, a positive effect, and fails to explain the border trend. The positive interaction effect may be due to an increased attractiveness of low income SEAs in more regulated states, as migrants substitute toward places with more affordable housing options.

As this paper uses bilateral migration data, I can decompose the negative effect of the land use regulations on net migration found in Ganong and Shoag (2017), and also investigate the effect of regulations at the origin. Consistent with their findings, the results in Table 8 have shown that the interstate migration inflow is reduced as regulation at destination state increases. At origin, the state residents may exit as the cost of living rises due to land use regulations. On the other hand, the land use regulations are often in favor of the incumbents, and induce residents to stay. In Appendix Table 4, I find that the land use regulations at origin have a similar but a weaker effect as the regulations at destination, and if significant, the regulation decreases the interstate outflow relative to the intrastate flows. In column 2 and 3, the regulations at high income origin areas will have reduced the interstate outflow of migration, whereas for low income origins, the interstate outflow weakly increases in regulation in Column 4. This may be driven by the demographic profiles of the existing residents, and for high income areas, the incumbents can still afford the higher living cost, but the middle or lower income earners in low income areas are affected. The land use regulations are largely insignificant for low income origins in Column 4 and 5, and the insignificant border effects for the areas with income in the bottom quantile in column 5 seem to be driven more by the sample itself than the land use regulations, as the coefficients are largely inisignificant.

There are two main concerns regarding the possible endogeniety of land use regulations: omitted variables and simultaneity. The way in which Ganong and Shoag (2017) address the endogneity issues for land use regulation on the income convergence is twofold. First, the authors run a placebo test using the total number of court cases, and show that the effects of regulations are not driven by some change in the general litigious climate. Second, they test for reverse causality by using the regulation measure in 1965, the period after which the land use regulations begin to gain popularity. The results show that the level of regulations in 1965 do not have differential effects on the income convergence rates in the pre-period, but in the post-period, there is a significant negative effect on income convergence. This shows that the increase in regulations cannot have been driven by the lower income convergence. Following their paper, I also provide the results for the placebo test in Appendix Table 2, and show that the total number of cases, which reflect the legal climate of the state, have insignificant or a positive effect on the border. This is the opposite of the effect of land use regulations, and if any, it will downward bias my results.

The pre-trend test, unfortunately, is not possible for this paper, as the migration data starts from 1960 and there are not enough data to test prior to the increased regulatons. But my migration data is at the SEA level while the regulation measure is at the state level. There is a large heterogeneity in migration within states, and I also use lagged land use regulation data, to reduce some of reverse causality issues. Also, the use of pair fixed effects will abosrb much of omitted variables as only within-pair over time variations are used.

Endogeneity concerns remain if there are any changes that are correlated with both regulation changes and migration changes. There is a possibility of some unobservables that are correlated with the increase of land use regulation, and at the same time, reduce migration. To address this, in Appendix Table 5, I include the border interaction terms with dissimilarity measures in addition to the specification in Column 3 in Table 8, to test whether the effect of land use regulations are absorbed by other socioeconomic differences. I find that the land use regulation at destination survives. For example, one possible concern is racially segregating and culturally conflicting places may have been more likely to implement regulations, and this may drive the results. Researchers argue that the change in climate toward land use regulations in 1960s can be attributed to racial desegregation in the aftermath of the Civil Rights Act (Fischel 2004). By including the difference in share of blacks as a control, I find that the effect of regulations on border effect survives.

1.5 Conclusion

The bilateral migration data from the decennial Census Published Volumes show that the decline in interstate migration led to an overall decrease in internal migration since the 1980s, but conversely, intrastate migration has increased since the 1960s. By using the gravity framework, I measure the border effect at the state line and quantify the home bias for migrants. Following Silva and Tenreyro (2006), I employ the PPML estimator with fixed effects to account for bias in the traditional OLS estimates. Despite lack of any formal border barriers, a significant and substantial border effect is found, and it is robust to different specifications. What is more, the estimates show that the border effect has increased over time and it has expanded with the differences in house prices. By using measures of land use regulations, I show that the more land use restricted states have higher borders for incoming migrants, and for high income areas, the increase of land use regulations can explain all of the growth from the border effect. My findings suggest that the popularity of land use regulations hinder migration, but further research is needed to fully understand the root of the border effect.

While this paper has addressed the decline in cross-state migration, the increase in within state migration has not been explained. The findings of this study suggest that non-economic factors such as party preferences affect intrastate migration, but further research is needed to identify the determinants of short-distance moves.

Appendix A

Table A.1: Baseline Regressions: OLS(ln(Migrants+1))

	All			Contiguous
	(1)	(2)	(3)	(4)
Logdistance	-1.396***	4.879***	2.650***	3.441***
	(0.0226)	(0.919)	(0.908)	(1.093)
Logdistance2		-1.179***	-0.734***	-1.099***
		(0.158)	(0.156)	(0.257)
Logdistance3		0.0711***	0.0443***	0.0772***
		(0.00893)	(0.00887)	(0.0201)
=1 if SEA Contiguous	1.012***	0.950***	1.009***	
	(0.0428)	(0.0510)	(0.0506)	
=1 if State Contiguous			0.396***	
			(0.0205)	
=1 if State Border	-1.527***	-1.396***	-1.324***	-1.077***
	(0.0467)	(0.0486)	(0.0499)	(0.0322)
Observations	1,292,860	1,292,860	1,292,860	13,930
R-squared	0.666	0.668	0.670	0.885
Destination*Year-FE	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y
Border Effect	4.604	4.041	3.759	2.935
Border(distance)	1874	1899	1980	79
Standardized Beta(%)	-71.27	-65.17	-61.80	
Standardized Beta(%)contig	47.25	44.33	47.09	
Standardized Beta(%)statecontig			18.50	

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level.

Table A.2: Zoning Regulations

Destination SEAs by Income	(1) All	(2) Above Median	(3) Above 75th	(4) Below Median	(5) Below 25th
Border1970	-0.138***	-0.0872**	-0.0841	-0.0930**	-0.119**
	(0.0350)	(0.0401)	(0.0544)	(0.0401)	(0.0475)
Border1980	-0.186***	-0.0747	0.00533	-0.199***	-0.250***
	(0.0533)	(0.0602)	(0.0701)	(0.0474)	(0.0559)
Border1990	-0.131*	-0.0533	0.0455	-0.0933*	-0.0771
	(0.0745)	(0.0874)	(0.122)	(0.0551)	(0.0702)
Border2000	-0.303***	-0.172**	-0.134	-0.340***	-0.313***
	(0.0701)	(0.0829)	(0.111)	(0.0664)	(0.0876)
ZoningXBorder1970	0.149**	0.158**	0.140	-0.0639	0.0581
	(0.0618)	(0.0744)	(0.0886)	(0.0892)	(0.121)
ZoningXBorder1980	0.0523	-0.0745	-0.256**	0.123*	0.295***
	(0.0730)	(0.0904)	(0.105)	(0.0731)	(0.0927)
ZoningXBorder1990	-0.0706	-0.171	-0.347**	-0.0137	0.0160
	(0.0909)	(0.108)	(0.151)	(0.0833)	(0.108)
ZoningXBorder2000	0.0291	-0.139	-0.203	0.177*	0.142
	(0.0841)	(0.0988)	(0.133)	(0.0994)	(0.142)
Observations	1,166,923	582,570	291,524	555,136	270,369
R-squared	0.981	0.985	0.988	0.982	0.981
Destination*Year-FE	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level. The measure of zoning cases is from Ganong and Shoag (2017).

Table A.3: Total Cases

Destination SEAs by Income	(1) All	(2) Above Median	(3) Above 75th	(4) Below Median	(5) Below 25th
Border1970	-0.00977	0.0421	0.0603	-0.0596**	-0.102**
	(0.0277)	(0.0283)	(0.0385)	(0.0290)	(0.0444)
Border1980	-0.200***	-0.145***	-0.0954**	-0.165***	-0.194***
	(0.0390)	(0.0380)	(0.0408)	(0.0390)	(0.0475)
Border1990	-0.256***	-0.205***	-0.158**	-0.139**	-0.0996
	(0.0593)	(0.0560)	(0.0680)	(0.0553)	(0.0699)
Border2000	-0.328***	-0.207**	-0.199**	-0.268***	-0.199**
	(0.0783)	(0.0833)	(0.0965)	(0.0836)	(0.101)
TotalXBorder1970	-0.220***	-0.200***	-0.307***	-0.168**	0.00700
	(0.0604)	(0.0754)	(0.107)	(0.0659)	(0.0946)
TotalXBorder1980	0.0955*	0.0719	-0.120	0.0537	0.158**
	(0.0552)	(0.0686)	(0.0779)	(0.0623)	(0.0747)
TotalXBorder1990	0.140**	0.0800	-0.0263	0.0583	0.0423
	(0.0702)	(0.0702)	(0.0794)	(0.0712)	(0.0879)
TotalXBorder2000	0.0621	-0.0807	-0.102	0.0302	-0.0642
	(0.0919)	(0.0997)	(0.110)	(0.104)	(0.118)
Observations	1,166,923	582,570	291,524	555,136	270,369
R-squared	0.981	0.986	0.988	0.982	0.981
Destination*Year-FE	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level. The measure of the total number of cases is from Ganong and Shoag (2017).

Table A.4: Land Use Regulations at Origin State

Destination SEAs by Income	(1) All	(2) Above Median	(3) Above 75th	(4) Below Median	(5) Below 25th
Border1970	-0.182***	-0.305***	-0.228***	-0.0223	-0.00677
	(0.0425)	(0.0649)	(0.0520)	(0.0455)	(0.0575)
Border1980	-0.176***	-0.324***	-0.277***	-0.0556	-0.0412
	(0.0555)	(0.0833)	(0.0757)	(0.0536)	(0.0488)
Border1990	-0.0352	-0.132	0.0398	0.0122	0.0469
	(0.0735)	(0.122)	(0.168)	(0.0615)	(0.0709)
Border2000	-0.199**	-0.330**	-0.227	-0.158*	-0.108
	(0.0858)	(0.133)	(0.164)	(0.0829)	(0.0988)
Landuse at OriginXBorder1970	0.259***	0.234***	0.293***	-0.0452	0.0831
	(0.0698)	(0.0847)	(0.0857)	(0.119)	(0.190)
Landuse at OriginXBorder1980	-0.00757	-0.0417	-0.00617	0.120*	0.235**
	(0.0556)	(0.0742)	(0.114)	(0.0655)	(0.0923)
Landuse at OriginXBorder1990	-0.241***	-0.349***	-0.479**	0.0157	0.0990
	(0.0776)	(0.132)	(0.224)	(0.0689)	(0.117)
Landuse at OriginXBorder2000	-0.156	-0.209	-0.207	0.0661	0.0467
	(0.0960)	(0.143)	(0.198)	(0.102)	(0.164)
Observations	1,166,931	585,143	293,434	551,708	268,928
R-squared	0.981	0.985	0.989	0.978	0.977
Destination*Year-FE	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level. The measure of the land use regulations is from Ganong and Shoag (2017).

Table A.5: Land Use Regulations And Controls

	(1) Population	(2) Income	(3) Education	(4) Rent	(5) House Value	(6) Unemployment	(7) Urban	(8) Black	(9) Republican
Border1970	-0.0306	-0.125	-0.148***	-0.126**	-0.295***	-0.0662	-0.148***	-0.114**	-0.121**
	(0.144)	(0.0800)	(0.0485)	(0.0518)	(0.0828)	(0.0497)	(0.0567)	(0.0488)	(0.0563)
Border1980	0.379**	-0.103	-0.0825	-0.00449	-0.0367	-0.0322	-0.0324	-0.0107	0.00255
	(0.150)	(0.107)	(0.0933)	(0.0832)	(0.121)	(0.0806)	(0.0869)	(0.0831)	(0.0858)
Border1990	0.777***	0.185	0.110	0.267	0.287	0.156	0.189	0.145	0.214
	(0.222)	(0.212)	(0.177)	(0.164)	(0.194)	(0.138)	(0.161)	(0.160)	(0.156)
Border2000	0.714***	0.00128	-0.246	0.158	0.150	-0.0626	-0.0410	-0.107	-0.0198
	(0.255)	(0.209)	(0.184)	(0.162)	(0.197)	(0.159)	(0.171)	(0.159)	(0.161)
LanduseXBorder1970	0.289***	0.261***	0.271***	0.276***	0.276***	0.242***	0.270***	0.256***	0.263***
	(0.0875)	(0.0845)	(0.0827)	(0.0857)	(0.0847)	(0.0841)	(0.0840)	(0.0847)	(0.0882)
LanduseXBorder1980	-0.200	-0.217	-0.188	-0.210	-0.213*	-0.200	-0.218*	-0.204	-0.229*
	(0.129)	(0.132)	(0.129)	(0.129)	(0.128)	(0.126)	(0.128)	(0.128)	(0.130)
LanduseXBorder1990	-0.558***	-0.564***	-0.505**	-0.556***	-0.556***	-0.484***	-0.561***	-0.486***	-0.562***
	(0.190)	(0.191)	(0.202)	(0.192)	(0.185)	(0.170)	(0.188)	(0.188)	(0.187)
LanduseXBorder2000	-0.361**	-0.345*	-0.205	-0.337*	-0.313*	-0.303*	-0.340*	-0.256	-0.350*
	(0.182)	(0.182)	(0.192)	(0.182)	(0.184)	(0.179)	(0.186)	(0.178)	(0.182)
|Control(Dest. - Origin)| XBorder1970	-0.00743	0.000819	0.281	0.000251	0.0221*	-3.421***	0.0884	-0.329	0.0172
	(0.0114)	(0.0108)	(0.215)	(0.0150)	(0.0118)	(1.302)	(0.0768)	(0.402)	(0.322)
|Control(Dest. - Origin)| XBorder1980	-0.0297***	0.0117	0.495*	-0.00726	0.00170	0.744	0.0541	-0.635	-0.240
	(0.0100)	(0.00957)	(0.282)	(0.0124)	(0.0114)	(1.277)	(0.0842)	(0.400)	(0.254)
|Control(Dest. - Origin)| XBorder1990	-0.0409***	0.00420	0.873**	-0.0164	-0.00762	-0.0439	0.101	-0.0263	0.0346
	(0.0135)	(0.0122)	(0.403)	(0.0125)	(0.0129)	(0.745)	(0.104)	(0.445)	(0.256)
|Control(Dest. - Origin)|XBorder2000	-0.0512***	-0.000868	2.139***	-0.0380**	-0.0167	1.686	0.108	0.183	0.225
	(0.0171)	(0.0136)	(0.628)	(0.0165)	(0.0166)	(1.491)	(0.131)	(0.421)	(0.320)
Observations	291,524	291,495	291,524	290,605	291,249	291,524	291,524	291,524	291,524
R-squared	0.989	0.988	0.989	0.988	0.988	0.988	0.988	0.988	0.988
Destination*Year-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Origin*Year-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Pair-FE	SEA	SEA	SEA	SEA	SEA	SEA	SEA	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Two-way clustering is used at destination-origin SEA level. The measure of the land use regulations is from Ganong and Shoag (2017).

CHAPTER 2

Partisan Geographic Sorting in the United States 1960-2000

2.1 Introduction

Is America divided? Republican and Democrats today are so different that not only are their political views and policy outlooks diverging, but even their thoughts on non-political issues such as science differ. Gallup finds 69% of Republicans believe global warming is exaggerated, while only 4% of Democrats think so.[1] Hatemi, McDermott, Eaves, Kendler, and Neale (2013) finds two partisan groups have different brains. And they are increasingly hostile to the people in other party (Iyengar, Sood, and Lelkes 2012; Iyengar and Westwood 2014), considering them as 'closed-minded' and 'dishonest' (Doherty, Kiley, and Jameson (2016)). Gentzkow (2016) state, "Political polarization is real... what divides them politically is increasingly personal."

Have politics become so personal that it matters in choosing where to live? Do Americans prefer to live in communities of similar ideology and political views? One way to test whether political preferences affect residential choices is to look at the geographical patterns in migration. If people selectively migrate based on political similarities, areas with different political preferences will have less migration as the dissimilarities between destination and origin can act as a political

[1] http://news.gallup.com/poll/231530/global-warming-concern-steady-despite-partisan-shifts.aspx

"distance" that may deter migrants.

This paper studies geographic sorting in the United States for 1960-2000. I estimate the effects of different political party preferences between areas on the migration flows. I use the presidential election outcomes to measure the support for the Democratic Party, which varies across geography and over time. The political dissimilarity measure is combined with place-to-place migration data at the State Economic Area (SEA) level, to test whether political differences between areas decrease migration. I use the gravity framework for migration, which explains the migration flows using destination and origin characteristics. This makes it possible to observe the relative importance of political dissimilarity in comparison to other differential factors that affect migration flows, such as urbanization or race.

Using the time series of the data, I also test for the "big sort" hypothesis. In his bestseller, *The Big Sort*, Bill Bishop claims that Americans are increasingly self-segregating into clusters that align with their political preferences, as sociocultural values have become more divided and political post-Civil Rights Act. He warns that such political sorting is increasing the divisions between communities and will tear the country apart. However, researchers have found mixed evidences on Bishop's hypothesis.

My findings provide evidences of partisan sorting in the U.S. migration. I show that different party affiliations reduced migration flows, and migration between same party areas was 1.03 times higher than migration between different party areas. One percentage point increase in vote share differences decreased migration by -0.38%. Even with the most demanding specification, although it is weakly significant, migration between a pair of areas was 1.007 times higher compared to migration between a pair of areas of which party affiliations change and become different over time. One percent increase in vote share differences over time also decreased migration by -0.07%.

However, I do not find evidences of the big sort, and Americans do not seem to have increasingly sorted geographically for the period 1960-2000. While I do find an increase in partisan sorting for areas with large differences in vote shares in 2000, this increase is more driven by across-SEAs differences and not robust to different specifications.

This paper is related to the literature on partisan geographic sorting and political segregation. Bishop (2009)'s "big sort" hypothesis attracted great interest from the general public and academics alike, yet the literature is not conclusive on this matter. Some researchers find big sort largely skeptical (Abrams and Fiorina 2012; Fiorina and Abrams 2008; Gentzkow, Shapiro, and Taddy 2016; Glaeser and Ward 2006), while others find supporting evidences of increasing political segregation (Abramowitz and Saunders 2008; Gimpel and Hui 2015; Hui 2013; Lang and Pearson-Merkowitz 2015; McDonald 2011; Sussell 2013).

Fiorina and Abrams (2008) claim that the big sort has been overstated. Repeating Bishop's analysis with voter registration data, Abrams and Fiorina (2012) find that the share of Independents has increased while population living in landslide counties has declined, and claim that the U.S. counties have actually become more politically heterogeneous. Glaeser and Ward (2006) examine political segregation over a longer span of time, and also conclude that there is no increasing trend in polarization, arguing that America is no more polarized than it has been historically. Gentzkow, Shapiro, and Taddy (2016) uses a novel method to estimate residential segregation by political parties from party identification survey and the campaign contribution data, and do not find an increase in partisanship over time. While these studies criticize the claim of increasing political polarization and segregation over time, the authors do not deny the presence of geographic sorting.

In contrast, Abramowitz and Saunders (2008) argue that politically active and well-informed Americans have become more divided on ideological orientations, increasingly partisan, and the states have become more politically polarized, social and cultural characteristics have become more correlated with party identification, and this has brought about more participation in politics. Sussell (2013) builds political segregation indices for California using both presidential election outcomes and the party registration data defined at sub-county levels, and shows that the political segregation has increased during the period 1992-2010. Lang and Pearson-Merkowitz (2015) explore partisan sorting between 1970-2012 using county-level presidential election data, and find that geographical polarization has increased starting from 1996. The authors use IRS in-migration and out-migration data, and find that larger in-migration contributes to the increasing political polarization.

While confirming the trend of political polarization and segregation is meaningful and interesting, this paper does not comment on the outcome of partisan sorting. The migration data set used in this study, while unique, is not granular enough to study the growth of political polarization over time. Instead, I focus on identifying migration patterns and describing the extent of partisan sorting in the aggregate data from 1960 to 2000. Thus, it is more closely related to studies that empirically examine the evidences of geographic sorting in migration data.

McDonald (2011) tracks the migration choices of election survey respondents for 2000-2006, and observe both the individual party identification and the subsequent choice of destination at the congressional district level. He finds that migrants choose destinations that align with their preferences. Cho, Gimpel, and Hui (2013) use voter files of seven states for 2004, 2006, and 2008, and find that voters are indeed migrating into places that are aligned with their political preferences. Both papers can identify individual level partisan sorting at a disaggregated geographical level, but due to limited data availability, only selective samples for recent years can be examined. While this study uses more aggregated migration flows data, it is able to analyze partisan sorting consistently for a longer period of time. It is also meaningful to observe sorting patterns in earlier years, and test whether partisan sorting is on the rise.

There are largely two possible mechanisms through which partisan geographic sorting occurs. First is the preference for political homophily. Migrants care about the political preferences of the potential neighbors and want to move to places with like-minded people (Gimpel and Hui 2015; Hui 2013). Second, migrants move according to other socio-cultural factors, which have become more political and correlated with partisanship over time. In this paper, I compare the political dissimilarities with other place-based characteristics, and show the relative importance of partisanship in migration.

The rest of the paper is organized as follows. Section 2 describes the data and the empirical strategy. Section 3 discusses the effect of different party affiliations on migration, and Section 4 concludes.

2.2 Data and Empirical Strategy

2.2.1 Data

This study uses data on migration, presidential elections, and other socioeconomic control variables. The 5-year migration data is collected from the Decennial Published Census Volumes for every decade between 1960 and 2000, and is consistently defined at the State Economic Area (SEA) level. The United States is divided into 509 SEAs in total, as a SEA is comprised of one or a group of contiguous counties contained within state. The observation of the data is at the destination and origin SEA pair level for each census year, and thus, a 509-by-508 matrix of migration flows for five periods is constructed.

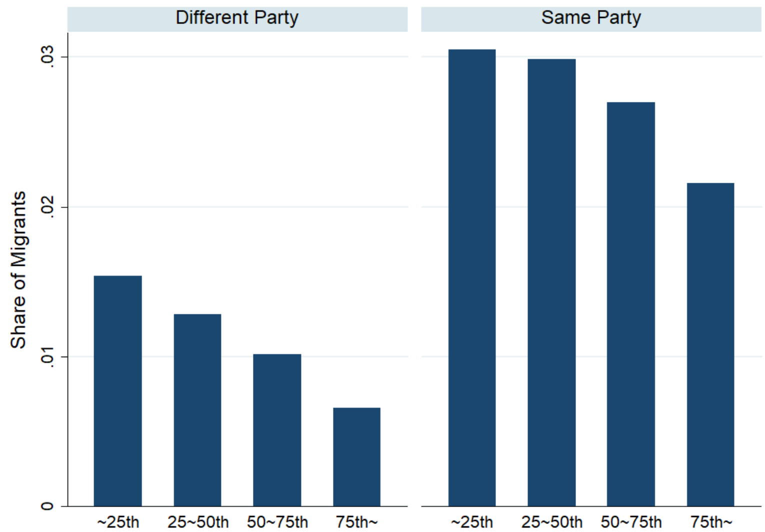
In order to measure the political party affiliation at destination and origin, the presidential election outcomes of 1956, 1964, 1976, 1984, and 1996 are used for each census year. I chose the election years that are nearest to the census years, as the aim of this study is to observe whether the party preferences exhibited in recent elections affect the migration flows.[2] The voting data are collected at the state and the county levels, and the SEA level data are constructed as the population weighted averages of the county vote shares. I use two different measures of dissimilarity in the party preferences of the SEAs. First, I define a dummy variable which equals 1 if the destination and origin have different winning political parties. For example, if the Republican party was the winning party in 1996 for a pair of SEAs, the dummy variable will be equal to 0 at the corresponding census year of 2000. The limitation of using a dummy variable is that this will treat all the pairs of different parties as the same, regardless of whether the vote share differences are at 1% or at 20%. Thus, in order to fully utilize the data, I also use the absolute differences in vote shares for the Democratic party. This will be a continuous variable that reflects how much the two areas differ in support for the Democratic party.[3]

As there are many factors that may also affect migration flows, I include traditional gravity controls such as bilateral distance, contiguity, and the state border. I also collect data on various socioeconomic factors, in order to account for other dissimilarities between places that can affect migration flows. Economic variables, such as population, family income, education, rent, unemployment, and urbanization, and non-economic variables, such as the number of churches, and the shares of black population are controlled for.

[2]The presidential election data for 1950-1990 are from ICPSR 13, and the data for 1996 is collected from CQ Press Voting and Elections Collection.

[3]To measure partisanship, political segregation and geographic sorting literature use voter registration data, presidential election outcomes, or survey data on party identification such as the ANES (American National Election Survey) or the CCES (Cooperative Congressional Election Study). The use of presidential elections as a measure of partisanship has been criticized by some researchers (Fiorian and Abrahms 2008), as it can be affected by other factors such as candidates. However, for the purpose of this study, presidential election outcomes are the only available data source that is representative of the whole country, and consistently defined at the county level from 1960.

Figure 2.1: Aggregate Migration Patterns and Partisanship



Note: Author's calculation using migration data from the U.S. decennial Censuses and the presidential election outcomes. Within groups of different party and same party SEAs, the areas are further categorized by the differences in vote shares. The migrants are in shares of total population.

2.2.2 Data Description

In Figure 1, I categorize the aggregate migration flows using the two measures of political dissimilarity, the $DifferentParty$ and the differences in vote shares. The height of each bar shows the shares of population who moved on average over the period 1960-2000 according to the sub-category. The migration flows are first divided into two groups, depending on whether the move is between SEAs that have the same or different party affiliations. Within the two groups, the migration flows are further classified according to the vote share differences. The vote share differences increase from left to right, meaning the last sub-category for both groups are the SEAs with large differences in vote shares (larger than 75th percentile). This uncovers two patterns in the aggregate data that are noteworthy. First, there are more migrants moving between same party areas than different party areas. All four bars on the left are lower than the bars on the right. Second, for both groups, the migration flows are consistently declining as the vote share differences increase. While the first point is likely driven by the fact that there are more number of same party pairs, second point suggests that migration flows are larger between areas that have similar political party preferences.

How different are the Republican versus Democratic areas? As social and demographic cleavages between the supporters of the two parties deepen, there is a concern that the significance of political dissimilarity measure is driven by other differences that are correlated with political preferences between the areas. Table 1 presents the summary statistics of the SEAs grouped by its winning party. There are no significant differences in the number of migrants that move to either party affiliated areas, but the two groups are different on some of the characteristics. The most significant differences between the two groups are in the shares of high school graduates and black population. On average, the Republican areas have higher income and rent, more educated and less black population, and less populated and urbanized, with slightly more numbers of churches. The comparison of the two groups show that other factors need to be sufficiently controlled for in order to correctly identify the partisan sorting effect.

2.2.3 Empirical Strategy

I use the modified gravity model for place-to-place migration, commonly used in migration literature (Greenwood and Hunt 2003), where migration flow between a destination and an origin is inverse of distance, and a function of size of the population, and the push and pull factors at destination and origin. For estimation, I use OLS because it allows for more flexible fixed effects estimations.

The observations are at the origin i and destination j SEA pairs over time t.

Table 2.1: Republican versus Democratic SEAs

	All	Republican	Democratic	Difference	P-Value
Migrants	125	127	121	6	0.341
Population	447,406	414,590	489,714	-75,124	0.025
Median Family Income	21,255	22,125	20,133	1,992	0.002
%High School	0.60	0.62	0.58	0.04	0.000
Median Rent	245	256	232	24	0.001
%Unemployment	0.06	0.06	0.06	0.00	0.079
%Urban	0.53	0.52	0.54	-0.02	0.050
Church	1.43	1.46	1.39	0.07	0.032
%Black	0.10	0.08	0.11	-0.03	0.000
Observations	2,541	1,431	1,110		

[a] There are 509 SEAs for 5 decades. AK, HI, and DC are not included in 1960, because they did not vote. Data are the mean values across SEAs during 1960-2000.

The baseline specification is as below.

$$\log(migrants_{i,j,t} + 1) = \alpha X_{i,j} + \beta DifferentParty_{i,j,t} + \delta_{i,t} + \theta_{j,t} + Pair_{i,j} + \epsilon_{i,j,t} \tag{2.1}$$

The dependent variable is the log of migrants between SEA pairs in census year, for which 1 is added to prevent loss of zero flows data. As defined above, $DifferentParty_{i,j,t}$ equals 1 if destination and origin SEAs have different party affiliations. Despite the advancement of technology and transportation, distance remains one of the key variables that affect migration, and the variable $X_{i,j}$ includes the traditional gravity variables, such as the bilateral distance polynomial terms and the contiguity of the SEAs. By including time-varying destination and origin fixed effects, $\delta_{i,t}$ and $\theta_{j,t}$, the area-specific characteristics that can affect migration such as population, job opportunities, climate, or amenities are controlled for. A group of state-pair dummies, $Pair_{i,j}$, are also added to account for pair-specific factors, such as historical relationship, past migration flows, or time zone differences between the destination and origin states that can affect migration flows. The most rigorous specification includes the SEA destination-origin pair fixed effects, 258,572 dummies for each pair, and only time variations within pairs will be used for identification.

The estimator of interest is β, which measures the effect of different party affiliations on the migration flows. A negative β would suggest the presence of geographic sorting, as there are less migration between places that have different political party preferences. Alternative to $DifferentParty_{i,j,t}$, the absolute differences of the vote shares, $\mid \%\triangle Democrat_{i,j,t} \mid$, or $Vote75_{i,j,t}$, an indicator variable for large vote share differences, are also used. The vote shares will allow for more variation in measuring political dissimilarities, and because vote shares do not have a directional implication, the absolute difference between pairs can be thought of as a political "distance" between the destination and origin. A negative coefficient for $\mid \%\triangle Democrat_{i,j,t} \mid$ will imply that larger differences in vote shares will reduce migration flows.

I also estimate how partisan sorting varies with destination characteristics by using interaction specification.

$$\begin{aligned} \log(migrants_{i,j,t} + 1) &= \alpha X_{i,j} + \beta DifferentParty_{i,j,t} + \gamma DifferentParty_{i,j,t} \times Y_{j,t} \\ &\quad + \delta_{i,t} + \theta_{j,t} + Pair_{i,j} + \epsilon_{i,j,t} \qquad (2.2) \end{aligned}$$

The vector of destination characteristics $Y_{j,t}$ includes population, median family income, the share of high school graduates, median rent, unemployment rate, urbanization rate, availability of churches, and the share of black population. If the interaction coefficient γ is negative, this indicates that the geographic sorting is increasing in the corresponding destination characteristics.

2.3 Results

2.3.1 Baseline

Table 2.2: Baseline Regression: Geographic Sorting

	(1)	(2)	(3)	(4)	(5)	(6)
=1 if Different Party	-0.0401***		-0.0342***		-0.00735*	
	(0.00673)		(0.00493)		(0.00391)	
|%Democratic(Dest. - Origin)|		-0.400***		-0.387***		-0.0706**
		(0.0524)		(0.0420)		(0.0296)
Log(Distance)	4.820***	4.790***	4.365***	4.360***		
	(0.928)	(0.926)	(0.811)	(0.810)		
Log(Distance)2	-1.169***	-1.163***	-1.081***	-1.080***		
	(0.160)	(0.159)	(0.151)	(0.150)		
Log(Distance)3	0.0706***	0.0702***	0.0674***	0.0674***		
	(0.00904)	(0.00902)	(0.00933)	(0.00932)		
=1 if Contiguous	0.946***	0.942***	0.939***	0.935***		
	(0.0509)	(0.0508)	(0.0345)	(0.0345)		
=1 if State Border	-1.395***	-1.393***				
	(0.0485)	(0.0485)				
Observations	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.668	0.668	0.689	0.689	0.787	0.787
Destination*Year FE	Y	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y	Y
State-pair FE			Y	Y		
SEA-pair FE					Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.

First, I test whether there are less migration flows between the SEAs with different political party preferences. Table 2 presents the baseline results on the political dissimilarities and migration flows. In the first column, the coefficient for $DifferentParty_{i,j,t}$ is negative and significant, indicating that if destination and origin SEAs support different political parties, the migration flows between areas are lower. The result implies that on average, moving between different party SEAs is exp(0.0401) = 1.04 times lower than moving between same party SEAs, after controlling for the nonlinear effects of distances and the contiguity of areas, as well as the push and pull factors at destination and origin captured by the time-varying fixed effects. The gravity variables, distance, contiguity, and the state border, are strongly significant as expected, and longer distances, non-contiguity, and cross-state borders reduce migration. The rest of the columns in Table 2 include the time-varying destination and origin fixed effects, and the gravity controls, which will be absorbed by additional pair fixed effects.

In order to utilize the richer information on the political characteristics of the area, I also use $\mid \%\triangle Democrat_{i,j,t} \mid$, the absolute differences in the vote shares for the Democratic party, as a measure of political dissimilarities. Column 2 presents the results for vote share differences using the same specification as Column 1, and shows that $\mid \%\triangle Democrat_{i,j,t} \mid$ also has a negative and significant effect on migration flows. The coefficient -0.4 can be interpreted in the similar way, as the elasticity of the political dissimilarity, implying that one percentage point increase in vote share differences decreases the migration flows by -0.4%.[4] The larger the differences in vote shares, the smaller the migration flows will be.

In the following Columns 3 and 4, I include state-pair fixed effects, because the migration flows can be affected by bilateral characteristics specific to the destination and origin states, e.g., sharing same regions, same time zone, or similar environments. Column 3 shows that the coefficient for $DifferentParty_{i,j,t}$ is smaller but remain significant at -0.0342, which indicates that the number of movers between same party SEAs is $\exp(0.0342) = 1.03$ times higher than the number of movers across different party SEAs. In Column 4, the coefficient for $\mid \%\triangle Democrat_{i,j,t} \mid$ is also slightly smaller but strongly significant.

In the last two columns, the results with most rigorous specification are reported. The SEA-pair fixed effects reduce much of the variation in the data and control for all destination-origin pair time-invariant factors. All of the grav-

[4]Vote share differences are rescaled to take value between [0,1]. $\exp((-0.4/100) - 1) \times 100 = -0.399$

ity controls become collinear, and only the political dissimilarity measures survive. The party affiliations can change over time for a destination-origin pair, and $DifferentParty_{i,j,t}$ measures how a switch of party affiliations over time may affect the migration flows. Column 5 shows that with SEA-pair fixed effects, $DifferentParty_{i,j,t}$ is weakly significant, and the effect is substantially reduced at $\exp(0.00735) = 1.007$. This may be because there are more cross-sectional variations and too little variations over time.

The last column shows that $\mid \%\triangle Democrat_{i,j,t} \mid$ also remains significant after controlling for SEA-pair fixed effects, although the magnitude of the coefficient is now greatly reduced to -0.0706. The interpretation of the coefficient is that one percentage point change in the vote share differences over time decreases the migration flows by -0.07%. Thus, if the differences in vote shares for a given pair increase over time, the migration flows between the two areas will decline. Evaluated at 17.97 percentage point, the mean value of the maximum differences in vote shares over time for SEA pairs, the effect on migration will be -1.25%. While the magnitude of the actual geographical sorting found in this study is small, this can have a greater impact on political polarization if residential sorting promotes higher political participation (Perez-Truglia 2018).

In the Appendix, Table A.1 presents the same specification as Table 1 for alternative years. As presidential election of 1964 was exceptional due to Southern realignment, I confirm that the results are not driven by the selection of presidential election years.

To summarize, Table 1 shows that there are less migrants moving between areas with different political party preferences, and if areas grow further apart in terms of vote shares for the Democratic party, the number of migrants also falls.

Table 2.3: Geographic Sorting and Other Controls

	(1)	(2)	(3)	(4)
=1 if Different Party	-0.0228***		-0.00790**	
	(0.00462)		(0.00384)	
|%Democratic(Dest. - Origin)|		-0.256***		-0.0766***
		(0.0351)		(0.0284)
|Population(Dest. - Origin)|	-0.258***	-0.258***	-0.194***	-0.194***
	(0.0150)	(0.0150)	(0.0172)	(0.0172)
|Family Income Med.(Dest. - Origin)|	0.186***	0.185***	0.269***	0.269***
	(0.0395)	(0.0395)	(0.0327)	(0.0327)
|%High School(Dest.-Origin)|	-0.183**	-0.177**	0.120	0.122
	(0.0855)	(0.0856)	(0.0808)	(0.0808)
|Rent(Dest. - Origin)|	-0.211***	-0.209***	-0.202***	-0.201***
	(0.0352)	(0.0352)	(0.0274)	(0.0274)
|%Unemployed(Dest.-Origin)|	-2.342***	-2.330***	-0.247	-0.250
	(0.235)	(0.235)	(0.159)	(0.159)
|%Urban(Dest.-Origin)|	-0.260***	-0.259***	-0.226***	-0.227***
	(0.0260)	(0.0260)	(0.0359)	(0.0359)
|Church(Dest. - Origin)|	-0.0448***	-0.0450***	-0.0671***	-0.0674***
	(0.00948)	(0.00949)	(0.0129)	(0.0130)
|%Black(Dest.-Origin)|	-0.842***	-0.834***	-0.894***	-0.890***
	(0.0785)	(0.0783)	(0.130)	(0.130)
Observations	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.696	0.696	0.788	0.788
Gravity Controls	Y	Y	Y	Y
Destination*Year FE	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y
State-pair FE	Y	Y		
SEA-pair FE			Y	Y
StdBeta(%) Party			-0.177	-0.367
StdBeta(%) Urban			-2.410	-2.419
StdBeta(%) Church			-2.171	-2.182
StdBeta(%) Black			-5.480	-5.453

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.

2.3.2 Geographic Sorting and Other Controls

One concern with the baseline findings is that other factors correlated with partisanship may drive the correlation between migration and the political dissimilarities. Table 1 has shown that there are some significant differences between Republican and Democratic areas. Since there are many factors that can simultaneously affect an individual's migration decision, other differences between destination and origin, that were not fully controlled for in the baseline regression, can affect migration flows. For example, typical Republican areas are known to be rural and religious, whereas Democrat areas are more urban and secular. Migrants may be more attracted to urban and educated areas, and this can lead to finding a partisan sorting effect.

In order to test for other non-political factors correlated with election outcomes and migration, I include various socioeconomic differences in addition to the political dissimilarity measure in Table 3. Each of the columns in Table 3 corresponds to the Columns 3-6 in Table 2 in the same order, with the inclusion of the absolute differences in population, family median income and rent, the number of churches per person, and the share of blacks, unemployed, urban, and educated population. Compared to the baseline results, the coefficients of political dissimilarity are largely unchanged and remain significant, especially with SEA-pair fixed effects. In Columns 1 and 2 with state-pair fixed effects, the size of the coefficients for $DifferentParty_{i,j,t}$ and $\mid \%\triangle Democrat_{i,j,t} \mid$ each drop from -0.0342 to -0.0228, and from -0.387 to -0.256 respectively. However, with SEA-pair fixed effects, the estimates are almost unchanged. Column 4 shows that $\mid \%\triangle Democrat_{i,j,t} \mid$ is significant after including other dissimilarity controls, and one percent point increase in vote share differences over time will decrease migration flows by -0.07%, same as the baseline estimate.

While Table 3 demonstrates the significance of political differences between areas on migration, it also shows how other factors play a role. The differences in the shares of educated population and the unemployment rates are significant in Column 1 and 2, however, both variables lose significance once SEA-pair fixed effects are included, implying that their temporal variations do not significantly affect migration flows over time. All other variables remain significant across different specifications. The results indicate that there are less migrants moving between areas that have increasing differences in population, house rent, urbanization, availability of churches, and the share of black population. This suggests that there are more people moving between places such as Chicago and Los Angeles, for example, than between Los Angeles and some rural town in Iowa. The median family income is the only variable that has a positive sign, which indicates that there are more people moving between areas with growing differences in income over time. All of the control variables are in absolute differences and therefore, cannot be interpreted in directional effects, but a possible explanation would be that migrants move for better economic returns.

It is also possible to compare the relative importance of the political dissimilarity measures with other characteristics. The beta coefficients of selected variables are reported in the last rows. A one standard deviation increase in the different-party is associated with beta/100 standard deviations in migration. The comparison shows that different party affiliation has about one-thirteenth of the effect of differences in churches, and the Democratic vote share difference has about one-seventh of the effect of churches and urban, but much smaller than black.

As with baseline results, I present the same specifications using alternative election years in Table A.2. I confirm that the results are not driven by the selection of presidential election years.

Table 2.4: Geographic Sorting at Destination

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
=1 if Different Party	-0.162***	-0.0588	-0.0177	-0.0328	0.0110	-0.0234***	0.0135*	-0.00890*
	(0.0551)	(0.0492)	(0.0154)	(0.0283)	(0.00954)	(0.00845)	(0.00792)	(0.00489)
Interact with Destination Characteristics								
X Population	0.0123***							
	(0.00441)							
X Median Family Income		0.00528						
		(0.00499)						
X %High School			0.0166					
			(0.0230)					
X Median Rent				0.00478				
				(0.00518)				
X %Unemployed					-0.331**			
					(0.153)			
X %Urban						0.0303**		
						(0.0138)		
X Church							-0.0144***	
							(0.00463)	
X %Black								0.0139
								(0.0350)
Observations	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787
Destination*Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y	Y	Y	Y
SEA-pair FE	Y	Y	Y	Y	Y	Y	Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.

2.3.2.1 Other Characteristics at Destination

Next, I examine how the partisan sorting effect changes with destination characteristics, by interacting $DifferentParty_{i,j,t}$ with destination controls. Different demographic groups are attracted to different pull factors at destinations. For example, old retirees may care more about climates and environment, while young migrants may be more attracted to urban cities with more job opportunities. As such, partisans may also be attracted to contrasting characteristics. Destinations with high income and education, for example, may attract more migrants who are politically conscious and strongly partisan. Thus, using interaction terms, I am able to observe whether migration to certain areas are more likely to be sorting on party preferences.

In Table 4, each columns show interacted $DifferentParty_{i,j,t}$ with various controls at destination. A negative interaction term implies that partisan sorting is increasing in the control variable, and a positive term, vice versa. Surprisingly, income, education, rent, and black population do not have a significant interaction effect, which indicates that the partisan sorting effect does not differ on the corresponding characteristics. However, the interaction terms with population, unemployment, urbanization, and church availability are significant. The results show that the negative effect of party differences on migration is stronger if destination is less populated, more unemployed, less urbanized, and has more number of churches. This suggests that migration for non-economic reasons are more likely to sort on political preferences, where the destination areas do not have attractive economic characteristics. The church variable shows an interesting result in Column 7. If destination has more number of churches per person, there is an increase in partisan sorting as migration from different party is reduced.

2.3.3 Geographic Sorting and Partisanship

Is different political party sorting effect greater for more partisan areas? If a destination area is strongly Democratic or Republican, migrants who support the opposite party will be even more reluctant to move to the area. Using the baseline result with vote share differences from Table 2, I show that the geographic sorting is stronger and more robust when the areas are more partisan. In other words, if the destination and origin have a large difference in support for the Democratic Party, the different party affiliations will have a greater effect.

Table 2.5: Geographic Sorting and Partisanship

	All				Outlier
	(1)	(2)	(3)	(4)	(5)
=1 if Different Party	-0.00735*			-0.00004	-0.0383***
	(0.00391)			(0.00389)	(0.0141)
(%Dem.$\geq p50$ at Destination)		-0.00872***			
		(0.00326)			
(%Dem.$\geq p75$ at Destination)			-0.0122***	-0.00154	
			(0.00437)	(0.00624)	
X (%Dem.$\geq p75$ at Destination)				-0.0156**	
				(0.00691)	
Observations	1,288,808	1,288,808	1,288,808	1,288,808	136,824
R-squared	0.787	0.787	0.787	0.787	0.850
Destination*Year FE	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y
SEA-pair FE	Y	Y	Y	Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The standard errors are two-way clustered at destination-origin SEA level.
[b] Last column includes all migration flows to outlier SEAs that voted differently from the state.

The first column in Table 5 is the baseline result, that different party affiliations have a weakly significant effect of -0.007 and reduce migration within a pair. In Columns 2 and 3, I define new dummy variables, $Vote50_{i,j,t}$ and $Vote75_{i,j,t}$ which equal 1 if the vote share differences are larger than, or equal to the 50th and the 75th percentiles. Both coefficients are significant and negative. The result in Column 3 indicates that if there is a big increase in the vote share differences for a pair over time, this will significantly reduce the migration flows by -1.2%. In Column 4, I interact $DifferentParty_{i,j,t}$ with $Vote75_{i,j,t}$ to test whether the geographic sorting is driven by large difference pairs in the tail. The result confirms this, and shows that the interaction term is significant and negative. Having a different party affiliation over time with a large vote share differences will reduce the migration flows by -1.5%.

In the last column, I define the SEAs that vote differently from its states as "outlier" SEAs. The idea is to capture areas such as Orange County in California, or Austin in Texas, where the political preferences of the area is well-known and strongly partisan. The geographical level of the data is not as fine, however, and aggregation at the SEA level averages out some of these political differences. In the data, more than 80% of the SEAs vote for the same party as the state, while 18% of the SEAs vote differently. Areas like Riverside County that voted Republican in Democratic-supporting California would be counted as an outlier in this data. I expect that migration flows to outlier SEAs are more likely to be partisan, and limit the sample to outlier SEA destinations only. The regression result shows that the partisan sorting effect is indeed stronger, and the coefficient has increased to -0.0383, indicating that the change of party preference over time will reduce migration flows. For migration flows to outlier SEAs, moving from same party areas are 1.03 times larger than moving from different party areas. Thus, Table 5 shows that the geographic sorting increases with the partisanship of the areas.

In Table 6, I divide the sample by red and blue SEAs, and examine if partisan sorting effect differs with the political party itself. The migration flows are categorized by red versus blue SEAs, first at origin, and then at destination.[5] Because limiting sample this way makes $DifferentParty_{i,j,t}$ collinear with the fixed ef-

[5] Comparing between red versus blue states also delivers the same results.

Table 2.6: Geographic Sorting by Types of Migration

	By Origin		By Destination	
	Rep	Dem	Rep	Dem
|%Democratic(Dest. - Origin)|	0.0510	-0.216***	0.0631	-0.132*
	(0.0604)	(0.0620)	(0.0558)	(0.0719)
Observations	691,572	497,848	691,572	497,848
R-squared	0.815	0.833	0.820	0.826
Destination*Year FE	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y
SEA-pair FE	Y	Y	Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.

fects, I use $\mid \%\triangle Democrat_{i,j,t} \mid$, the vote share differences, instead. In Column 1, for all migrant flows moving away from Republican SEAs, $\mid \%\triangle Democrat_{i,j,t} \mid$ is insignificant. For migrants moving from Democratic SEAs, $\mid \%\triangle Democrat_{i,j,t} \mid$ is highly significant at -0.216. When the sample is split by the destination winning party, $\mid \%\triangle Democrat_{i,j,t} \mid$ is again only significant for the Democratic SEAs, but the effect is weaker. This seems to suggest that geographic sorting is driven by migrants moving from Democratic areas. Because the data used in this study is at the aggregate level and does not include information on the migrant's own party identification, it cannot fully explain the reason behind the patterns found in aggregate flows, but the results show some suggestive evidences of heterogeneity in the geographic sorting of the Republican migrants and the Democratic migrants. Previous research on political sorting also have noticed some differences in migration choices depending on the party affiliation of the migrant.[6]

Table 2.7: Geographic Sorting over Time

	(1)	(2)	(3)	(4)	(5)	(6)
=1 if Different Party	-0.0388***	-0.0525***	0.00740			
Interaction with Year	(0.0127)	(0.0137)	(0.0112)			
x 1970	0.0124	0.0434*	-0.0327**			
	(0.0188)	(0.0232)	(0.0151)			
x 1980	0.00992	0.0305**	-0.0210*			
	(0.0140)	(0.0150)	(0.0125)			
x 1990	-0.0226	-0.00929	-0.0117			
	(0.0223)	(0.0236)	(0.0216)			
x 2000	0.00117	0.00853	-0.00860			
	(0.0147)	(0.0155)	(0.0135)			
(%Dem.$\geq p75$ at Destination)				-0.0523***	-0.0587***	-0.00743
Interaction with Year				(0.0112)	(0.0116)	(0.0101)
x 1970				0.0142	0.0308*	-0.0209
				(0.0163)	(0.0181)	(0.0148)
x 1980				0.00682	0.0272*	-0.00948
				(0.0141)	(0.0146)	(0.0129)
x 1990				0.00463	0.00734	0.0144
				(0.0143)	(0.0147)	(0.0127)
x 2000				-0.0472***	-0.0444***	-0.0112
				(0.0159)	(0.0164)	(0.0131)
Observations	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.689	0.692	0.787	0.689	0.692	0.787
Gravity	Y	Y	Y	Y	Y	Y
Destination*Year FE	Y	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y	Y
State-pair FE	Y			Y		
State-pair*Year FE		Y			Y	
SEA-pair FE			Y			Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.

Table 2.8: Geographic Sorting: Cross-Sections

	1960	1970	1980	1990	2000	1960	1970	1980	1990	2000
=1 if Different Party	-0.0524***	-0.00894	-0.0213**	-0.0618***	-0.0440***					
	(0.0137)	(0.0174)	(0.00850)	(0.0206)	(0.00780)					
(%Dem.$\geq p75$ at Destination)						-0.0584***	-0.0285**	-0.0302***	-0.0514***	-0.103***
						(0.0116)	(0.0129)	(0.0111)	(0.0102)	(0.0121)
Observations	254,520	258,572	258,572	258,572	258,572	254,520	258,572	258,572	258,572	258,572
R-squared	0.717	0.681	0.648	0.706	0.704	0.717	0.681	0.648	0.706	0.705
Gravity	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Destination*Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-pair FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The standard errors are two-way clustered at destination-origin SEA level.

2.3.4 Is Big Sort Real?

While showing the evidence of geographic sorting is meaningful, the more interesting empirical question is how does it change over time. By using the time series component in the data, the partisan sorting effect can be tracked starting from 1960 to evaluate the "big sort" hypothesis.

Table 7 reports the regression results of time-interacted $DifferentParty_{i,j,t}$ with several fixed effects specifications. The interaction terms are estimated relative to the coefficient in 1960. If the big sort hypothesis is true, the size of the coefficients of $DifferentParty_{i,j,t}$ should increase with time, and the interaction terms should be significant and negative. Overall, the results in Table 8 show no increasing trend over time. Column 1 shows that within state pairs, partisan sorting effect does not significantly differ over time. In Column 2, I include a time-varying state pair fixed effects. It is possible that there are other temporal patterns that may offset partisan sorting effects, due to changes such as states entering reciprocal agreements, or forming interstate compacts, etc. The state border effect has also increased during the same period, and this may also confound results (See Chapter 1). Column 2 shows that the negative effect of different party affiliations has weakly decreased over time after 1960. Once SEA-pair fixed effects are included in Column 3, however, I find that the partisan sorting effect is no longer significant in 1960, and weakly increased during 1970 and 1980, contrary to the results in Column 2. These effects may be the aftermath of the Civil Rights Act, in particular due to the realignment of the South.

In Columns 4 to 6, I estimate the time-interacted $Vote75_{i,j,t}$, the effect of having a large difference in vote shares. While both Columns 4 and 5 show a significantly increased partisan sorting effect in 2000, it loses significance once the SEA-pairs are controlled for, which indicates that this was more driven by party

[6] For example, Cho, et al. (2013) find that the Republicans have a stronger tendency to move to areas with copartisans.

differences across SEAs than over time. The cross-sectional results in Table 8 confirms this.

2.4 Conclusion

This study looks at the effect of political dissimilarities between areas, namely, different political party preferences, on migration flows for the period 1960-2000. By using presidential election outcomes combined with internal migration data, it provides evidences of partisan sorting that different party affiliations lowered the migration flows between the areas. I also found that increasing differences in party vote shares lowered the number of migrants. The results show that migration between same party areas were on average 1.03 times higher than between different party areas. Within SEA pairs, although the effect is weak, migration flows were 1.007 times lower if party affiliations became different over time, and one percent increase in vote share differences lowered migration flows by -0.07%. The partisan sorting effect also changes with destination characteristics, and is stronger if destination is less populated, more unemployed, less urbanized, and has more number of churches. The migrants are also less likely to move to different party areas if the areas are more partisan and greatly differ on support for the Democratic Party.

Contrary to the big sort hypothesis, I do not find an increase in partisan sorting over time. However, more locally defined data on political identification and migration is needed to correctly evaluate the temporal pattern in residential sorting. Also, extended data may find stronger evidences of sorting, as the data used in this study ends in 2000.

Although the partisan sorting effect found in this paper is small, it is significant and sheds light on earlier patterns in the aggregate migration data. The findings of this paper suggest that geographic sorting was present in the earlier years, and partisanship had effect on the migration choices of Americans.

Appendix B

Table B.1: Baseline Regression: Geographic Sorting for Alternative Years

	(1)	(2)	(3)	(4)	(5)
=1 if Different Party	-0.0411***	-0.0361***		0.000125	
	(0.00595)	(0.00463)		(0.00339)	
|%Democratic(Dest. - Origin)|			-0.435***		-0.106***
			(0.0418)		(0.0316)
Log(Distance)	4.826***	4.366***	4.341***		
	(0.928)	(0.811)	(0.811)		
Log(Distance)2	-1.170***	-1.081***	-1.076***		
	(0.160)	(0.151)	(0.151)		
Log(Distance)3	0.0706***	0.0675***	0.0671***		
	(0.00903)	(0.00933)	(0.00934)		
=1 if Contiguous	0.944***	0.938***	0.935***		
	(0.0509)	(0.0345)	(0.0345)		
=1 if State Border	-1.396***				
	(0.0485)				
Observations	1,288,808	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.668	0.689	0.689	0.787	0.787
Destination*Year FE	Y	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y	Y
Pair FE		State	State	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.
[b] Alternative election years of 1956, 1968, 1976, 1988, and 1996 are used.

Table B.2: Geographic Sorting and Other Controls for Alternative Years

	(1)	(2)	(3)	(4)
=1 if Different Party	-0.0226***		-0.000386	
	(0.00407)		(0.00335)	
|%Democratic(Dest. - Origin)|		-0.297***		-0.103***
		(0.0353)		(0.0304)
|Population(Dest. - Origin)|	-0.258***	-0.258***	-0.193***	-0.194***
	(0.0150)	(0.0150)	(0.0172)	(0.0172)
|Family Income Med.(Dest. - Origin)|	0.186***	0.185***	0.268***	0.269***
	(0.0395)	(0.0395)	(0.0327)	(0.0327)
|%High School(Dest.-Origin)|	-0.181**	-0.173**	0.118	0.125
	(0.0855)	(0.0857)	(0.0808)	(0.0807)
|Rent(Dest. - Origin)|	-0.211***	-0.209***	-0.203***	-0.201***
	(0.0352)	(0.0352)	(0.0273)	(0.0274)
|%Unemployed(Dest.-Origin)|	-2.340***	-2.322***	-0.242	-0.248
	(0.235)	(0.235)	(0.159)	(0.159)
|%Urban(Dest.-Origin)|	-0.259***	-0.258***	-0.226***	-0.227***
	(0.0261)	(0.0261)	(0.0359)	(0.0359)
|Church(Dest. - Origin)|	-0.0448***	-0.0446***	-0.0666***	-0.0676***
	(0.00948)	(0.00948)	(0.0129)	(0.0130)
|%Black(Dest.-Origin)|	-0.839***	-0.843***	-0.899***	-0.883***
	(0.0785)	(0.0785)	(0.131)	(0.130)
Observations	1,288,808	1,288,808	1,288,808	1,288,808
R-squared	0.695	0.696	0.788	0.788
Gravity Controls	Y	Y	Y	Y
Destination*Year FE	Y	Y	Y	Y
Origin*Year FE	Y	Y	Y	Y
Pair FE	State	State	SEA	SEA

[a] Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are two-way clustered at destination-origin SEA level.
[b] Alternative election years of 1956, 1968, 1976, 1988, and 1996 are used.

CHAPTER 3

Trust and Natural Disasters

3.1 Introduction

Trust improves economic performances as lower transaction costs and better government credibility facilitates innovation and investment (Knack and Keefer (1997)). How the trust itself is formed and promoted, however, is unclear and open to question. Trust is a persistent cultural norm that is transmitted from previous generations to the next (Dohmen et al. (2012); Tabellini (2008)), affected by past experiences, such as historical institutions, regulations or slave trade experiences (Aghion et al. (2010); Guiso et al. (2008); Nunn and Wantchekon (2011)), and also correlated with individual and community characteristics (Alesina and La Ferrara (2002); Alesina and Giuliano (2011)).

This study uses natural disasters to identify the effects on trust. Natural disaster provides a unique opportunity to study trust, as 1) it is an exogenous event; and 2) it damages lives and properties, and disrupts existing social connections. The U.S. is one of the top five countries that are most frequently hit by natural disasters. On average, the annual economic losses from natural disasters are estimated to be $11 billion and more than 400,000 people are affected every year.[1] What impact would such shocks have on interpersonal trust? In a state of emergency, the disaster victims will be forced to interact with people outside their usual social circle, and the experiences of increased social interactions may have

[1] EMDAT database (1950 - 2014)

a positive or negative effect on trust. It can be predicted that the experiences of prosocial behaviors such as cooperating and sharing with others can increase trust, whereas experiences of rivalry for resources or moral hazard such as looting can decrease trust.

Using two U.S. survey data sets, this paper provides evidences that support the former prediction. I show that individuals who recently have experienced natural disasters are more likely to trust others, and cooperate with neighbors. First, I use the self-reported trust values from the General Social Survey (GSS) for 1973-2010 and the disaster declaration data from the Federal Emergency Management Agency (FEMA) for 1953-2013. Using the regional differences in disaster frequencies, I find positive correlation between the level of trust and the number of contemporaneous natural disasters in the survey year. This effect is stronger for shocks that happen nearer to the survey date and the regressions using lagged shocks show that the effects are temporary. The results are robust after including individual and state controls.

I also use data from the Current Population Survey Volunteer Supplements for 2006-2013 with the FEMA declaration data to test whether disasters encourage more cooperation, which would promote trust. I find that the individuals who have experienced more number of natural disasters are more likely to have cooperated with their neighbors. This positive correlation is consistently significant after including individual and state controls.

This paper contributes to three strands of literature. First, it is related to the empirical studies on the short-run determinants of trust. Using the General Social Survey data, Alesina and La Ferrara (2002) find that trust is correlated with individual characteristics such as recent traumatic experiences, race, gender, income, education, etc, and community characteristics such as income and racial heterogeneity. In particular, they find that the racial and ethnic heterogeneity and income disparities have negative effects on trust. Alesina and Giuliano (2011) find that stronger family ties lower political participation and trust toward others.

This paper also adds to the literature studying the effects of natural disasters on individual preferences and economic outcomes. Natural disasters can have long-run effect on economic growth.[2] There are also several papers on the implications of disasters on risk, trust or time preferences.[3] This paper is especially related to empirical studies on the effect of natural disasters on interpersonal trust and social behaviors. To measure trust, studies either use field experiments or survey data. The empirical findings are contradictory. Whitt and Wilson (2007) find increased cooperation in Katrina evacuees. Becchetti et al. (2012) observe the long-term effect of Tsunami in Sri Lanka and find that the victims who suffer more losses also give and expect more in a dictator game. Castillo and Carter (2011) find nonlinear relationship between pro-social behaviors and negative weather shocks. They find that the victims of Hurricane Mitch are more likely to show cooperative behavior, but too severe a shock can decrease cooperation. Fleming et al. (2011) find no differences in trust level but decreases in trustworthiness for villagers affected by the Chilean Earthquake in 2010, whereas Andrabi and Das (2010) find increase in trust toward foreigners among those who receive foreign aids. Yamamura (2014) also finds that natural disasters increase individual investment in social capital. After the 1995 Kobe Earthquake, Japanese people were more likely to participate in voluntary community-building activities.

This is also related to recent literature that study the impact of lifetime shocks on socioeconomic behaviors using survey data. Giuliano and Spilimbergo (2014) show that the experience of recessions in "impressionable age" has effect on the preferences for redistribution as an adult. Madestam and Yanagizawa-Drott

[2] Previous literature on the long-run effect of natural disasters on economic growth are well surveyed in the working paper of Hsiang and Jina (2014).

[3] Several papers find that disaster victims are more risk-averse and impatient (Cassar et al. (2011); Cameron and Shah (2013)), whereas Eckel et al. (2009) claim that Katrina victims are more risk-loving and Callen (2011) finds increase in patience of those affected by the 2004 Indian Ocean Earthquake.

(2012) use rainfalls on Fourth of July as a proxy for childhood participation in political events and show that it has effect on political preferences as an adult. Shah and Steinberg (2013) also use rainfalls in India as productivity shocks, and show evidence for counter-cyclicality of human capital investment. They find that when there is drought, children are more likely to be enrolled and educated by lower opportunity cost of schooling. BenYishay (2013) observes the effect of early-life rainfall shocks on trust level in adulthood in sub-Saharan African countries, and finds that in regions with historically low trust level, low rainfall in early childhood has negative effect on trust as low trust is transmitted.

Two papers closely related to this study are Toya and Skidmore (2014) and Durante (2009). Toya and Skidmore (2014) observe the effect of the past experiences of natural disasters by disaster types on the trust score at the country level. Using the World Value Survey data, they find that the countries experiencing high number of storms will have higher trust level, whereas the countries with high number of floods will have lower trust level. Durante (2009) observes the historical weather variability in Europe using climate data from 1500-2000, and finds that the social trust today is higher in the regions with more variability. He explains this positive relationship by the persistent norms formed from the past experiences of cooperation and collective actions among farmers to insure against weather risks. While both studies look at the impact of historical weather shocks, this paper looks at the short-run effects of the contemporaneous natural disasters.

The rest of the paper is organized as follows. Section 2 describes data. Section 3 discusses the empirical strategy and explains results, and Section 4 concludes.

3.2 Data Description

3.2.1 Natural Disasters

The FEMA disaster declarations dataset includes information on the declaration dates, disaster incident types, declared areas, assistant preograms, incident begin and end dates. The state-level data is available from 1953 to 2013, and the county-level data is from 1959 to 2013.

When a state is hit by a natural disaster, the governor of the state can decide to request for a declaration to the President. Once the President determines that the disaster is "beyond the capabilities of the State" and makes a declaration, the federal assistance is provided to the state in need through FEMA.[4]

For the measure of natural diasters, I count the number of incidents declared as major disasters or emergency by FEMA within an year or month at the state or county level. Because the dataset does not include information on the actual magnitude of the disasters, I am limited to relying on the frequency of the disasters. However, it is reasonable to assume that all the events included in the data are of such magnitude that they have community-wide impacts as to require federal aids.

Figure 1 shows the distribution of the natural disasters by decades at the state level. The coastal states in the south have consistently high number of disasters. Texas, California, New York, and Oklahoma are the four most affected states.

The number of incidents per different types of disasters are shown in Table 1. Severe storms and floods are the top natural hazards that occur most frequently in the U.S., followed by Hurricanes. As shown in Figure 2, the number of declarations

[4]Stafford Act §401(a). All requests for a declaration by the President that a major disaster exists shall be made by the Governor of the affected State. Such a request shall be based on a finding that the disaster is of such severity and magnitude that effective response is beyond the capabilities of the State and the affected local governments and that Federal assistance is necessary.

have increased drastically over time, and two-thirds of the declarations have been made after 1990. This surge can be explained partly by the actual increase in extreme weather incidents and improvements in weather tracking technology, but it is also related with population increase, the expansion of federal role in disaster responses, and policy changes. FEMA declarations are also not free from political considerations. Elections and increasing media coverage of natural disasters can inflate political pressure on the president to declare.[5] For analysis, all of the regressions include year fixed effects to control for any common year specific trends.

Table 3.1: Disaster Types in FEMA Declarations

Disaster Types	
Coastal Storm	23
Dam/Levee Break	3
Drought	46
Earthquake	28
Fire	69
Fishing Losses	6
Flood	742
Freezing	18
Hurricane	303
Mud/Landslide	5
Severe Ice Storm	58
Severe Storm	846
Snow	152
Tornado	157
Tsunami	4
Typhoon	54
Volcano	4
Others	27
Total	2545

[a] FEMA Declarations 1953-2013.

[5]See Lindsay and McCarthy (2012).

3.2.2 Social Capital

For social capital measure, I use two U.S. survey data sets, the General Social Survey (GSS) and the Current Population Survey (CPS) Volunteer Supplements. The General Social Survey(GSS) is conducted by the National Opinion Research Center, and provides various social indicators by interviewing samples of national representatives annually between 1972 and 1993 (except in 1979, 1981, and 1992) and bi-annually since 1994. I use all the data from 1973 to 2010, as 1972 does not include geographic identifiers. Overall, this includes in total 24 surveys of 55,087 observations.

The Current Population Survey (CPS) is a monthly labor force survey conducted by the Bureau of Census. The Volunteer Supplement is carried out as a supplement to the September CPS every year and is available from 2002 to 2013. More than 100,000 persons age 15 years and over are interviewed annually from approximately 56,000 households. All the data from 2006 to 2012 are used, as the variable of interest, $Neighbor$, is available from 2006.

3.2.2.1 Trust

Trust is a cultural norm and is difficult to quantify or measure. Previous studies use trust games and experiments, surveys or other social indicators as proxies to measure trust. This paper follows the second approach, and use self-reported responses to survey questions on trust.[6]

I use the response to the following trust question from the General Social Survey (GSS): "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" As Alesina and La Ferrara (2002), my main variable $Trust$ takes a value of 1 if respondent chooses "Most

[6] There are literature that discusses whether the trust question actually measures trust (Glaeser et al. (2000); Sapienza et al. (2007)).

people can be trusted," and 0 if respondent chooses "Can't be too careful" or "Other, depends." The GSS provides geographic information of the respondents at the state level and also at the county level from 1994. Trust is persistent and demonstrates little variance over time. Figure 3 shows the average trust score by states in the periods 1972 - 2010.

Additional questions about helpfulness and fairness are also used for alternative measures of prosocial attitudes. The GSS includes the following two questions: "Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?" and "Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?" As with $Trust$, $Helpful$ is defined as 1 if respondent answers "Helpful," and 0 if the respondent chooses "Looks out for self" or "Depends." Similarly, $Fair$ is equal to 1 if respondent chooses "Fair," and 0 if repondent answers "Take advantage" or "Depends."

3.2.2.2 Neighbor

Social capital can also be measured by civic and volunteer activities. The Current Population Survey (CPS) Volunteer Supplements asks the following question: "Since September 1st of last year, have you worked with other people in your neighborhood to fix or improve something?" The variable *Neighbor* takes value of 1 if the respondent chose "Yes," and 0 if "No." I use this variable as a proxy for cooperation and social interactions.

I can benefit from using both survey data, as the GSS provides a longer time series of changes in attitude, while the CPS provides a larger cross-section of behavioral changes.

3.2.3 Individual and State Controls

Individual controls include education, family income, employment status, marital status, children, religion, gender, and race. *Education* is measured by years of schooling. $FamilyIncome$ is classified in 12 categories. $Children$ is a dummy that equals 1 if the respondent has any children. *Employment* status includes being employed, unemployed, or out of labor force. $Married$ is a dummy that equals 1 if the respondent is married, and 0 otherwise. $Male$ is a dummy for male. *Race* corresponds to being black, white, or others. Lastly, *Religion* is recoded as 1 for Protestant, 2 for Catholic, 3 for Jewish, and 4 for any other religions. It equals 0 if the respondent has no religion.

State controls consist of the state GDP, unemployment rate, the Gini coefficient, and the crime rates. State level GDPs are available from 1963 and are provided by the Bureau of Economic Analysis (BEA). Unemployment rates are collected by the Bureau of Labor Statistics (BLS) since 1976. For the state level Gini coefficients, I use the U.S.state-level income inequality data.[7] The crime statistics are from the FBI.

3.3 Empirical Strategy and Results

3.3.1 General Social Survey

3.3.1.1 Empirical Strategy

The empirical strategy is as follows:

$$Trust_{i,s,t} = \beta_0 + \beta_1 NaturalDisasters_{s,t} + \beta_2 X_i + \beta_3 Y_{s,t} + age_i + state_s + year_t + \varepsilon_{i,s,t}$$

where i,s and t corresponds to individual, state, and time accordingly. $Trust_{i,s,t}$ equals 1 if individual i in state s surveyed at year t answers that most people can

[7]http://www.shsu.edu/eco_mwf/inequality.html

be trusted. $NaturalDisasters_{s,t}$ is the number of natural disasters that hit state s at year t. X_i is a set of individual controls such as education, family income, employment and marital status, having any children, religious denomination, gender, and race. $Y_{s,t}$ is a set of state controls that include unemployment rate, GDP, Gini coefficient, and crime rates. All regressions include age, state and year fixed effects to control for any common trends in age groups, within state or within year. The standard errors are clustered at the state level and $Trust_{i,s,t}$ dummy is rescaled to take value between 0 and 100.

3.3.1.2 Results

Table 2 shows the baseline result of regressing contemporaneous natural disasters at the survey year on the level of trust. In column 1, I regress the trust variable on the number of natural disasters that happened in the contemporaneous year, and race, age, state, and year dummies. The coefficient for natural disasters is 0.889, positive and significant at the 1% level. This implies that the level of trust is rising in the number of natural distaters. Column 2 and 3 include the individual and state controls respectively, and in the last column, all specifications are included. The coefficient of interest remains positive and statistically significant at the 5% level after including all controls. Individuals residing in a state which was hit by more natural disasters in the survey year, are more likely to trust other people.

The individual characteristics included in the specification also demonstrates the correlation with trust. Consistent with Alesina and La Ferrara (2002), individuals with more education and higher income are more likely to be trusting, and the religious denominations are insignificant. The coefficients for employment status show that being out of labor force is negatively correlated with trust, while unemployed is insignificant. The state level controls such as the state GDP, the Gini coefficient, crime rates, and unemployment rate was included. While not reported in Table 2, all of the state level controls were insignificant, except the

Table 3.2: GSS: Trust and Natural Disasters

	(1)	(2)	(3)	(4)	Beta
Natural Disasters	0.889***	0.661**	0.736**	0.562**	0.0140
	(0.292)	(0.248)	(0.282)	(0.248)	
Education		3.524***		3.494***	0.2232
		(0.121)		(0.123)	
Family Income		1.019***		0.897***	0.0511
		(0.111)		(0.116)	
Unemployed		-1.085		-1.133	-0.0051
		(1.061)		(1.236)	
Out of Labor Force		-1.717***		-2.485***	-0.0239
		(0.598)		(0.649)	
Married		2.811***		2.964***	0.0305
		(0.622)		(0.682)	
Children		-2.059***		-2.692***	-0.0249
		(0.578)		(0.671)	
Protestant		0.931		1.140	0.0116
		(0.973)		(1.022)	
Catholic		-0.794		-0.783	-0.0069
		(1.336)		(1.348)	
Jewish		-2.886		-1.000	-0.0028
		(2.291)		(2.435)	
Other Religions		0.522		0.696	0.0028
		(1.042)		(1.008)	
Male		2.034***		1.716**	0.0176
		(0.711)		(0.780)	
Black		-18.16***		-17.93***	-0.1251
		(1.333)		(1.383)	
Other Races		-9.084***		-9.530***	-0.0437
		(1.029)		(1.077)	
State Controls			Y	Y	
State FE	Y	Y	Y	Y	
Year FE	Y	Y	Y	Y	
Age FE	Y	Y	Y	Y	
Observations	34,451	30,871	31,483	28,106	
R-squared	0.076	0.135	0.077	0.135	
Clusters	49	49	49	49	

[a] All specifications include age, state, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

state GDP variable, which was positive and significant. Due to inclusion of state fixed effects, the state level controls may not have large variation over time.

To understand the magnitude of the coefficients, the standardized beta coefficients for Column 4 are also reported in the last column. One standard deviation increase in natural disasters corresponds to 0.014 standard deviations increase in trust. This is comparable to half the effect of being married, and similar to the effect of being male, but much smaller compared to the effect of education or race.

Although significant, using natural disasters that happen in the concurrent year may weaken the relationship between trust and the events of natural disasters. The survey respondent may have not been residing in the state at the time of disaster. Also, the effect of natural disasters will decrease with time as people forget. Thus, I can further use the survey date information and examine the effects of disasters that happened more recently. The natural disasters can also be matched with the trust data using the month of the survey date to increase the probability of the individuals being present in the state at given period. This will also reduce the possibility of including disasters that happen after the survey as the GSS interviews are conducted throughout the year. I use shorter time periods, and count the number of natural disasters that occur in the same quarter, half-a-year, or three-quarter of the survey date.

In Table 3 Column 1, I show the effects of natural disasters that happened within the last three month from the surveyed month. The size of the coefficient increases by almost twice as much from 0.889 to 1.629, and it is strongly significant. The remaining columns show that with increase of the periods, the effect of natural disasters are reduced. This suggests that the positive effect of natural disasters on trust is largely driven by the recent events.

Table 3.3: GSS: Recent Disasters

	3 months	6 months	9 months	12 months
Natural Disasters	1.629***	0.827**	0.753***	0.474**
	(0.538)	(0.390)	(0.262)	(0.228)
Education	3.493***	3.493***	3.494***	3.493***
	(0.122)	(0.122)	(0.122)	(0.122)
Family Income	0.894***	0.896***	0.894***	0.894***
	(0.117)	(0.116)	(0.116)	(0.116)
Unemployed	-1.103	-1.113	-1.103	-1.121
	(1.240)	(1.242)	(1.237)	(1.235)
Out of Labor Force	-2.504***	-2.526***	-2.533***	-2.526***
	(0.651)	(0.652)	(0.654)	(0.652)
Married	2.991***	2.982***	3.001***	2.990***
	(0.679)	(0.676)	(0.677)	(0.677)
Children	-2.680***	-2.675***	-2.678***	-2.668***
	(0.670)	(0.670)	(0.671)	(0.671)
Protestant	1.144	1.128	1.116	1.124
	(1.040)	(1.033)	(1.034)	(1.034)
Catholic	-0.763	-0.771	-0.785	-0.782
	(1.352)	(1.346)	(1.349)	(1.350)
Jewish	-0.973	-0.990	-1.020	-1.000
	(2.473)	(2.458)	(2.466)	(2.455)
Other Religions	0.750	0.733	0.725	0.729
	(1.011)	(1.004)	(1.008)	(1.004)
Male	1.688**	1.682**	1.680**	1.687**
	(0.774)	(0.776)	(0.777)	(0.775)
Black	-17.94***	-17.93***	-17.96***	-17.95***
	(1.381)	(1.384)	(1.386)	(1.387)
Other Races	-9.643***	-9.643***	-9.650***	-9.638***
	(1.125)	(1.122)	(1.120)	(1.121)
State Controls	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Observations	28,059	28,059	28,059	28,059
R-squared	0.135	0.135	0.135	0.135
Clusters	49	49	49	49

[a] All specifications include age, state, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.

[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

Table 3.4: GSS: Heterogeneity of Natural Disasters

	(1)	(2)	(3)	(4)	(5)	(6)
Natural Disasters	0.642	2.391***	0.460	0.695*	0.761**	1.016***
	(0.804)	(0.757)	(0.283)	(0.353)	(0.327)	(0.376)
Education	3.501***					
	(0.156)					
(Education)xDisaster	-0.00617					
	(0.0592)					
Family Income		1.084***				
		(0.151)				
(Family Income)xDisaster		-0.177***				
		(0.0638)				
Unemployed			-2.743			
			(1.745)			
Out of Labor Force			-2.621***			
			(0.938)			
(Unemp.)xDisaster			1.376			
			(0.905)			
(Out of L.F.)xDisaster			0.118			
			(0.487)			
Married				3.262***		
				(0.947)		
(Married)xDisaster				-0.256		
				(0.444)		
Male					2.206**	
					(0.861)	
(Male)xDisaster					-0.430	
					(0.398)	
Children						-1.979**
						(0.901)
(Children)xDisaster						-0.623
						(0.459)
State Controls	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Observations	28,106	28,106	28,106	28,106	28,106	28,106
R-squared	0.135	0.135	0.135	0.135	0.135	0.135
Clusters	49	49	49	49	49	49

[a] All specifications include age, state, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

3.3.1.3 Heterogeneous Effects of Natural Diasasters

In this section, I interact $NaturalDisasters_{i,s,t}$ with various individual controls, and observe the potential heterogeneous effects of natural disasters. The effects of the natural disasters on trust can differ according to different subgroups of characteristics such as race or income. To test this, I include the interaction terms bewteen the natural disasters and the individual characteristics in addition to the baseline specification from Table 2 Column 4. The regression results for education, income, employment status, marital status, gender, and having children, are reported in Table 4. A negative interaction coefficient would imply that the positive effect of natural disasters is decreasing in the corresponding individual characteristic. The results show that only income has a significant interaction term. A negative and significant coefficient for income-interacted term indicates that having a high income will reduce the effect of natural disasters on trust. This can be explained by the fact that as high-income earners are better insured and are less likely to be affected by disasters, they are also likely to have less interactaction with other people at the time of disasters.

3.3.1.4 Robustness Checks: Alternative Measures

For robustness, I use two alternative measures for the societal trust. The GSS asks whether the respondents consider other people helpful and whether other people are fair. As with $Trust$, $Helpful$ and $Fair$ are dummies that equal 1 if the respondent answers that people are helpful or fair respectively. The regression results of the baseline specification using $Helpful$ and $Fair$ as the dependent variables are reported in Table 5. I find that the disaster shock is positive and significant for $Fair$, but insignificant for $Helpful$. In the last column, I include the average effect size (AES) coefficient for $Trust$, $Helpful$, and $Fair$. AES coefficient is calculated following Kling et al. (2007). I find a significant and positive AES

Table 3.5: GSS: Helpful and Fair

	Helful	Fair	AES
Natural Disasters	0.378	0.640**	0.0119***
	(0.308)	(0.259)	(0.00349)
Education	2.465***	2.723***	
	(0.132)	(0.118)	
Family Income	1.052***	1.023***	
	(0.118)	(0.113)	
Unemployed	1.934	-1.768	
	(1.204)	(1.418)	
Out of Labor Force	0.116	-0.696	
	(0.655)	(0.623)	
Married	2.024***	3.948***	
	(0.539)	(0.515)	
Children	-1.644***	-3.525***	
	(0.609)	(0.664)	
Protestant	3.003***	3.099***	
	(1.002)	(0.869)	
Catholic	1.935	3.477***	
	(1.396)	(1.124)	
Jewish	-4.165	0.540	
	(2.749)	(2.064)	
Other Religions	-1.407	0.497	
	(1.860)	(1.555)	
Male	-7.818***	-4.321***	
	(0.622)	(0.661)	
Black	-13.74***	-19.60***	
	(1.232)	(1.589)	
Other Races	-5.722***	-7.119***	
	(1.563)	(1.669)	
State Controls	Y	Y	Y
State FE	Y	Y	Y
Year FE	Y	Y	Y
Age FE	Y	Y	Y
Observations	26,706	26,612	25,740
R-squared	0.090	0.121	
Clusters	49	49	49

[a] All specifications include age, MSA, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.
[c] Dependent variables are helpful and fair. The last column reports the average effect size (AES) coefficient for all three measures of prosocial attitudes.

coefficient.

3.3.1.5 Robustness Checks: MSA Level Data

Using the geographic identifier variable, I further identify the disaster effects at a more disaggregated level. The GSS includes the geographic information of the respondents at the MSA level for all samples from 1973, and at the county level from 1994. As an event of natural disaster in one part of the state may have zero effect on the residents of other parts in the state, using data at the MSA level would increase the likelihood of the survey respondents being an actual disaster victim, or a part of the community hit by natural disasters. As demonstrated in Tables 6 and 7, I run the same specifications using MSA level shocks, with year and MSA fixed effects and clustered standard errors at the MSA level. In order to avoid losing observations, the non-MSA areas in the states are included as a single non-MSA for each state (Luttmer (2001)). I find that $NaturalDisasters_{i,s,t}$ remains significant and positive at 5% level after including all the individual and the state controls.

In Table 7, I repeat the exercise using recent disaster shocks. Similar to Table 3, the coefficients are reduced as more later disasters are counted. In the last column, all of natural disasters that happened in the last 12 months prior to surveyed month are included, and the positive effect on trust becomes insignificant. This confirms that the positive relationship is driven by recent shocks. Individuals are most affected by natural disasters that took place in the recent few months.

3.3.2 Current Population Survey Volunteer Supplement

Using the GSS data in the previous section, I showed that natural disasters have a positive and significant effect on trust, and it is robust across different specifications. Suggested explanation for this positive association was a possible increase

Table 3.6: GSS: Trust and Natural Disasters at MSA

	(1)	(2)	(3)	(4)
Natural Disasters	0.613**	0.528**	0.652**	0.618**
	(0.268)	(0.260)	(0.291)	(0.270)
Education		3.445***		3.420***
		(0.103)		(0.109)
Family Income		0.903***		0.788***
		(0.120)		(0.124)
Unemployed		-0.700		-0.506
		(1.334)		(1.365)
Out of Labor Force		-2.488***		-2.878***
		(0.728)		(0.720)
Married		3.304***		3.354***
		(0.611)		(0.616)
Children		-2.452***		-2.856***
		(0.610)		(0.613)
Protestant		0.601		0.866
		(1.031)		(1.074)
Catholic		-1.205		-1.237
		(1.214)		(1.253)
Jewish		-2.185		-0.527
		(2.080)		(2.041)
Other Religions		-0.145		-0.140
		(1.418)		(1.468)
Male		1.660***		1.586**
		(0.622)		(0.653)
Black		-17.64***		-17.26***
		(1.035)		(1.077)
Other Races		-8.956***		-8.961***
		(1.229)		(1.291)
State Controls			Y	Y
MSA FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Race FE	Y	Y	Y	Y
Observations	30,527	27,215	29,026	25,821
R-squared	0.088	0.144	0.089	0.144
Clusters	190	190	190	190

[a] All specifications include age, MSA, and year fixed effects. Standard errors are in paranthesis and clustered at the MSA level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

Table 3.7: GSS: Recent Disasters at MSA

	3 months	6 months	9 months	12 months
Natural Disasters	1.223**	0.927**	0.632**	0.255
	(0.517)	(0.389)	(0.302)	(0.279)
Education	3.421***	3.421***	3.421***	3.420***
	(0.108)	(0.108)	(0.108)	(0.108)
Family Income	0.786***	0.785***	0.786***	0.787***
	(0.124)	(0.124)	(0.123)	(0.123)
Unemployed	-0.502	-0.506	-0.489	-0.495
	(1.366)	(1.368)	(1.365)	(1.366)
Out of Labor Force	-2.932***	-2.939***	-2.933***	-2.924***
	(0.721)	(0.721)	(0.721)	(0.720)
Married	3.387***	3.386***	3.395***	3.390***
	(0.614)	(0.613)	(0.611)	(0.613)
Children	-2.861***	-2.861***	-2.859***	-2.852***
	(0.616)	(0.615)	(0.614)	(0.614)
Protestant	0.866	0.843	0.856	0.855
	(1.076)	(1.074)	(1.075)	(1.075)
Catholic	-1.204	-1.227	-1.221	-1.218
	(1.258)	(1.255)	(1.255)	(1.256)
Jewish	-0.513	-0.515	-0.490	-0.485
	(2.041)	(2.036)	(2.037)	(2.037)
Other Religions	-0.0870	-0.112	-0.116	-0.115
	(1.467)	(1.466)	(1.464)	(1.465)
Male	1.573**	1.564**	1.561**	1.558**
	(0.656)	(0.656)	(0.656)	(0.656)
Black	-17.21***	-17.20***	-17.23***	-17.22***
	(1.078)	(1.082)	(1.079)	(1.077)
Other Races	-9.055***	-9.066***	-9.070***	-9.061***
	(1.293)	(1.294)	(1.294)	(1.293)
State Controls	Y	Y	Y	Y
MSA FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Race FE	Y	Y	Y	Y
Observations	25,777	25,777	25,777	25,777
R-squared	0.144	0.144	0.144	0.144
Clusters	190	190	190	190

[a] All specifications include age, MSA, and year fixed effects. Standard errors are in paranthesis and clustered at the MSA level.*** p<0.01, ** p<0.05, * p<0.1.

[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

in positive social interactions. In this section, I use the CPS data and provide supporting evidences of increased cooperative behaviors.

3.3.2.1 Empirical Strategy

The baseline regression for the CPS is as follows:

$$Neighbor_{i,s,t} = \beta_0 + \beta_1 NaturalDisasters_{s,t} + \beta_2 X_i + \beta_3 Y_{s,t} + age_i + state_s + year_t + \varepsilon_{i,s,t}$$

where i,s and t corresponds to individual, state, and time accordingly. $Neighbor_{i,s,t}$ equals 1 if individual i in state s surveyed at year t answers that i has worked with neighbors in the last 12 months. As with the GSS, $NaturalDisasters_{s,t}$ is the number of natural disasters that hit state s at year t. The set of all individual controls and state controls, X_i and $Y_{s,t}$, are also the same excluding the religion denominations which are not asked in the CPS. All the regressions include age, state and year fixed effects to control for any common trends in age groups, within state or within year. The standard errors are clustered at the state level and $Neighbor_{i,s,t}$ dummy is rescaled to take value between 0 and 100.

3.3.2.2 Results

The baseline regression result is shown in Table 8. I find that $NaturalDisasters_{i,s,t}$ is positively correlated with working with neighbors and is consistently significant at 1% level after including all individual and state controls, and the set of fixed effects. The result shows that individual i in a state hit by more natural disasters are more likely to have worked with one's neighbors to fix or improve things. This result is robust across specifications through Column 1-4.

The result on individual characteristics show who are more likely to have interacted with neighbors. Individual with higher income and more education are more likely to have cooperated with neighbors. Being out of labor force is nega-

Table 3.8: CPS: Helping Neighbors and Natural Disasters

	(1)	(2)	(3)	(4)
Natural Disasters	0.199***	0.191***	0.175***	0.168***
	(0.0464)	(0.0501)	(0.0402)	(0.0446)
Education		1.343***		1.343***
		(0.0602)		(0.0604)
Family Income		0.207***		0.207***
		(0.0200)		(0.0201)
Unemployed		1.560***		1.558***
		(0.171)		(0.171)
Out of Labor Force		-1.181***		-1.181***
		(0.143)		(0.142)
Married		1.026***		1.027***
		(0.122)		(0.122)
Children		1.836***		1.837***
		(0.199)		(0.198)
Male		1.159***		1.159***
		(0.135)		(0.135)
Black		0.188		0.192
		(0.341)		(0.342)
Other Races		-2.853***		-2.855***
		(0.282)		(0.282)
State Controls			Y	Y
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Observations	651,799	591,122	651,799	591,122
R-squared	0.021	0.041	0.020	0.041
Clusters	51	51	51	51

[a] All specifications include age, state, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

Table 3.9: CPS: Recent Disasters

	3months	6months	9months	12months
Natural Disasters	0.198***	0.235***	0.193***	0.209***
	(0.0703)	(0.0670)	(0.0518)	(0.0418)
Education	1.343***	1.343***	1.343***	1.343***
	(0.0604)	(0.0604)	(0.0604)	(0.0604)
Family Income	0.207***	0.207***	0.207***	0.207***
	(0.0201)	(0.0201)	(0.0201)	(0.0201)
Unemployed	1.557***	1.557***	1.559***	1.559***
	(0.171)	(0.171)	(0.171)	(0.171)
Out of Labor Force	-1.182***	-1.182***	-1.182***	-1.182***
	(0.142)	(0.142)	(0.142)	(0.142)
Married	1.028***	1.027***	1.027***	1.027***
	(0.122)	(0.122)	(0.122)	(0.122)
Children	1.836***	1.836***	1.837***	1.837***
	(0.198)	(0.198)	(0.198)	(0.198)
Male	1.159***	1.159***	1.159***	1.159***
	(0.135)	(0.135)	(0.135)	(0.135)
Black	0.191	0.192	0.192	0.193
	(0.342)	(0.342)	(0.342)	(0.342)
Other Races	-2.856***	-2.855***	-2.855***	-2.855***
	(0.282)	(0.282)	(0.282)	(0.282)
State Controls	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Observations	591,122	591,122	591,122	591,122
R-squared	0.041	0.041	0.041	0.041
Clusters	51	51	51	51

[a] All specifications include age, MSA, and year fixed effects. Standard errors are in paranthesis and clustered at the MSA level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

tively correlated with cooperation, as the individuals in this group include retirees and housewives who are socially less active. Males are significantly more likely to cooperate, but this may be because the question asks whether the respondent has worked with neighbors to "fix or improve things" which can be male-biased activities. Being married and having children also increases the likelihood of interactions with neighbors. Interstingly, individuals of other races, non-white and non-black, are less likely to have worked with neighbors.

Table 9 includes the result with disaster shocks that occurred more recently. As the CPS Volunteer Supplement is always conducted in September and most disasters happen before winter, it is likely that the yearly variable $NaturalDisasters_{i,s,t}$ are natural disasters that occur before the survey date. Nonetheless, to minimize the possibility of including forward disasters and out-of-state migrations, I use the number of disasters that happen before September within shorter time periods. Regression using shocks that happen in contemporaneous quarter, half-a-year, and three-quarter are reported in Table 9. Interestingly, for $Neighbor_{i,s,t}$, the effects remain largely the same, and I do not find a decrease over time. I find that the positive behavior is more persistent, and lagged disaster shocks continue to have a positive effect on $Neighbor_{i,s,t}$. While not reported in Table 9,

I also test for heterogeneous effects in Table 10. I find no sigificant effects for any of the interaction terms between the natural disasters and the individual controls.

3.3.3 Disaster Types

In this section, I show the effect of natural disasters on trust by the disaster types. The types of natural disasters can widely differ in severity of the damage, the length of affected time, the coverage of the affected area, and such. Also, some types of disasters, such as hurricane, can be more anticipated than an earthquake,

Table 3.10: CPS: Heterogeneity of Natural Disasters

	(1)	(2)	(3)	(4)	(5)	(6)
Natural Disasters	0.638	0.111	0.156***	0.177***	0.180***	0.145***
	(0.726)	(0.116)	(0.0555)	(0.0603)	(0.0528)	(0.0520)
Education	1.362***					
	(0.0682)					
(Education)xDisaster	-0.0117					
	(0.0182)					
Family Income		0.198***				
		(0.0278)				
(Family Income)xDisaster		0.00542				
		(0.0106)				
Unemployed			1.536***			
			(0.219)			
Out of Labor Force			-1.238***			
			(0.167)			
(Unemp.)xDisaster			0.0135			
			(0.115)			
(Out of L.F.)xDisaster			0.0348			
			(0.0751)			
Married				1.050***		
				(0.205)		
(Married)*ND				-0.0141		
				(0.0731)		
Male					1.198***	
					(0.166)	
(Married)xDisaster					-0.0241	
					(0.0460)	
Children						1.692***
						(0.231)
(Children)xDisaster						0.0897
						(0.0823)
State Controls	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Observations	591,122	591,122	591,122	591,122	591,122	591,122
R-squared	0.041	0.041	0.041	0.041	0.041	0.041
Clusters	51	51	51	51	51	51

[a] All specifications include age, state, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.
[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates.

Table 3.11: GSS: By Disaster Types

	(1) Storm	(2) Flood	(3) Hurricane	(4) Tornado	(5) Earthquake	(6) Fire	(7) Volcano
Natural Disaster	0.893**	-0.0259	-0.0327	0.756	-0.485	1.624**	-13.04***
	(0.368)	(0.536)	(0.704)	(1.192)	(1.822)	(0.781)	(1.856)
Individual Controls	Y	Y	Y	Y	Y	Y	Y
State Controls	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y	Y
Race FE	Y	Y	Y	Y	Y	Y	Y
Observations	28,106	28,106	28,106	28,106	28,106	28,106	28,106
R-squared	0.135	0.135	0.135	0.135	0.135	0.135	0.135
Clusters	49	49	49	49	49	49	49

[a] All specifications include age, MSA, and year fixed effects. Standard errors are in paranthesis and clustered at the state level.*** p<0.01, ** p<0.05, * p<0.1.

[b] State level controls include state GDP, Gini coefficient, violent crime rate, property crime rate, and unemployment rates. Individual controls include education, income, employment status, marital status, having children, religion, gender, and race.

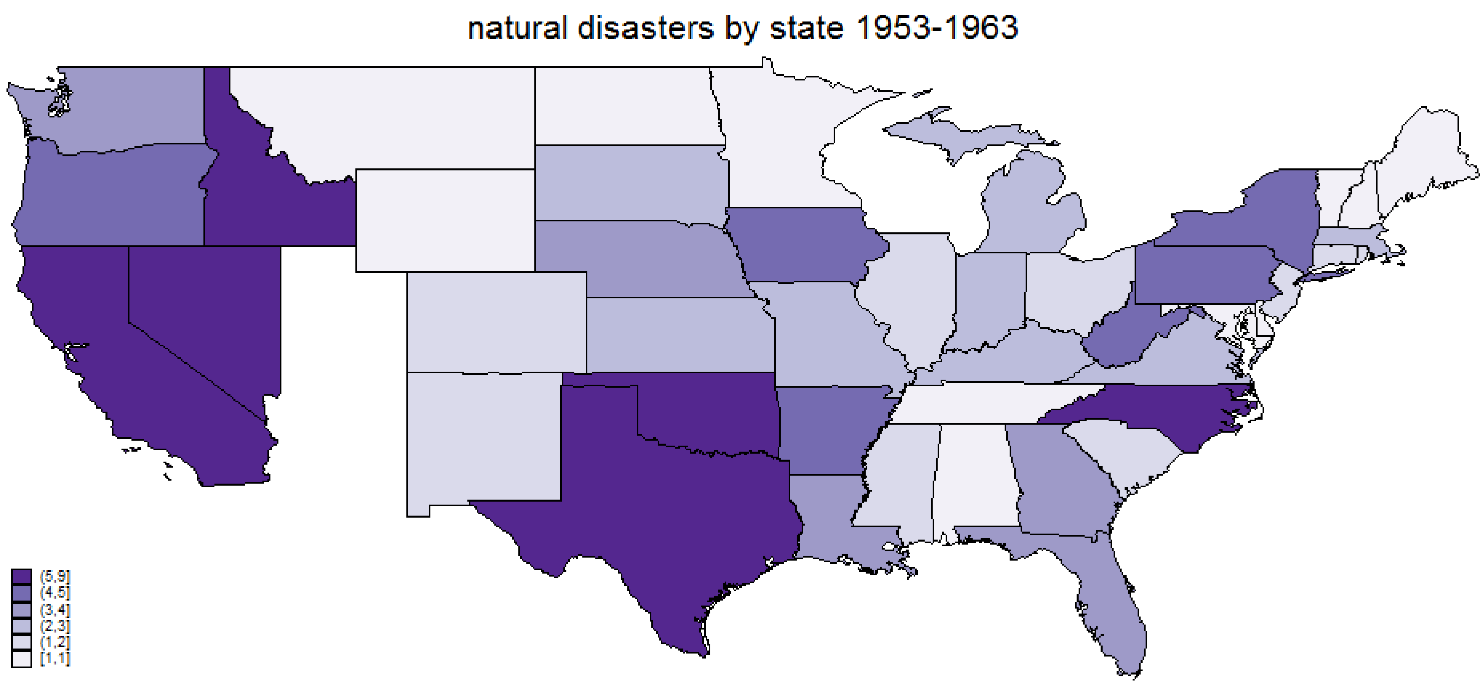
for example. The frequencies of the disasters are very different as well.

In Table 11, I show the baseline regression for different types of disasters. Storm is one of the most frequent natural disasters in FEMA declarations, and I find that it has a positive effect on trust at 0.893. Fire is another type of disasters that is highly significant and positive in increasing trust. Interestingly, the coefficient of volcano is strongly negative, but there were only 4 volcanic incidences counted in the data.

3.4 Conclusion

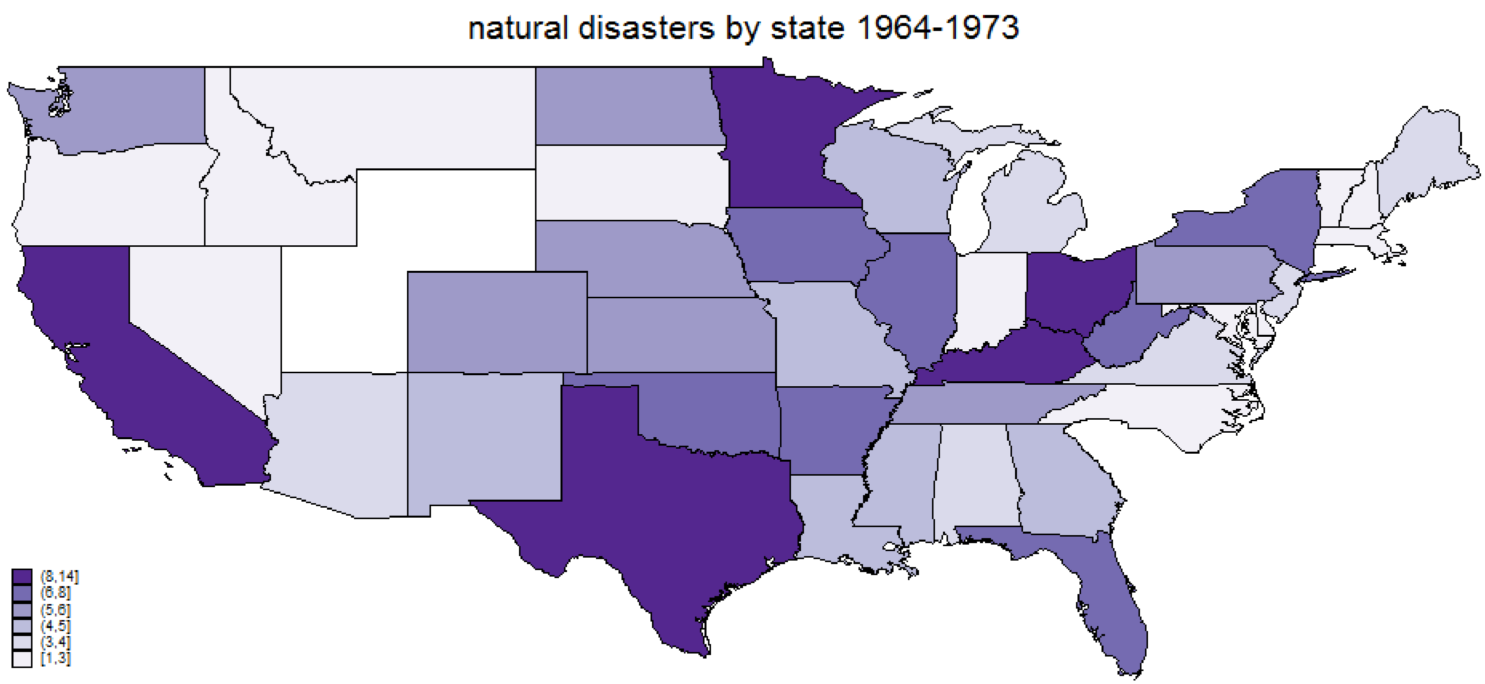
Natural disasters are exogeneous shocks that can impact our behaviors and preferences. In the time of natural disasters, people tend to receive or give help to those around them. This experience of positive social interaction can have positive effect on trust.

Using the GSS and the CPS data, I show that the experience of recent natural disasters increase the prosocial attitude and behaviors. Individuals who recently experienced natural disasters are more likely to trust other people. I show evidences that the increase in trust can be explained by positive social interactions. I find that individuals in the disaster affected areas were more likely to have cooperated with neighbors. This result is robust controlling for different individual and state controls. This paper provides suggestive evidence of short-run positive correlation between the natural disasters and trust. Further research using improved historical data on a more disaggregated level will shed light on the long-run dynamics between the natural disasters and social capital.



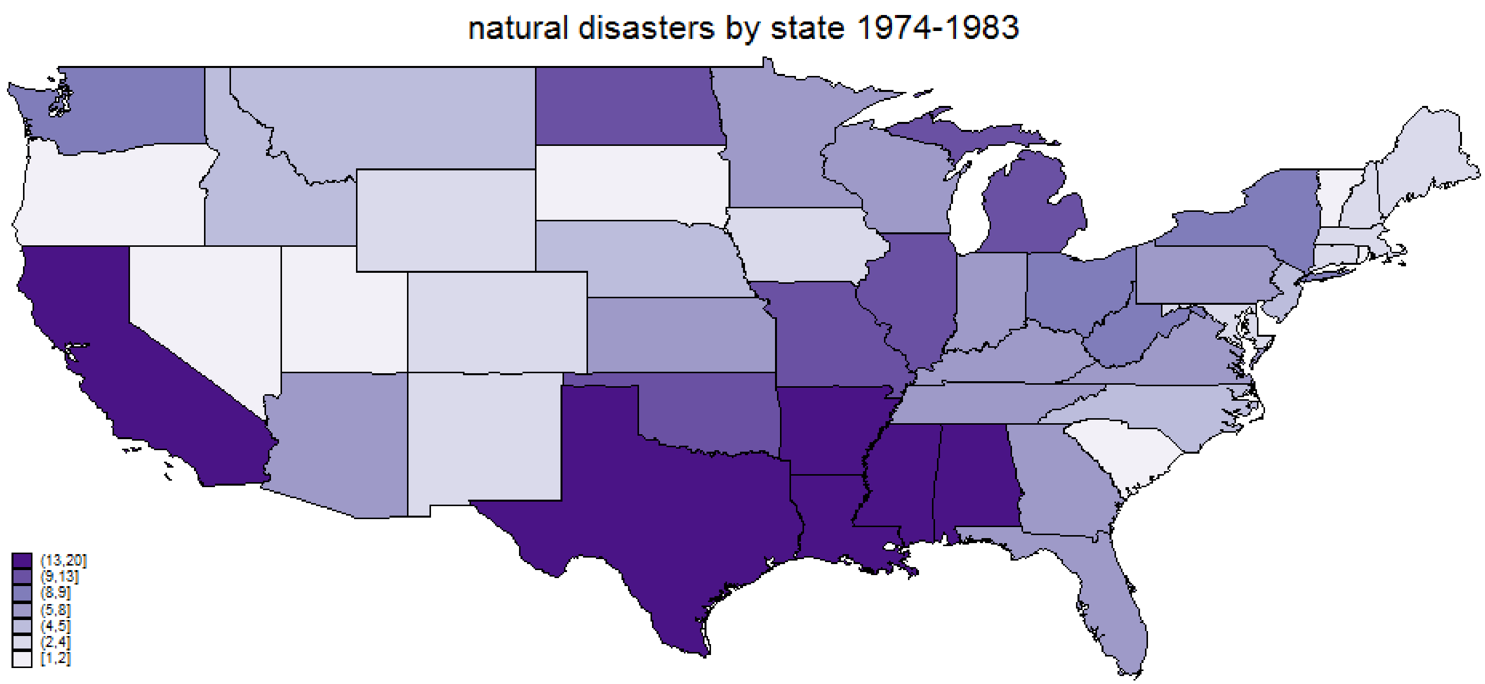
(a) 1953-1963

Figure 3.1: State Level Natural Disasters by Decade



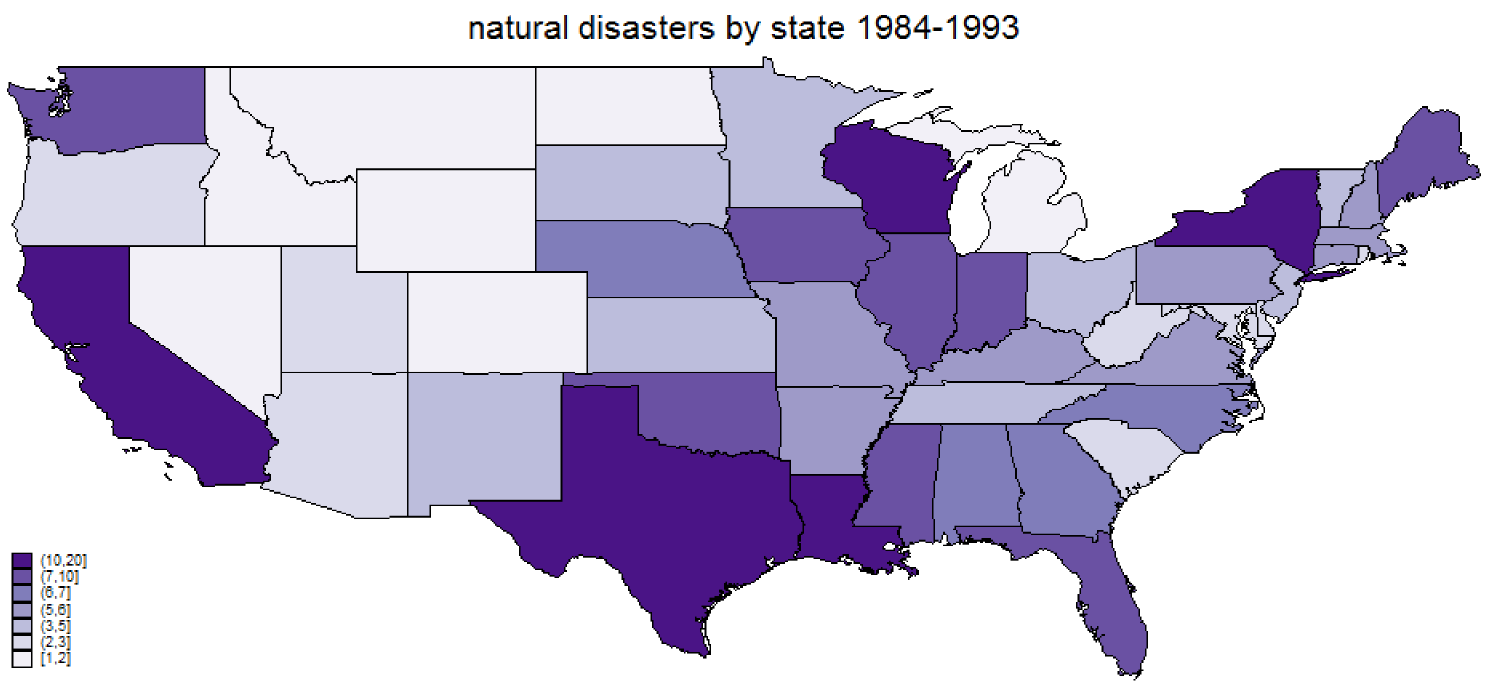
(a) 1964-1973

Figure 3.1: State Level Natural Disasters by Decade



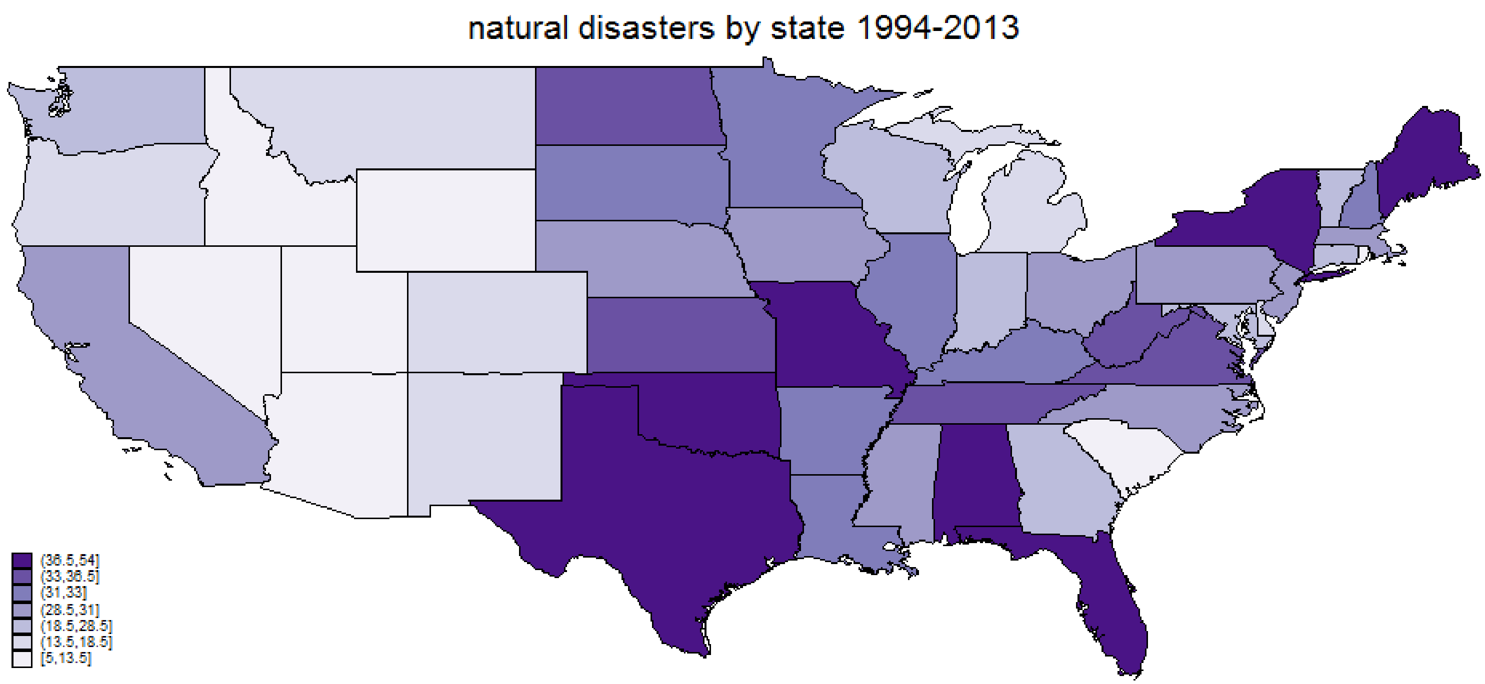
(a) 1974-1983

Figure 3.1: State Level Natural Disasters by Decade



(a) 1984-1993

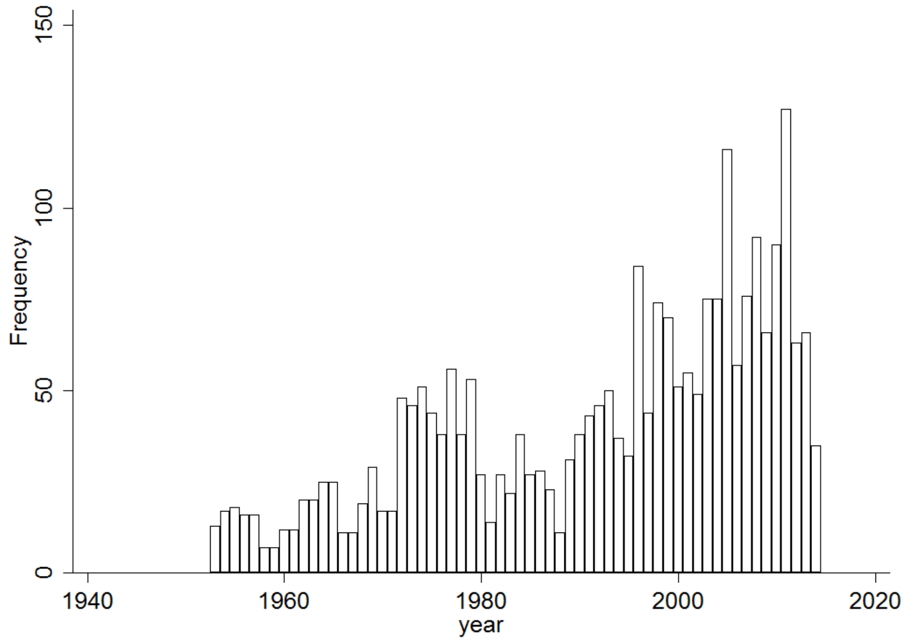
Figure 3.1: State Level Natural Disasters by Decade



(a) 1994-2013

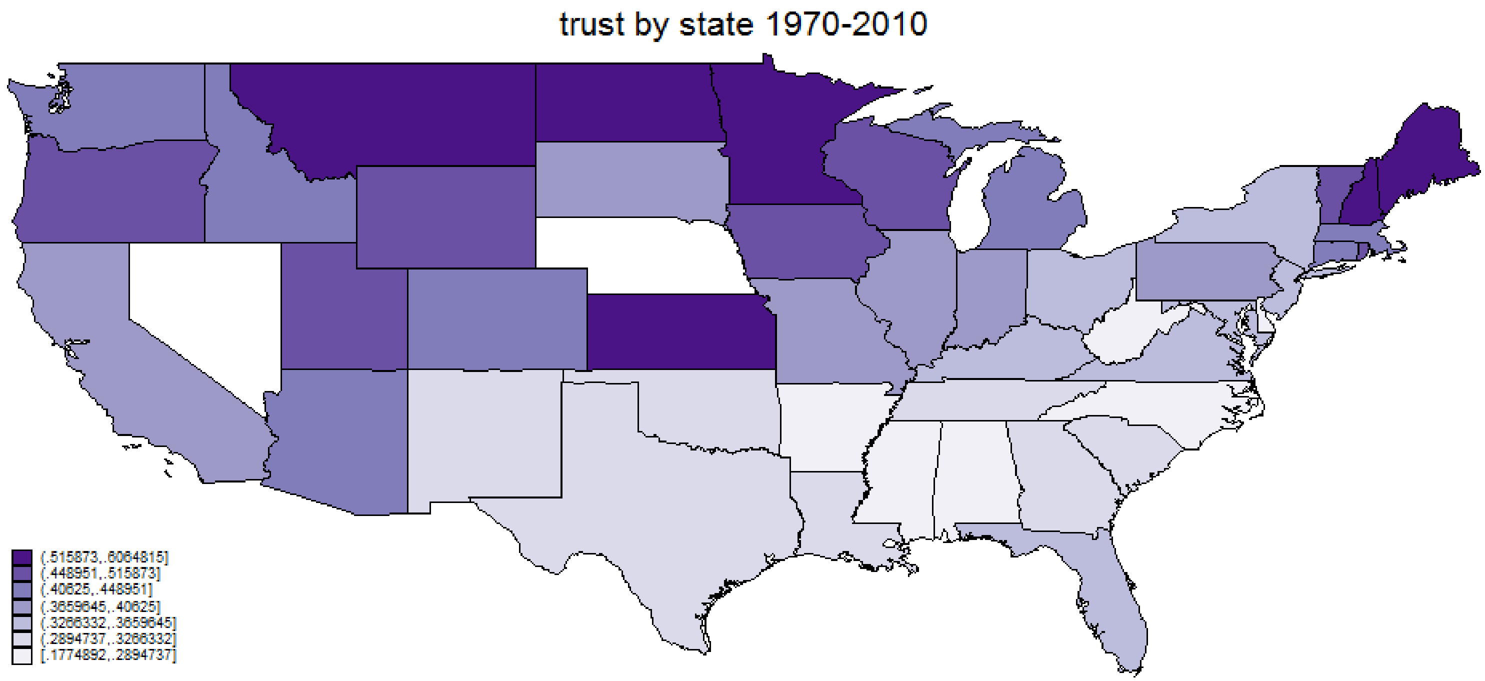
Figure 3.1: State Level Natural Disasters by Decade

Figure 3.2: Distribution of FEMA Declaration between 1953-2014



Note: Author's calculation using migration data from the U.S. decennial Censuses and the presidential election outcomes. Within groups of different party and same party SEAs, the areas are further categorized by the differences in vote shares. The migrants are in shares of total population.

Figure 3.3: Trust by State 1970-2010



Note: Author's calculation of survey responses from GSS between 1970-2010.

Bibliography

Alan I. Abramowitz and Kyle L. Saunders. Is polarization a myth? The Journal of Politics, 70(2):542–555, 2008. doi: 10.1017/S0022381608080493. URL https://doi.org/10.1017/S0022381608080493.

Samuel J Abrams and Morris P Fiorina. "the big sort" that wasn't: A skeptical reexamination. PS: Political Science & Politics, 45(2):203–210, 2012.

Philippe Aghion, Yann Algan, Pierre Cahuc, and Andrei Shleifer. Regulation and Distrust. Quarterly Journal of Economics, 125(3):1015–1049, August 2010.

Delina E. Agnosteva, James E. Anderson, and Yoto V. Yotov. Intra-national trade costs: Measurement and aggregation. Working Paper 19872, National Bureau of Economic Research, January 2014. URL http://www.nber.org/papers/w19872.

Alberto Alesina and Paola Giuliano. Family ties and political participation. Journal of the European Economic Association, 9(5):817–839, 2011.

Alberto Alesina and Eliana La Ferrara. Who trusts others? Journal of Public Economics, 85(2):207–234, 2002.

James E. Anderson and Eric van Wincoop. Gravity with gravitas: A solution to the border puzzle. American Economic Review, 93(1):170–192, March 2003. doi: 10.1257/000282803321455214. URL http://www.aeaweb.org/articles?id=10.1257/000282803321455214.

Tahir Andrabi and Jishnu Das. In aid we trust: Hearts and minds and the pakistan earthquake of 2005. World Bank Policy Research Working Paper Series, Vol, 2010.

Mohammad Arzaghi and Anil Rupasingha. Migration as a way to diversify: Evidence from rural to urban migration in the u.s. Journal of Regional

Science, 53(4):690–711, 2013. ISSN 1467-9787. doi: 10.1111/jors.12055. URL http://dx.doi.org/10.1111/jors.12055.

Leonardo Becchetti, Stefano Castriota, and Pierluigi Conzo. Calamity, aid and indirect reciprocity: The long run impact of tsunami on altruism. 2012.

Ariel BenYishay. Childhood shocks and trust in adulthood: Impacts of early–life rainfall in sub–saharan africa. 2013.

Jeffrey H. Bergstrand, Mario Larch, and Yoto V. Yotov. Economic integration agreements, border effects, and distance elasticities in the gravity equation. European Economic Review, 78(Supplement C):307 – 327, 2015. ISSN 0014-2921. doi: https://doi.org/10.1016/j.euroecorev.2015.06.003. URL http://www.sciencedirect.com/science/article/pii/S0014292115000847.

Simon Bertoli and Jesus Fernandez-Huertas Moraga. Multilateral resistance to migration. Journal of Development Economics, 102:79 – 100, 2013. ISSN 0304-3878. doi: http://dx.doi.org/10.1016/j.jdeveco.2012.12.001. URL http://www.sciencedirect.com/science/article/pii/S0304387812001034. Migration and Development.

Bill Bishop. The Big Sort: Why the Clustering of Like-Minded America Is Tearing Us Apart. Houghton Mifflin Harcourt, 2009.

Michael Callen. Catastrophes and time preference: Evidence from the indian ocean earthquake. Unpublished Manuscript, University of California at San Diego, 2011.

Lisa Cameron and Manisha Shah. Risk–taking behavior in the wake of natural disasters. Technical report, National Bureau of Economic Research, 2013.

Alessandra Cassar, Andrew Healy, and Carl Von Kessler. Trust, risk, and time preferences after a natural disaster: Experimental evidence from thailand. Unpublished manuscript, 2011.

Marco Castillo and Michael Carter. Behavioral responses to natural disasters. Technical Report 1026, George Mason University, Interdisciplinary Center for Economic Science, 2011.

Wendy K. Tam Cho, James G. Gimpel, and Iris S. Hui. Voter migration and the geographic sorting of the american electorate. Annals of the Association of American Geographers, 103(4):856–870, 2013. doi: 10.1080/00045608.2012.720229. URL https://doi.org/10.1080/00045608.2012.720229.

Pierre-Philippe Combes, Miren Lafourcade, and Thierry Mayer. The trade-creating effects of business and social networks: evidence from france. Journal of International Economics, 66(1):1 – 29, 2005. ISSN 0022-1996. doi: http://dx.doi.org/10.1016/j.jinteco.2004.07.003. URL http://www.sciencedirect.com/science/article/pii/S0022199604001138.

Paul S. Davies, Michael J. Greenwood, and Haizheng Li. A conditional logit approach to u.s. state-to-state migration. Journal of Regional Science, 41(2):337–360, 2001. ISSN 1467-9787. doi: 10.1111/0022-4146.00220. URL http://dx.doi.org/10.1111/0022-4146.00220.

Carroll Doherty, Jocelyn Kiley, and Bridget Jameson. Partisanship and political animosity in 2016. Washington, DC: Pew Research Center, 2016.

Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. The intergenerational transmission of risk and trust attitudes. 79(2):645–677, 2012.

Ruben Durante. Risk, cooperation and the economic origins of social trust: an empirical investigation. Available at SSRN 1576774, 2009.

Catherine C. Eckel, Mahmoud A. El-Gamal, and Rick K. Wilson. Risk loving after the storm: A bayesian–network study of hurricane katrina evacuees. Journal of Economic Behavior & Organization, 69(2):110–124, 2009.

Morris P. Fiorina and Samuel J. Abrams. Political polarization in the american public. Annual Review of Political Science, 11(1):563–588, 2008. doi: 10.1146/annurev.polisci.11.053106.153836. URL https://doi.org/10.1146/annurev.polisci.11.053106.153836.

David Fleming, Alberto Chong, and Hernan Bejarano. Do natural disasters affect trust/trustworthiness? evidence from the 2010 chilean earthquake. Technical Report 104522, Agricultural and Applied Economics Association, 2011.

Peter Ganong and Daniel Shoag. Why has regional income convergence in the u.s. declined? Journal of Urban Economics, 102:76 – 90, 2017. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2017.07.002. URL https://www.sciencedirect.com/science/article/pii/S0094119017300591.

Matthew Gentzkow. Polarization in 2016. Toulouse Network of Information Technology white paper, 2016.

Matthew Gentzkow, Jesse M. Shapiro, and Matt Taddy. Measuring polarization in high-dimensional data: Method and application to congressional speech. Working Paper 22423, National Bureau of Economic Research, July 2016. URL http://www.nber.org/papers/w22423.

James G. Gimpel and Iris S. Hui. Seeking politically compatible neighbors? the role of neighborhood partisan composition in residential sorting. Political Geography, 48:130 – 142, 2015. ISSN 0962-6298. doi: https://doi.org/10.1016/j.polgeo.2014.11.003. URL http://www.sciencedirect.com/science/article/pii/S0962629815000049.

Paola Giuliano and Antonio Spilimbergo. Growing up in a recession. Review of Economic Studies, 81(2):787–817, 2014.

Edward L. Glaeser and Bryce A. Ward. Myths and realities of american political geography. Journal of Economic Perspectives, 20(2):119–144, June 2006. doi: 10.1257/jep.20.2.119. URL http://www.aeaweb.org/articles?id=10.1257/jep.20.2.119.

Edward L Glaeser and Bryce A Ward. The causes and consequences of land use regulation: Evidence from greater boston. Journal of Urban Economics, 65(3): 265–278, 2009.

Edward L. Glaeser, David I. Laibson, Jose A. Scheinkman, and Christine L. Soutter. Measuring trust. Quarterly Journal of Economics, 115(3):pp. 811–846, 2000.

Edward L. Glaeser, Joseph Gyourko, and Raven E. Saks. Why have housing prices gone up? The American Economic Review, 95(2):329–333, 2005. ISSN 00028282. URL http://www.jstor.org/stable/4132842.

Michael J. Greenwood. Chapter 12 internal migration in developed countries. volume 1, Part B of Handbook of Population and Family Economics, pages 647 – 720. Elsevier, 1997. doi: http://dx.doi.org/10.1016/S1574-003X(97)80004-9. URL http://www.sciencedirect.com/science/article/pii/S1574003X97800049.

Michael J. Greenwood and Gary L. Hunt. The early history of migration research. International Regional Science Review, 26(1):3–37, 2003. URL https://EconPapers.repec.org/RePEc:sae:inrsre:v:26:y:2003:i:1:p:3-37.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Long term persistence. Technical report, National Bureau of Economic Research, 2008.

Peter K. Hatemi, Rose McDermott, Lindon J. Eaves, Kenneth S. Kendler, and Michael C. Neale. Fear as a disposition and an emotional state: A genetic and environmental approach to out–group political preferences. American Journal of Political Science, 57(2):279–293, 2013. doi: 10.1111/ajps.12016. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12016.

Elhanan Helpman, Marc Melitz, and Yona Rubinstein. Estimating trade flows: Trading partners and trading volumes. The Quarterly Journal of Economics, 123(2):441–487, 2008.

Russell Hillberry and David Hummels. Intranational home bias: Some explanations. The Review of Economics and Statistics, 85(4):1089–1092, 2003. ISSN 00346535, 15309142. URL http://www.jstor.org/stable/3211830.

Solomon M. Hsiang and Amir S. Jina. The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones. Technical report, National Bureau of Economic Research, 2014.

Iris Hui. Who is your preferred neighbor? partisan residential preferences and neighborhood satisfaction. American Politics Research, 41(6):997–1021, 2013. doi: 10.1177/1532673X13482573. URL https://doi.org/10.1177/1532673X13482573.

Shanto Iyengar and Sean J. Westwood. Fear and loathing across party lines: New evidence on group polarization. American Journal of Political Science, 59(3):690–707, 2014. doi: 10.1111/ajps.12152. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12152.

Shanto Iyengar, Gaurav Sood, and Yphtach Lelkes. Affect, not ideologya social identity perspective on polarization. Public Opinion Quarterly, 76(3):405–431, 2012. doi: 10.1093/poq/nfs038. URL http://dx.doi.org/10.1093/poq/nfs038.

Janna Johnson and Morris M Kleiner. Does occupational licensing reduce interstate migration? WE Upjohn Institute for Employment Research, July, 30, 2015.

Greg Kaplan and Sam Schulhofer-Wohl. Understanding the long-run decline in interstate migration. International Economic Review, 58(1):57–94, 2017. ISSN 1468-2354. doi: 10.1111/iere.12209. URL http://dx.doi.org/10.1111/iere.12209.

Jeffrey R. Kling, Jeffrey B. Liebman, and Lawrence F. Katz. Experimental analysis of neighborhood effects. Econometrica, 75(1):83–119, 2007.

Stephen Knack and Philip Keefer. Does social capital have an economic payoff? a cross-country investigation. Quarterly Journal of Economics, pages 1251–1288, 1997.

Zovanga L. Kone, Maggie Y. Liu, Asditya Matto, Caglar Ozden, and Siddharth Sharma. Internal borders and migration in india, 2016.

Corey Lang and Shanna Pearson-Merkowitz. Partisan sorting in the united states, 1972–2012: New evidence from a dynamic analysis. Political Geography, 48: 119 – 129, 2015. ISSN 0962-6298. doi: https://doi.org/10.1016/j.polgeo.2014.09.015. URL http://www.sciencedirect.com/science/article/pii/S0962629814001000.

Bruce R. Lindsay and Francis X. McCarthy. Stafford Act Declarations, 1953–2011: Trends and Analyses, and Implications for Congress, volume 31. Congressional Research Service, 2012.

Erzo F.P. Luttmer. Group loyalty and the taste for redistribution. Journal of Political Economy, 109(3):500–528, 2001.

Andreas Madestam and David Yanagizawa-Drott. Shaping of the Nation: The Effect of Fourth of July on Political Preferences and Behavior in the United States. Working paper series, Harvard University, John F. Kennedy School of Government, August 2012.

John McCallum. National borders matter: Canada-u.s. regional trade patterns. The American Economic Review, 85(3):615–623, 1995. ISSN 00028282. URL http://www.jstor.org/stable/2118191.

Ian McDonald. Migration and sorting in the american electorate: Evidence from the 2006 cooperative congressional election study. American Politics Research, 39(3):512–533, 2011. doi: 10.1177/1532673X10396303. URL https://doi.org/10.1177/1532673X10396303.

Daniel L. Millimet and Thomas Osang. Do state borders matter for u.s. intranational trade? the role of history and internal migration. Canadian Journal of Economics, 40(1):93–126, 2007. ISSN 1540-5982. doi: 10.1111/j.1365-2966.2007.00401.x. URL http://dx.doi.org/10.1111/j.1365-2966.2007.00401.x.

Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Internal migration in the united states. Journal of Economic Perspectives, 25(3):173–96, September 2011. doi: 10.1257/jep.25.3.173. URL http://www.aeaweb.org/articles?id=10.1257/jep.25.3.173.

Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Job changing and the decline in long-distance migration in the united states. Demography, 54(2): 631–653, Apr 2017. ISSN 1533-7790. doi: 10.1007/s13524-017-0551-9. URL https://doi.org/10.1007/s13524-017-0551-9.

Nathan Nunn and Leonard Wantchekon. The Slave Trade and the Origins of Mistrust in Africa. American Economic Review, 101(7):3221–52, December 2011.

David C. Parsley and Shang-Jin Wei. Explaining the border effect: the role of exchange rate variability, shipping costs, and geography. Journal of International Economics, 55(1):87 – 105, 2001. ISSN 0022-1996.

Ricardo Perez-Truglia. Political conformity: Event–study evidence from the united states. The Review of Economics and Statistics, 100(1):14–28, 2018.

Sandra Poncet. Provincial migration dynamics in china: Borders, costs and economic motivations. Regional Science and Urban Economics, 36(3):385–398, 2006.

Serena Rhee and Fatih Karahan. Population aging, migration spillovers and the decline in interstate migration. 2015 Meeting Papers 1177, Society for Economic Dynamics, 2015. URL https://EconPapers.repec.org/RePEc:red:sed015:1177.

Joshua L. Rosenbloom and William A. Sundstrom. The decline and rise of interstate migration in the united states: Evidence from the ipums, 1850–1990. In Research in Economic History, pages 289–325. Emerald Group Publishing Limited, 2004.

Paola Sapienza, Anna Toldra, and Luigi Zingales. Understanding trust. Working Paper 13387, National Bureau of Economic Research, September 2007.

Manisha Shah and Bryce Millett Steinberg. Drought of opportunities: Contemporaneous and long term impacts of rainfall shocks on human capital. Technical Report 19140, National Bureau of Economic Research, 2013.

J.M.C. Santos Silva and Silvana Tenreyro. The log of gravity. The Review of Economics and Statistics, 88(4):641–658, 2006.

Jesse Sussell. New support for the big sort hypothesis: An assessment of partisan geographic sorting in california, 1992–2010. PS: Political Science & Politics, 46 (4):768–773, 2013. doi: 10.1017/S1049096513001042.

Guido Tabellini. The scope of cooperation: Values and incentives. 123(3):905–950, 2008.

Hideki Toya and Mark Skidmore. Do natural disasters enhance societal trust? Kyklos, 67(2):255–279, 2014. ISSN 1467-6435.

Sam Whitt and Rick K. Wilson. Public goods in the field: Katrina evacuees in houston. Southern Economic Journal, 74(2):pp. 377–387, 2007. ISSN 00384038. URL http://www.jstor.org/stable/20111973.

Holger C. Wolf. Intranational home bias in trade. Review of Economics and Statistics, 82(4):555–563, 2000.

Eiji Yamamura. Natural disasters and social capital formation: The impact of the great hanshin–awaji earthquake. Papers in Regional Science, 2014.