Social Network Effects in Financial Intermediation

Dissertation zur Erlangung des akademischen Grades
doctor rerum politicarum (Dr. rer. pol.)

Technische Universität Dortmund
Fakultät Wirtschaftswissenschaften

von
Felix Dornseifer, M.Sc.

September 2024

Acknowledgements

I would like to thank all the people who have contributed to the completion of this thesis. First of all, I would like to thank my supervisors, Christiane Hellmanzik and Oliver Rehbein. I am grateful that Christiane provided me with the opportunity and environment to continue studying the topics that Oliver motivated me to explore during my studies at the University of Bonn. I am grateful for the countless discussions, the helpful suggestions, and the trust in my abilities.

I thank all members of the RTG Regional Disparities & Economic Policy, especially Jens Wrona for creating a research environment in which I received constructive criticism, feedback, and new ideas at all times. The group has broadened my perspective and encouraged me - in a positive way - to engage with new topics and perspectives. Funding from the German Research Foundation through the RTG is greatly appreciated.

I would like to thank a group of former fellow students who continue to enjoy thinking and deliberating with me about questions from this thesis and beyond - preferably over a game of dice.

Special thanks go to my colleagues at the Chair of Urban, Regional, and International Economics. In particular, Maren Kaliske and Vanessa Hellwig discussed recurring problems with me and paved the way for me with regard to countless things. The discussions were always helpful, motivating, and a major reason why I will always remember the time during which this thesis was developed as very enjoyable. Additionally, Jens Kirsten, Hannah Bellgardt, and Hannah Hofschröer provided excellent research assistance.

Finally, I want to thank my parents and Lisa for giving me the freedom and support to pursue the paths I felt were most worthwhile, for listening to me endlessly, offering helpful advice, and being willing to support me in the best ways they could.

Felix Dornseifer Wenden, 26 September 2024

Contents

[1] Unpublished

List of Figures

List of Tables

Introduction

The transmission of information via social networks has always played a pivotal role in economic decision making, influencing behavior across markets and geographies. Social networks facilitate the flow of information, reduce frictions, and foster trust factors that are particularly important in environments where formal institutions or traditional information channels may be underdeveloped or inefficient. In recent years, the availability of detailed data on social networks has revolutionized how researchers approach the study of economic interactions and financial decisions. This thesis explores the effects of social networks across three distinct players in financial intermediation: firms, banks and insurance companies.

All three chapters of this thesis leverage Facebook's Social Connectedness Index (SCI) that measures the strength of real-world social connections. The index provides a novel and comprehensive measure of social ties between geographic regions, constructed from online friendship links. Specifically, it quantifies the relative intensity of connections between pairs of locations, both within and across countries. It maps the connectedness of regions shaped by historical, geographic, and economic factors. Given the network's global reach, with over 2.1 billion active users, the SCI offers an unprecedented and representative view of social connectedness at a large scale. The data allows to address previously intractable questions about the role of social ties in information transmission and the mitigation of economic frictions. The granularity of the data allows to evaluate network effects at multiple geographic levels, ranging from links between countries on an international scale to county-to-county connections within the United States or European countries. It opens up new possibilities to explore how social ties influence various forms of financial intermediation. This thesis sheds light on the moderating effects of social connectedness on investment decisions, mortgage financing and insurance uptake, addressing issues arising from climate risks, macroeconomic uncertainty and technological advances.

Financial intermediation, traditionally viewed as the process by which financial institutions such as banks or insurance companies match supply and demand by facilitating the flow of funds is a dynamic and continuously evolving research field. In its most fundamental form, financial intermediation reduces transaction costs, manages risks, and resolves (information) frictions arising between economic agents. These intermediaries play a crucial role in fostering economic growth and stability by enabling capital allocation and risk management. Since the Great Financial Crisis, new challenges and opportunities that change the landscape of financial intermediation have emerged. The questions covered in this thesis address issues that will become increasingly important in the future. These include the growing relevance of climate risks and macroeconomic uncertainty in foreign direct investment decisions, technological disruptions of traditional banking systems and climate risks for private households' assets. Understanding and managing these dynamics will be a critical factor for an ongoing financial integration and the transition towards a sustainable economy.

Climate-related risks, both physical and transitional, are becoming increasingly integrated into financial intermediation processes. Physical risks stem from climate-induced events such as floods, droughts, and hurricanes, while transition risks arise from the shift towards a low-carbon economy, introducing new regulatory and operational risks. In the first and third chapters of this thesis, these risks are examined within the context of FDI decisions and flood insurance markets. As natural disasters become more frequent and regulation increasingly addresses climate change, it is crucial to develop a sound understanding of how these risks impact financial intermediaries, firms, and households. All stakeholders must be prepared for these risks, as climate change will play an even greater role in shaping global investment flows and risk management strategies. The empirical findings of this thesis contribute to the understanding of these forms of uncertainty by highlighting the effects of social networks.

Next to climate risks and uncertainty, financial intermediation is subject to rapid technological change. Technical innovations with a disruptive character have already led to major changes in the traditional banking business over the last decade. As technology continues to evolve, traditional banks face increasing pressure to innovate and to adapt their business model. Disruptive technologies have a significant impact on financial intermediation, the transmission of monetary policy and the risk management of banks. In a future where technology will play an ever-increasing role, it is important to understand the consequences of this development.

The topics addressed in this thesis are not only timely but also foundational to the future of financial intermediation. Combined with the recently available data on social connectedness, this allows to provide valuable insights for policy makers.

The first chapter investigates how cross-country social networks can mitigate frictions in foreign direct investment (FDI). Facebook's Social Connectedness Index is used to quantify cross-border social ties and to assess their importance in explaining FDI flows. Investors often encounter substantial challenges related to legal, cultural, and political factors when investing in foreign markets. Information on navigating these differences is not always easily accessible. Destination countries may be subject to risks and uncertainties that are hard for external investors to evaluate, which complicates the process of making cross-border investments. This chapter presents evidence that social connectedness is associated with a significant increase in FDI, even when accounting for factors that are known to be the main determinants of FDI, such as physical distance. The results show that by bridging information asymmetries, real-world social connections can overcome a wide variety of investment frictions such as not sharing common business relationships or being culturally dissimilar. In addition, the negative consequences of climate-related risks can be attenuated. Finally, social connectedness can mitigate macroeconomic uncertainty, independent from institutional differences. This underscores the role of social ties in bridging informational and institutional gaps in international capital allocation.

The second chapter focuses on technological developments in the provision of residential mortgages and its repercussions on the availability of local bank branches. Over the last decade, financial technology ("fintech") companies have flourished while banks significantly downsized their branch networks. Although many products and services are available online, local bank branches still play an important role: They provide small business loans, payment products, investment advice and help in managing personal finances. Banks are key players in financial intermediation and their branches are crucial to support this role. This chapter shows a negative relationship between rising market shares of fintech lenders in the residential mortgage market and the number of traditional brick and mortar bank branches. The results suggest that a significant share of branch closures can be attributed to the rise of fintechs. Smaller branch networks do not only mean harder access to financial products for consumers, but also have implications for banks and the transmission of monetary policy. Not only emerging fintechs, but also increasing regulatory requirements contribute to the restructuring. To disentangle these opposing effects, I use the Social Connectedness Index as an instrumental variable. The intensity of social ties between regions is used to identify variation in local fintech adoption.

Climate risks have received a significant rise in attention. As natural disasters become more frequent, a debate has developed on how to ex-ante achieve a better asset protection through insurance. Although residential properties account for the largest share of private household assets, insurance coverage against natural hazards is relatively low. The third chapter helps to better understand the factors that lead private homeowners to insure their property against elemental damage. Using the catastrophic 2021 flood in Germany as a natural experiment, the paper demonstrates how social connectedness influences insurance uptake rates in unaffected regions. The results show that stronger social ties into the affected areas increase the likelihood of homeowners purchasing additional insurance coverage. People receive an update about their own flood risk via a social learning mechanism, triggered by first-hand experiences of those affected. The effect is moderated by regional disparities in climate policy attitudes and the intensity of different types of social capital. Peer effects in insurance uptake are less pronounced if bridging social capital is large. Bridging social capital is associated with the ability to obtain information from a wider range of sources and higher levels of trust, in the government and institutions. This most likely results in a smaller information update. Bonding social capital has a positive impact, reinforcing peer effects.

In essence, this thesis uses granular data on social networks to gain a deeper understanding of how information flows across borders and communities. It addresses various areas of financial intermediation, showing how social ties increase cross-border investments, propagate the use of new lending technologies and cause peer effects in insurance markets. The insights derived from this thesis have implications for policymakers and financial institutions seeking to leverage the power of social networks to accomplish many goals. Among many others, some of these goals could be to mitigate frictions in financial markets, to improve (international) economic integration, to retain access to financial services or to hedge risks for banks, firms and households.

CHAPTER 1

Social Connectedness and Foreign Direct Investment[1]

Abstract

This paper shows that cross-country social networks can help to mitigate frictions in foreign direct investment (FDI). First, we document an economically important relationship between real-world social networks, measured through Facebook, and FDI. We exploit common genetic ancestors as a novel instrument to establish a plausible causal link. Next, we show that social connectedness is more important if countries are legally and culturally different and if institutional frictions in the destination country are high. Social connections also mitigate investment frictions arising from climate risk. The effect of macroeconomic uncertainty on FDI is also significantly reduced by social connections.

[1] Unpublished. This chapter is joint work with Oliver Rehbein.

1.1 Introduction

How do investors overcome information frictions when deciding where to invest? Investors frequently face significant obstacles in legal, cultural and political dimensions when investing abroad. Information on how to handle legal and cultural differences is often not readily available and some investors struggle to accurately anticipate these differences in the destination country. Destination countries also face risks and uncertainties that may be difficult to anticipate for an outside investor. Such frictions frequently create issues with cross-border investments. Often, these investment frictions are also highly policy relevant. For example, the Taiwanese chip manufacturer TSMC currently faces significant problems in establishing chip factories in Arizona (Liu, 2024). And while many frictions have been shown to be important for foreign direct investment (e.g., Ly et al., 2018; Jardet et al., 2023), little attention has been paid to factors that might be able to mitigate the effects of such investment frictions.

In this paper, we present evidence that social ties can be such a mitigating factor. Using friendship links from Facebook as a proxy for real-world social connections across countries, we first show that foreign direct investment (FDI) increases in social connectedness, holding constant other known determinants of FDI. Crucially, social connectedness can significantly dampen the negative effects of investment frictions on FDI. We specifically show that social connectedness gains importance in explaining FDI flows when cultural or legal differences are large, when institutions in the destination country are weak or costly to overcome, and when climate risk and macroeconomic uncertainty is high. Overall, the results point strongly toward the fact that social networks provide a potential mechanism to overcome a wide variety of investment frictions in foreign direct investment.

To measure social connectedness, we use a data set on Facebook friendship links between countries (Bailey et al., 2018). The Social Connectedness Index (SCI) provides a snapshot of the probability that a person in country A is friends with a person in country B. Friendship links on Facebook provide a picture of social links on an unprecedented scale, with over 2 billion active users worldwide (Facebook, 2020). Kuchler et al. (2022b) demonstrate that COVID-19 spread along Facebook links, providing strong evidence that friendship links on Facebook reflect real-world social ties and do not merely reflect digital communication through the platform. We hypothesize that these real-world social connections are - directly or indirectly - leveraged in investment decisions, especially in the presence of major investment frictions. In other words, when such frictions are high, it helps to have social contacts "on the ground" which may be able to provide better information or appropriate context.

We combine this Facebook data with bilateral FDI data from the World Investment Report and add a large set of gravity and cultural variables from the Centre d'Études Prospectives et d'Informations Internationales (CEPII), the World Bank and the International Monetary Fund (IMF). We merge this data to a battery of different datasets indicating cultural and legal differences, the development of institutions, the presence of physical and transition risk stemming from climate change and state-of-the-art measures of macroeconomic uncertainty (Jurado et al., 2015; Ludvigson et al., 2021; Ahir et al., 2022; Caldara and Iacoviello, 2022).

We show that social connectedness is associated with a significant increase in FDI, even while controlling for a battery of factors that are known to be the main determinants of FDI, such as physical distance. The effect is sizable: For an increase in social connectedness by 10%, FDI increases by at least 1.2% overall and by 3.6% at the intensive margin. The bilateral panel structure of the data also allows us to include an extensive set of fixed effects that can control for the most relevant determinants and time trends of FDI in our large set of countries. Our results show, that without accounting for social connectedness, the effects of variables usually thought to matter greatly for FDI, such as distance and language, are significantly overestimated. Further, we present an analysis showing that omitted variables are unlikely to bias our results (Oster, 2019). We address further endogeneity concerns by using genetic distance as an instrumental variable (IV) for cross-country social connectedness. The results suggest a plausible causal interpretation of our findings and show that social connections can be traced back to the movement and interaction of people long before Facebook started to gather data on how individuals are linked across the globe.

We next present evidence that social connectedness is more important if countries do not share common business relationships or are culturally dissimilar. For example, social connectedness is significantly more important for foreign direct investment if trade between countries is low and if no regional trade agreements exist. It is also more important in countries that do not share a common language or legal history and are politically different. Finally, social connectedness is more relevant when the destination country does not have well-developed institutions, costly business processes or lacks developed financial markets. These sets of results suggest that social connectedness is much more important when the investment is subject to larger information asymmetries.

Climate risks have received a significant rise in attention from investors, but effects on foreign direct investment appear to be limited (Li and Gallagher, 2022; Gu and Hale, 2023). We provide evidence that social connectedness plays an important role in informing investors about climate risk in the destination country, thereby mitigating the effect of climate risk on FDI. We first show that the importance of social connectedness for FDI increases if the physical climate risk from natural disasters in the destination country is high. Interestingly, it similarly gains importance if the climate risk in the originating country is high, perhaps pointing to the fact that natural disasters affect investors (Raff et al., 2018) in the originating country or create awareness about this risk for investment elsewhere. Unexpectedly, social connectedness also helps to spur FDI when countries face high transition risk, for example when countries invest heavily into renewable energy. We find some evidence that this effect might be related to uncertainty about changes in emission related policies, but the precise channel might be an avenue for future research.

Further, we demonstrate that social connections can mitigate macroeconomic uncertainty, which is associated with declines in output in general (Baker et al., 2016) and FDI in particular (Jardet et al., 2023). As a first step, we interact social connectedness with text-based measures of uncertainty, such as the World Uncertainty Index (Ahir et al., 2022) and the Geopolitical Risk Index (Caldara and Iacoviello, 2022). While we show that social connectedness is more important if uncertainty in the destination country is high, our results suggest that social connectedness is most important in overcoming global uncertainty. With regard to global uncertainty, social connectedness increases its effect on investment by 50% when moving from very low to very high global uncertainty. The results are similar when using data-driven instead of text-based measures.

In sum, the results provide evidence that social connections provide a mechanism to mitigate and overcome a variety of frictions commonly faced by international investors. Thus, the results offer one plausible channel for the strong effects of migrant communities on trade and firm activity (Cohen et al., 2017) and for investment when information frictions are large (Burchardi et al., 2019). Our findings have implications for firms and investors, who might be able to strategically exploit social networks to their advantage. Furthermore, our findings also highlight positive effects of international travel and migration. The social footprint left behind by such cross-border movements removes frictions from international financial flows, likely leading to indirect benefits not easily apparent at first sight.

This paper relates to three separate strands of literature: the determinants of FDI, the importance of investment frictions and the role of social networks in general. Empirical research on FDI determinants has been mainly focused on either the destination country or the originating country (Biswas, 2002; Blonigen and Piger, 2014; Ly et al., 2018; Paul and Feliciano-Cestero, 2021). On the country-pair level, the most important determinants suggested by the literature are trade (Büthe and Milner, 2008; Berger et al., 2013; Kox and Rojas-Romagosa, 2020) and distance (Egger and Pfaffermayr, 2004; Alfaro and Chen, 2018). We expand this literature by highlighting the importance of social networks as a key determinant for FDI and show that social connectedness is in fact one of the most important cross-country determinants of FDI, similar to distance and trade volume between countries. Further, we show that social connectedness mitigates many investment frictions that are important for FDI, such as bad institutions (Wei, 2000), language differences (Ly et al., 2018), or the country's regulatory environment when it comes to business creation (Corcoran and Gillanders, 2015). While these frictions have been shown to affect foreign direct investment, we contribute to the literature by offering social networks as a highly relevant mitigating factor that has not received much attention in the FDI literature so far.

More recently, frictions arising from climate change risk have received increased attention.

Such frictions are likely to impact financial flows (see Giglio et al. (2021) for a review). Yet, strong effects of physical climate risks have not been found for FDI (Gu and Hale, 2023; Li and Gallagher, 2022). We contribute to this literature by showing that social connectedness mitigates risk stemming from physical climate risk, perhaps providing an explanation for why direct effects of physical risks are difficult to uncover. With respect to transition risk, it is often thought that countries with low environmental regulation are more likely to receive FDI inflows (see Cole et al. (2017) for a review). Most empirical evidence confirms the idea that environmental regulation decreases FDI inflows (Chung, 2014; Cai et al., 2016; Ni et al., 2022), while others find that more stringent environmental regulation can lead to higher capital inflows (Hanna, 2010; Fourné and Li, 2023). We contribute to this discussion by demonstrating that social connectedness is a relevant factor in the analysis: Wherever social connections are strong, they tend to mitigate frictions from environmental regulation in the destination country.

Global economic and political developments have created a heightened focus on the role of uncertainty in the decision-making of firms and other economic agents (Bloom, 2009, 2014). Increases in uncertainty negatively affect a variety of economic outcomes (Jurado et al., 2015; Novy and Taylor, 2020). Bernanke (1983) and Bloom et al. (2007) demonstrate that investment is highly sensitive to uncertainty because it is often irreversible. As a result, the investment might be delayed until the uncertainty is resolved (Julio and Yook, 2016). Evidence that uncertainty has significant negative effects on investment also exists at the firm level (Gulen and Ion, 2016). Recent advances in using text as data have generated an increasing variety of uncertainty measures (Jurado et al., 2015; Baker et al., 2016; Ahir et al., 2022) with similar results. Not surprisingly, uncertainty has thus also been shown to be an important determinant of foreign direct investment (Julio and Yook, 2016; Hsieh et al., 2019; Canh et al., 2020; Avom et al., 2020; Choi et al., 2021; Nguyen and Lee, 2021; Jardet et al., 2023).[2] Given the large importance of uncertainty for investment in general and for FDI in particular, we demonstrate that the existence of social networks can help to overcome uncertainty. Importantly, the results imply that weakening social networks might, on the other hand, amplify existing effects of uncertainty.

Social networks play an important role in many economic areas, especially when the exchange of information is important (see Kuchler and Stroebel (2021) for an overview). One strand of evidence points to the importance of networks for household financial decisions, for example for stock market participation, choice of particular stocks, and household debt (Brown et al., 2008; Georgarakos et al., 2014; Ouimet and Tate, 2020; Balakina et al., 2023). The literature also documents measurable effects of professional and political networks on investors (Fisman, 2001; Hong et al., 2005; Pool et al., 2015; Di Maggio et al., 2017), firm outcomes (Acemoglu et al., 2016), trade (Greif, 1989; Rauch, 2001; Combes et al., 2005),

[2] Cascaldi-Garcia et al. (2023) provide a summary of all measures and their numerous effects on various economic outcomes.

and growth (Burchardi and Hassan, 2013). While many of these papers rely on survey data or a relatively specific set of networks, newly available data from social networks, especially Facebook, allows measurement of social networks on a much broader scale. This type of broad social connectedness is strongly correlated with (housing market) investment (Bailey et al., 2019; Kuchler et al., 2022a) and bank lending (Rehbein and Rother, 2024). Strong evidence for the fact that data from social networks reflect real-world connections is provided by Kuchler et al. (2022b), who show that COVID-19 spread along links documented by Facebook.[3]

Most closely related to our paper is Bailey et al. (2021). The authors demonstrate that (real-world) social connectedness is positively associated with larger cross-country trade volumes, especially in the presence of large information frictions. Using the same data on social connectedness from Facebook, we add to this emerging literature in two key ways: First, we demonstrate the importance of social connections in foreign direct investment. Most importantly, we showcase that social connectedness mitigates investment frictions arising from bad institutions, physical and regulatory climate risk, and macroeconomic uncertainty.

1.2 Data and Empirical Strategy

Our baseline approach explains FDI flows from an origin country i to a destination country j by the social connectedness of these two countries. We control for a host of traditional controls such as distance and trade and several additional economic and cultural covariates. We use a panel data set covering FDI flows from 2010 to 2019 on a country-pair level.

1.2.1 Social Connectedness

To proxy for real-world social connections between countries we use Facebook's Social Connectedness Index. A full description of this data along with important correlations at the US county level can be found in Bailey et al. (2018). The Social Connectedness Index is based on friendship links on Facebook, but maps well to real-world social connections (Kuchler et al., 2022b). We argue – in line with other literature using the SCI – that connections on Facebook do not exclusively influence FDI via communication on the platform. Instead, they reflect existing friendships and social contacts that communicate via all possible channels available in the modern world. The connectedness between country i and country j is defined as follows:

$$SCI_{i,j} = \frac{FacebookConnections_{i,j}}{FacebookUsers_i \times FacebookUsers_j} \tag{1.1}$$

[3] Nevertheless literature on the economic effects of *online* social networks also exists. Field experiments (Allcott et al., 2020), the distribution of hyperlinks (Hellmanzik and Schmitz, 2017) or patent citations (Agrawal et al., 2008) have been used to study the effect on different outcome variables. Paniagua et al. (2017) describe the effect of the *use* of social media on FDI. All of these methods have the disadvantage of covering only a small fraction of global social connectedness.

The SCI can be interpreted as the relative probability of a Facebook user in country i being friends with a Facebook user in country j. To preserve confidentiality, Facebook multiplies the result of Equation 1.1 by a random factor. Accordingly, no conclusions can be drawn on the actual user data, and the SCI takes values between one and 1 billion. SCI enters the model as a time-invariant variable with data from August 2020, open to the public at the Humanitarian Data Exchange. To account for outliers in the distribution of the SCI, the data are winsorized at the 99$^{\text{th}}$ percentile. 2.74 billion monthly active users worldwide in the third quarter of 2020 (Facebook, 2020) suggests that the usage of the social network is internationally pervasive and thus provides an unprecedented picture of real-world social networks across the globe. The large scale also helps to make sure that social links are representative and are unlikely to be driven significantly by investment contacts.

To provide a picture of how social connectedness manifests across countries, we provide an overview of social links from the U.S. to other countries in Figure 1.1. The figure demonstrates that social links in the U.S. are strongest to its neighbors (Canada and Mexico) and other English-speaking countries (Ireland, the U.K., Australia, and New Zealand). The data also uncovers other relationships that might be driven by migration (Central America and Albania), ancestry (Nigeria, Kenya), and economic ties (Northern Europe). Sometimes connections arise due to historical factors, as can be seen by the above-average connection to the Philippines and to a lesser extent countries with significant U.S. military presence currently or in the past (Japan, South Korea, Singapore).

The force of history in forming social ties between countries is even more striking in Europe. Consider for example the breakup of the Austro-Hungarian Empire. Up until this day, strong real-world social networks persist between Austria and the former Eastern European parts of its Empire, while surrounding nations (Italy, Germany, Switzerland) lack these links. This pattern is clearly visible in the Facebook data, as shown in Figure 1.2. Austria shares significantly stronger social connections with countries that used to be part of the Austro-Hungarian Empire (see Figure A1.1) than it does with other European countries. For example, Austria shares much more social ties with Bosnia and Herzegovina than it does with France, despite the fact that France is economically and politically more similar.

The anecdote of Austria also provides a plausible link to foreign direct investment: It is likely due to its uniquely strong ties to the eastern parts of its former empire, that – after the Fall of the USSR – Austrians were among the first to invest in Eastern Europe. This trend is ongoing. Austria consistently invests more in absolute terms into many Eastern European countries than Germany, despite being much smaller. These investments are also often unusual and likely subject to significant information frictions. For example, the purchase of the Romanian "Banca Comerciala Romana" by Austria's "Erste Bank" in 2005 is one of the few cross-border acquisitions between relatively small banks in Europe.

Without Facebook's social connectedness data, such (historically created) links would be incredibly hard to uncover, especially in a cross-country context. In other words: While social connections between countries arise due to various circumstances, the precise ways in which they arise cannot easily be mapped across so many country-pairs. Facebook's Social Connectedness Index provides a picture of these complex historical, spatial, and economic patterns which is reflected in connections between individuals today.

1.2.2 Other Data

Inward Foreign Direct Investment: Every year, trends in FDI are examined and published in the World Investment Report by the UNCTAD. This is one of the most comprehensive databases on FDI and contains information on inflows as well as outflows. Our study concentrates on FDI inflows from 2010 to 2019. FDI is defined as an investment that gives its investor, resident in country i, control or a significant degree of influence over the management of a firm in country j (Giroud and Ivarsson, 2020). Since FDI flows per investment are presented on a net basis, reverse investment, negative retained earnings, or disinvestment can lead to negative FDI. To be able to estimate a gravity model, which depends on working with count data, we replace these negative values of FDI by zero.

Distance and Relationship Covariates: Distance covariates are obtained from the CEPII Gravity database. We include the population-weighted distance in km between countries i and j, which is the distance between the centroid of country i's population and the centroid of country j's population. They also include indicator variables on the country-pair level for a common border, a colonial relationship after 1945, a common colonizer, and the same official language.

Economic Covariates: As economic covariates we include total bilateral trade, taken from the Direction of Trade Statistics of the IMF, absolute GDP difference and the absolute difference in GDP growth rate in percentage points, both taken from World Bank Open Data. The latter controls for the economic cycle and for differences in the economic development of the origin and destination country. The more aligned the economic cycle of the two countries is, the better the investor's ability to evaluate the investment in the destination country. Total trade volume between the origin and destination country is constructed by the sum of country i's imports from country j and country j's imports from country i. Total trade is intended to control for the effect of existing business relationships. Because engaging in FDI and establishing a lasting interest in a foreign country requires more commitment than simply trading goods or services, investors might favor destinations that they already know well because of an existing trade record. Since investment decisions require a lead time, all these economic covariates enter as averaged values over the three years prior to the investment. We also control for the existence of a regional trade agreement using data from the CEPII Gravity database (also lagged by three years) and a dummy variable indicating whether a country pair shares the same currency to control for exchange rate uncertainty that may deter cross-border investments (Schiavo, 2007).

Culture and Institutions: We include a dummy for sharing a common ethnological language, spoken by at least 9% of the population and an indicator for a common legal origin from the CEPII Gravity database to control for cultural differences between countries. Religion is an important factor that influences the behavior of people and governments (La Porta et al., 1999). We use data on religious diversity from the Pew Research Center to construct a shared religion index. We also control for the proximity of the countries' political systems using data from the Varieties of Democracy Institute. Additionally, we control for the development of institutions using variables from the Worldwide Governance Indicators (WGI).

Instrumental Variable: To bolster identification we implement an instrumental variable approach using data on genetic distance from Spolaore and Wacziarg (2018). Genetic distance captures the evolutionary divergence between human populations. It builds on the time elapsed between a destination and origin country of investment since having the last common ancestors or, put differently, the time since since today's populations from two countries were the same population. Spolaore and Wacziarg (2018) combine different data sets that study microsatellite variations in human genes using DNA sequencing. They provide a data set of genetic relatedness between human populations on country level on an unprecedented scale. It captures genetic changes reflecting the historical relatedness between populations without being influenced by natural selection, providing a robust basis for inferring long-term social connectedness.

Climate Risk Measures: We rely on several data sources to address the impact of social connectedness on FDI with respect to physical climate risk. The Global Climate Risk Index annually published by Germanwatch focuses mostly on weather-related loss events. From the International Disaster Database (EM-DAT) we add information on the number of floods per country and year. We complement this disaster data with the percentage of population that is exposed to wildfires per country and year from the OECD Green Growth Indicators. Finally we include Climate Change Physical Risk as measured by Sautner et al. (2023). They use a machine learning approach to identify the physical climate risk exposure of firms based on their earnings calls. We aggregate their firm-year level data to the country-year level, based on the firms' headquarter location.
Next to physical climate risk, we use data on emissions, energy supply and regulation to assess transition risk. Therefore we add CO2-emissions in tons per capita from the Worldbank. Information about renewable energy supply, the development of environmental related policies

and energy related tax revenue is sourced from the OECD Green Growth Indicators. To address risks regarding regulation we rely on country-level year-on-year changes in the count of emission-related policies in place (Gu and Hale, 2023) and the exposure of firms to climate change regulatory risk, aggregated on the country-year level (Sautner et al., 2023).

1.2.3 Descriptive Statistics

The final data set for the baseline regression contains cross-sectional data on 157 origin countries and 114 destination countries. Overall, we cover a total of $6,552$ country pairs with $45,345$ observations over ten years from 2010 to 2019.[4] For almost 40% of the country pairs ($2,608$) the data spans all ten years of our panel. Table A1.1 in the appendix presents all data sources and provides a definition of each variable. Summary statistics are shown in Table 1.1.

The median FDI is 0.80 million USD and the distribution is highly right-skewed. Trade volume and differences in GDP and its growth rate show a large dispersion, which reflects a high variety in the economic performance of the countries across the world. Sizable differences reflect the fact that the data cover a large and diverse set of countries. The most important correlations for SCI are shown in Table A1.2 in the appendix: Social connectedness is most strongly correlated with distance (-0.65) and sharing a trade agreement (0.45).

1.2.4 Empirical Specification

For our baseline specification, we choose to estimate both a simple OLS model and a gravity model. Gravity models are often used in economics to explain various directional flows, such as the flow of goods, money, labor, and people between countries.They have also been frequently used to describe FDI flows (see for example Portes and Rey, 2005; Bergstrand and Egger, 2007; Head and Ries, 2008; Okawa and Van Wincoop, 2012). In its most basic form, a gravity model explains FDI flows by the GDP of the origin and the destination country and the distance between them. Thereby the GDP serves as a proxy for the supply and demand for cross-border investments while distance represents the frictions of such investments (i.e. transaction costs). Our estimates are specified by the following gravity equation:

$$FDI_{i,j,t} = \exp[\beta_1 \times log(SCI_{i,j}) + \beta_2 \times G_{i,j,t} + \beta_3 \times A_{i,j,t} + \delta_{i,t} + \delta_{j,t}] \times \epsilon_{i,j,t} \tag{1.2}$$

The left-hand side depicts FDI between country i and j at time t as the dependent variable. On the right-hand side, we include $log(SCI_{i,j})$ as the main explanatory variable, the vector of gravity ($G_{i,j,t}$), and the vector of economic and cultural covariates ($A_{i,j,t}$), both described in Section 1.2.2. $\delta_{i,t}$ and $\delta_{j,t}$ capture origin-time and destination-time fixed effects. These fixed

[4] Due to governmental censorship of Facebook, the SCI data is not available for China, Iran, North Korea, Tajikistan, and Turkmenistan. Additionally, some countries are not included due to a lack of data on covariates.

effects subsume all country-time varying variables which are usually deemed important for FDI inflows, such as population, economic vitality, and laws restricting inward and outward investments. Country-level differences in the use of Facebook that otherwise would have an impact on the SCI are one of the most important factors that are absorbed by these fixed effects. General differences between countries in the tendency to engage in FDI are also absorbed by these fixed effects.

We use a Pseudo Maximum Likelihood (PPML) estimator to evaluate Equation 1.2. Despite the fact that the underlying data do not follow a Poisson distribution, it yields robust results as long as the specification of the conditional mean is correct. One advantage of PPML estimation is that it is able to deal with count data that exhibits a large proportion of zeros. All continuous variables that are not scaled between zero and one enter the model logarithmized. We choose to do so because FDI and the SCI follow an approximately log-linear relationship, as plotted in Figure 1.3, and because it allows for the interpretation of coefficients as elasticities.

The gravity model allows for the correct estimation of the combined intensive and extensive margin with respect to FDI. Nevertheless, we want to specifically consider the intensive margin, for which social connectedness might be more relevant. We thus estimate an OLS model with fixed effects as displayed in Equation 1.3, using $log(FDI_{i,j,t})$ as the dependent variable. Using FDI in logs allows us to keep the interpretation of coefficients as elasticities in the OLS specification, which allows for easy comparison to the gravity specification. We employ the same control variables and also include origin-time and destination-time fixed effects and use clustered standard errors at the origin and destination country levels.

$$log(FDI_{i,j,t}) = \beta_1 \times log(SCI_{i,j}) + \beta_2 \times G_{i,j,t} + \beta_3 \times A_{i,j,t} + \delta_{i,t} + \delta_{j,t} + \epsilon_{i,j,t} \qquad (1.3)$$

1.3 Results: Social Connectedness and FDI

1.3.1 Baseline Estimation

We estimate Equations 1.2 and 1.3 and show the results in Table 1.2. Columns (1) and (2) report the estimate of the coefficient of $log(SCI_{i,j})$ without controls (but including origin-country-year and destination-country-year fixed effects). For both specifications, there is a significant correlation between social connectedness and FDI. For a 1% increase in social connectedness, FDI increases between 0.51% and 1.1% depending on the specification. A significant share of the variation in inward FDI is explained by fixed effects and SCI alone. We suggest that the larger elasticities in OLS may be explained by the fact that social connectedness might be more important at the intensive margin because the OLS estimation only includes non-zero country pairs. Supporting evidence for this is provided by Table A1.3, in which we use $log(1 + FDI_{i,j,t})$ as the dependent variable. In this case, the OLS coefficients are very similar to the PPML baseline estimation.

Columns (3) and (4) introduce our large set of country-pair level control variables, which include distance, trade, differences in economic development, and cultural factors. Including these controls reduces the size of the effect of the SCI to 0.12 and 0.36, but it remains statistically significant. Again, the difference between the two models suggests that social connections may be more important at the intensive margin. To put the size of this effect into perspective, consider the example of Austria, which has social ties to the members of its former Empire (see Section 1.2.1 for details). Austria's social connectedness to Romania is about 30% larger than its connectedness to France (despite the similar distance). As a result, our estimates imply that Austrian FDI to Romania will be about 4-10% larger than investment to France, simply due to this heterogeneity in real-world social ties, which are otherwise hard, if not impossible to measure.

Interestingly, distance, one of the most important control variables in predicting FDI is not statistically significant when including social connectedness as a control in the gravity specification. It is however highly significant without controlling for social connectedness (see Table A1.4). While it is very time- and cost-intensive to ship goods or to provide services over large distances, this does not necessarily apply to FDI. Intuitively, FDI is less reliant on transportation routes but more reliant on information. Alfaro and Chen (2018) suggest that the elasticity of FDI with respect to distance corresponds to half of the elasticity of trade with respect to distance. Controlling for the flow of information through social networks (and business networks via trade) might thus plausibly eliminate the effect of distance on FDI entirely.

Consistent with the literature, most other covariates enter with the expected sign. By far the most important economic covariate is the logarithm of trade volume showing an elasticity of 0.49 and 0.37, respectively. Differences in GDP between the origin and destination countries appear to be more important than differences in growth. Significant coefficients for the shared religion index and the political distance also suggest that cultural and political proximity ease FDI flows. A common border does not seem to be a promoting factor for investment in our setting.

One plausible interpretation of the effect of SCI on FDI is that social connections are a good proxy for existing real-world (potential) investment relationships that are otherwise almost impossible to measure and arise from many factors that are quite different in nature. Another plausible interpretation is that social connections are actively leveraged in the course of evaluating an FDI opportunity. For example, a U.S. FDI project in Mexico might be significantly easier to evaluate if a person in the decision-making team has (social) roots in Mexico, because that person may be able to acquire local knowledge that is otherwise very difficult to obtain (e.g. quality of the workforce, security, local culture).

To further highlight the point that social connectedness is one of the most important factors in explaining FDI, Tables A1.4 and A1.5 investigate the importance of each control variable separately. First, it is necessary to highlight the importance of fixed effects, which explain a significant share of the variation. We can get an impression of the importance of individual variables by comparing the contribution to the R^2, i.e. the explanatory power that the variable adds to the model beyond fixed effects. In PPML estimation (Table A1.4), SCI enters with the second largest explanatory power of all variables (added R^2 of 0.035), right behind trade volume (0.042), but ahead of distance (0.031). In OLS (Table A1.5) SCI is the most important predictor (added R^2 of 0.08) ahead of distance (0.079) and trade (0.074).

1.3.2 Omitted Variable Bias

While we do our best to control for the known determinants of FDI, it remains possible that the analysis suffers from an omitted variable bias. To mute this concern, we follow the procedure suggested in Oster (2019). The main idea is the following: If the coefficient of interest remains sufficiently stable when adding (meaningful for the model-fit) control variables, the coefficient is unlikely to be driven by omitted variables alone. To test this, Oster (2019) derives a simple formula based on regression coefficients and R^2 values:

$$\beta^* \approx \tilde{\beta} - \delta[\mathring{\beta} - \tilde{\beta}]\frac{R_{max} - \tilde{R}}{\tilde{R} - \mathring{R}} \tag{1.4}$$

$\mathring{\beta}$ and $\mathring{R}$ are the coefficient and R^2 from the uncontrolled regression (see Column (1) in Table A1.5) and $\tilde{\beta}$ and $\tilde{R}$ are coefficient and R^2 from the controlled regression (Column (4) of Table 1.2). We use the OLS estimations for this comparison, because the procedure suggested by Oster (2019) is derived for OLS models, but the logic (and the outcomes) extend to the PPML results. δ is set by assumption and represents the importance of the unobservables relative to the control variables. For example, a δ of 0.5 assumes that potential unobserved variables (if they could be observed) would be half as important as the current collection of control variables. Even though we have controlled for most known determinants of FDI, we take a very conservative assumption and set $\delta = 1$. In other words, we assume that any potentially omitted variables would explain as much variation as the collection of known FDI determinants. R_{max} is the maximum R^2 that can be explained by any variables. We assume $R_{max} = 1$.

Under these conservative assumptions, the bias-corrected elasticity of social connectedness is still 0.21. This remains an economically important effect. As suggested by Oster (2019), the intuition is even more important than the precise value: While our control variables (especially the country fixed effects) add significant explanatory power to the model, they do not do much to move the social connectedness effect towards zero.

Overall, these results indicate that social connectedness is significantly positively correlated with the FDI flows between countries. Despite controlling for a large set of fixed effects and a battery of economic and cultural covariates, the elasticity of FDI with respect to the SCI is 0.12 in the gravity specification. That means a doubling of social ties between country i and j is associated with an increase in FDI by 12%. In light of the fact that social connections between countries have not been investigated as a determinant of cross-border investments so far, this effect is large and speaks to the importance of social connectedness in gathering information about potential foreign direct investments.

1.3.3 Reverse Causality: Instrumental Variables

While we include a battery of known determinants of FDI to minimize omitted variable bias, there is also a potential reverse causality concern. The SCI data measures the probability of cross-border Facebook friendships in August 2020 and is used to explain FDI flows from 2010 to 2019. It is possible that cross-border friendships were established because an investment had been made. First, we argue that it is unlikely that the small share of individuals engaged in FDI decisions are driving the aggregate relationship between countries – which most often involve millions of individuals – to a statistically relevant degree. Furthermore, FDI involves fewer agents than trade and Bailey et al. (2021) provide comprehensive evidence that reverse causality is not a plausible concern in explaining trade flows.

Nevertheless, we complement this qualitative discussion with an instrumental variable approach. We implement the instrument using a control function approach for PPML (see Section A1.1 for details) and standard 2SLS for OLS. Finding an instrument for social connections is difficult because there is a broad range of reasons for cross-border friendships on Facebook (see Section 1.2.1). Many of these friendships can be traced back to single historical events or factors that are unique for single countries or country pairs. For example, Bailey et al. (2021) report a high social connectedness between Australia and Denmark, stemming from the marriage of the Danish crown prince to his Australian-born wife. The challenge is to find an instrument that influences all countries to a sufficient degree and is therefore valid for the whole data set.

We suggest genetic distance from Spolaore and Wacziarg (2018) as an instrument for the SCI. Genetic distance is a measure that builds on the time elapsed between a destination and origin country of investment having last common ancestors. Common ancestors from long-past time periods cannot be influenced by investment today, thus ruling out reverse causality. The literature suggests that sharing a common ancestor may lower socio-cultural barriers and consistently finds stronger economic co-movements in countries with more recent common ancestors (Spolaore and Wacziarg, 2013, 2018). We suggest that common ancestors facilitate this process because social relationships are easier to form. Based on these arguments, we argue that genetic ancestry is likely a good instrument for the effect of social connections on FDI. Due to its nature, it helps to rule out reverse causality in particular.

Columns (5) and (6) of Table 1.2 report the results using genetic distance as an instrument. The results of the first stage regressions can be found in Table A1.6. The instrument exhibits a very strong first stage (F-values > 120), which confirms the assumption that the relevance condition holds. Using genetic distance as an instrument reveals elasticities of SCI on FDI of 0.53 and 0.46. Both the PPML and the OLS IV approaches are pointing in the same direction as the baseline estimates including controls.

The results from our instrumental variable regressions suggest that there might be a causal effect of social connectedness on FDI. We argue that the instrument is plausibly exogenous, but nevertheless remain cautious in using causal interpretations throughout the paper. We note that even in the absence of causality, the relationship between social connectedness and FDI is novel and little explored, and is unlikely to be easily explained by simple omitted variable concerns.

1.3.4 Robustness and Placebo

A growing source of uncertainty in FDI data is the increasing use of tax routing and special-purpose-enterprises (SPEs), e.g. letterbox companies causing "phantom FDI". In those cases, the origin or destination of the FDI is not the ultimate designation which disturbs our analysis. Damgaard et al. (2024) use data from the IMF's Coordinated Direct Investment Survey to construct a global network of total FDI. This data set represents an alternative source for FDI flows to the UNCTAD data we use. While the IMF data has an initially smaller coverage, Damgaard et al. (2024) enrich the data leveraging mirror reports from counterpart economies. This leads to a data base covering 14,738 country pairs. We reconstruct our baseline results using this data set. The results (Table A1.7) are very close to our baseline estimates.

In a second step, we exploit the fact that Damgaard et al. (2024) decompose FDI into investments in non-SPEs ("real FDI") and SPEs ("phantom FDI"). This allows us to directly exclude "phantom FDI". Columns (1) and (2) of Table A1.8 confirm the baseline findings for FDI in non-SPEs, i.e. for "real FDI". The elasticity of SCI yields 0.30, documenting a higher effect compared to the baseline. Furthermore, the data also allows us to explicitly investigate the effect of social connectedness on "phantom FDI". such "phantom FDI" flows should not be affected by social connectedness between the originating country and the SPE country, because SPEs are simply a vehicle to route an investment into a different ultimate destination. Thus "phantom FDI" represents a suitable placebo test. As expected, the coefficient in Column (3) is insignificant and close to zero. Using genetic distance as an instrument in Column (4), the result is also statistically insignificant. Lastly, the coefficients are large and highly statistically significant for the case of "real FDI" by ultimate investor economy. In this case, we exclude the routing via the SPE to reconstruct the "true" FDI flows even when an SPE is used. For example, consider an investment from the US to Germany, passed through a holding firm (letterbox company) in Luxembourg. Correcting for the ultimate owner economy of the investment means that this investment is now is taken into account as an investment from the US to Germany, instead of an investment from Luxembourg to Germany. As expected, using this correction yields statistically significant results in line with the sample from Columns (1) and (2), where we simply remove FDI from SPEs.

Our results do not rely on the choice of a panel data set or the recording of FDI flows. We show this by estimating Equations 1.2 and 1.3 using cross-sectional data for the year 2020. Table A1.9 presents the results, which are comparable to our baseline findings. Countries have a larger incentive to monitor inward investment flows compared to outward flows. However, the World Investment Report does also provide outward FDI flows. Table A1.10 shows the results of using outward flows from 2010 to 2019. The results are again comparable to the baseline elasticities. To show that the results do not entirely capture the effect of existing FDI relationships, Table A1.11 controls for FDI flows from the previous years. The effect of social connectedness remains stable and statistically significant.

Moreover, we use additional data sets to contextualize and compare the impact of the SCI on investment decisions. First, we use the Coordinated Direct Investment Survey of the IMF (CDIS) to estimate the influence of the SCI on FDI stock data. Table A1.12 shows a coefficient of 0.29, significant on a 1% level, which is larger than in the baseline regression. It is the largest significant variable in this approach. Next, we use data of the Coordinated Portfolio Investment Survey of the IMF (CPIS) for 2019. Similar to above the CPIS represents stock data and will most likely date back several years. Next to that, portfolio investment is very different from FDI. It requires much less due diligence and commitment. Not only can it be smaller in size but it also does not necessarily fulfill the criteria of establishing lasting interest and managerial control. While for FDI personal visits are often unavoidable, portfolio investment is often made without the investor being on site. We thus expect the coefficient to be smaller and less significant. The results in Columns (4)-(6) confirm this expectation. The elasticity of the investment regarding the SCI yields 0.17, which is larger than the baseline, but the effect is less precisely estimated and is only significant at the 10%-level.

1.4 Social Connectedness and Information Frictions

One plausible interpretation of the effect of social connectedness on FDI is that social connections enable investors to gather more information about a destination country, or make the acquisition of information more easy. To provide supporting evidence for this hypothesis, this section investigates whether social connectedness is more relevant for FDI in cases where the need for information is increased. We start by investigating common frictions in international finance: The lack of existing trade relationships, the lack of institutions in the destination country and other country-pair differences that are likely to impair the flow of information. We then also test if social connectedness becomes more relevant when information frictions from climate change risk are present. Finally, we investigate the role of social connectedness when macroeconomic uncertainty is high.

1.4.1 Trade, Culture and Institutions

Trade: One reason that investment contracts between parties might be subject to fewer information asymmetries are existing business relationships. Between countries, one way to proxy for these relationships are bilateral trade flows between countries. We thus first interact SCI with the log of the trade volume between country-pairs. The results, presented in Table 1.3, Column (1) show that the effect of SCI for country pairs at the 75$^{\text{th}}$ percentile is slightly smaller than the baseline result. However, the estimated elasticity of 0.12 almost doubles for observations in the 25$^{\text{th}}$ percentile. Put differently, the effect of social connectedness significantly decreases in the strength of business ties. We complement the analysis by interacting the SCI with the dummy for a regional trade agreement, to investigate if barriers to trade might pose information asymmetries that can be overcome by social contacts. We again find significant effects. The estimated elasticity is 0.28 without a trade agreement in place and plummets to 0.01 when such an agreement exists. Both specifications demonstrate that in the presence of information frictions posed by regulation and prior business relationships, social connectedness plays an out-sized role in foreign direct investment.

Cultural differences: Another variant of information frictions may arise from cultural differences between the contracting parties: If similarities are small, information frictions may be significantly increased. We test this hypothesis by interacting SCI with three cultural covariates: a common legal origin, a common official language, and political distance. Column (3) demonstrates that a common legal origin more than halves the importance of social connections for FDI decisions. The variable is taken from the CEPII data and indicates if a country pair can look back on a similar legal origin. In that case, information might flow more easily, requiring a smaller role of social connectedness. Similar findings apply to sharing a common language (Column (4)). In this case, the effect of social connectedness on FDI vanishes entirely where countries share a language. When investing firms are unable to parse the (legal) environment in the destination country because they are unable to properly understand it due to language barriers, relationships to the destination country appear to be quite important.

Political differences: Similarly, information frictions might arise from differences in political systems. We thus merge a continuous indicator of political distance gathered from the Varieties of Democracy Institute. The indicator sorts countries into autocratic vs. democratic systems, with 0 indicating a closed autocracy and 9 indicating liberal democracies. We then divide these numbers by 9 and construct the squared distance between the originating and destination countries' score. We interact the resulting political distance indicator with social connectedness in Column (5). The outcome confirms the idea that this type of uncertainty plays an important role: The farther apart two countries are in terms of their political system, the more important social connections become. For politically very similar countries the effect of SCI on FDI is still 0.07 (25^{th} percentile), but it doubles to 0.18 for politically dissimilar country pairs (75^{th} percentile).

Institutions: We next draw on data from the Worldwide Governance Indicators (WGI). This index collects data on institutional quality for a large set of countries over our sample period and groups them into six different dimensions: Control of Corruption, Rule of Law, Voice and Accountability, Political Stability and Lack of Violence, Government Effectiveness and Regulatory Quality. Since these indicators are closely related, we only present results of the interaction with Regulatory Quality and Government Effectiveness in Table 1.3. The full set of interactions with all WGI indicators are presented in Table A1.13. Column (6) of Table 1.3 shows that institutions also appear to be an investment friction that can be bridged by social connections. If a destination country's regulatory quality is at the global median, a 1% increase in social connectedness is associated with a 0.13% increase in FDI in the destination country. However, when the regulatory quality is high (75^{th} percentile), this association decreases to almost zero. When, on the other hand, the destination country's regulatory quality is particularly low (25^{th} percentile), a 1% increase in social connectedness is associated with a 0.27% increase in FDI. The numbers are very similar when considering Government Effectiveness as an institution (Column (7)), or when looking at other institutional factors (Table A1.14). Overall, these interactions are consistent with the idea that businesses directly or indirectly exploit existing social relationships to different countries to overcome barriers posed by regulatory frictions in the destination country.

Ease of doing business: The CEPII Gravity Data include information on how many procedures and days are required in every country to start a business and how costly this process is. The literature shows that in countries where starting a business is faster and less costly, FDI increases (Jayasuriya, 2011; Corcoran and Gillanders, 2015). Faster and cheaper procedures may be a sign of a good contract and property rights enforcement and may thus present fewer frictions that require social connections to overcome. We thus interact the Social Connectedness Index with the ease of doing business variables.

Table 1.4 presents the results. Column (1) includes interactions with the time needed to start a business, Column (2) with the procedures required to start a business, and Column (3) with the associated costs of doing business. Since all countries require time and procedures greater than zero and almost all of them exhibit costs, we display the effect of social connectedness at different values of the indices in the destination countries at the bottom of Table 1.4. For example, in the median destination country, in terms of days to start a business, the elasticity of SCI with respect to FDI is 0.08. However, social connectedness is more important at the 75^{th} percentile, where the elasticity is 0.12 (Column (1)). The estimates for the number of procedures required to start a business and the costs are similar (Columns (2) and (3)). Interestingly, social connections seem to matter most strongly for the case of procedures which likely involve contact with somebody in the destination country. The important finding of this approach is that all three interaction terms of the SCI with the variables referring to the destination countries' ease of doing business are large in size and significant on a 1% level. This approach shows that the effect of social connectedness is highly sensitive to frictions arising when establishing a new business in the destination country. The results suggest that social relationships are able to reduce barriers by facilitating the process of investing abroad, especially when the hurdles to do so are large in terms of time, procedures, and costs.

Financial market development: Existing evidence suggests that financial development has a significant impact on attracting FDI (Desbordes and Wei, 2017). We thus interact SCI with a variable for financial development in the origin and destination country, taken from the IMF Financial Development Index. This index evaluates the financial market development by measuring its depth, access, and efficiency, combining a large number of variables and indicators (e.g. stock market capitalization to GDP, stocks traded to GDP or stock market turnover ratio). Financial development of a country proxies for the quality, accessibility, and trustworthiness of hard financial data. We hypothesize that better-developed financial markets reduce investment frictions. Consequently, we expect that social connectedness will play a smaller role in countries with well-developed financial markets.

Column (4) of Table 1.4 presents the results. The coefficient of SCI without taking the financial market development into account is large and statistically significant, yet its importance significantly decreases in financial market development. At the median financial development in the destination country, the elasticity of social connectedness w.r.t. FDI is 0.18, slightly higher than the baseline effect. Yet, the effect decreases to 0.05 when the destination country's financial market is at the 75^{th} percentile of development, and increases to 0.31 when financial development is only at the 25^{th} percentile. Simply put, better-developed financial markets decrease the need for social connectedness in foreign direct investment.

Altogether, these results point to the relevance of social connections in overcoming various different types of investment frictions in foreign direct investment. Whenever two countries are culturally or legally dissimilar, social connectedness becomes an important factor in exploiting FDI opportunities. The same holds when institutions in the destination country are not well developed. This logic appears to apply to both business-related and financing-related frictions.

1.4.2 Climate Risk: Physical

The importance of climate change risks has grown significantly due to the increasing prominence and salience of climate change. Yet, evidence of a potential effect on FDI has been hard to come by (Gu and Hale, 2023). One plausible explanation for this finding is that firms are able to overcome information frictions stemming from climate risk, thus avoiding some of the negative consequences. In this section we test if social connectedness may play a role in overcoming such investment frictions stemming from physical climate risk.

While Addoum et al. (2020) do not find any direct effect of temperature changes on firm sales or productivity, there are mounting concerns that physical risk can play a role in firms' investment decisions (Ginglinger and Moreau, 2023). We hypothesize that social connectedness between countries might be relevant in overcoming investment frictions from physical risk. We begin by interacting social connectedness with the Climate Risk Index from Germanwatch (CRI). The Climate Risk Index scores countries according to their relative exposure to extreme weather events such as floods and storms (Eckstein et al., 2021). The CRI measures covers a broad range of countries (180) and is available yearly over our entire sample from 2010-2019. It combines the number of deaths, the number of deaths relative to the population, the losses in USD (PPP), and the losses per unit of GDP. It gives larger weight to the relative measures in calculating the overall index. As it uses a combination of factors, the climate risk index provides a reasonable aggregated measure of physical climate risk on the country level. The CRI is coded as a ranking, meaning that larger values (lower ranks) mean lower climate change risk. We interact the CRI with social connectedness to investigate if social connectedness can help to overcome frictions stemming from physical climate risk.

Column (1) of Table 1.5 displays the results. As in our baseline regressions, social connectedness is highly relevant for FDI. The interaction between social connectedness and the Climate Risk Index of the destination country is significant and negative. This result implies that the importance of social connectedness for FDI increases as physical climate change risk increases (lower CRI scores). For more intuitive interpretation, we provide the elasticities of social connectedness on FDI at the bottom of the table. At median physical climate change risk, the effect of social connectedness on FDI is very similar to our baseline estimate: A 10% increase in social connectedness is associated with an increase in FDI by 0.9%. Yet, as physical climate risk increases (the CRI rank lowers), the effect of social connectedness increases to 0.19 at the 25^{th} percentile of the CRI index.

To confirm that our results are not driven by a quirk in the CRI data, we also gather country-level information on natural disasters from EM-DAT. This data base is operated by the Centre for Research on the Epidemiology of Disasters and compiles data on over 26,000 global disasters from 1900 to the present at the country level. It sources information from UN agencies, NGOs, insurance companies, research institutes, and media outlets. Since the data contains both natural hazards as well as technological disasters, we subsequently focus on the number of floods to capture climate risk. We interact the SCI with the number of floods at the country level. The results are displayed in Column (2) and are consistent with an increasing importance of social connectedness when physical climate risk is high: When the number of floods in both the destination and originating country increases, so does the importance of social connectedness for FDI. For the median disaster-affected country, the effect of social connectedness on FDI is 0.74% for a 10% increase in social connectedness. For countries at the 75$^{\text{th}}$ percentile of disaster prevalence, this increases to 1.28%. A similar finding holds when we simply use a dummy indicator of whether a country has experienced a flood in a given year or not (Column (3)). The effect of social connectedness on FDI is not significant in non-flooded countries, but roughly doubles in importance for countries which have experienced a flood.

We also investigate whether this finding extends to disasters other than floods. Using data on wildfires from the Green Growth Indicators by the OECD, we check whether countries with a larger share of the population exposed to wildfires respond similarly. Column (4) presents the results. Indeed, the interaction is significant at the 5% level, indicating that social connectedness is more important for FDI if natural disasters pose a more significant risk. Finally, we investigate investors' opinions directly. Sautner et al. (2023) provides a physical climate risk measure based on textual analysis of firms' quarterly earnings reports. If investors voice concerns about the physical risks of firms' investment projects abroad in earnings calls, we might expect a larger role of social connectedness. We aggregate this firm level measure on the country-year level and find some evidence in this direction, although the coefficient is not statistically significant (Column (5)) for the destination country.

It is notable that the effect of social connectedness on FDI increases with the physical risk in the originating country for most measures. There are two plausible channels for this. First, the effects of natural disasters in the originating country freeze-up funds that are now needed elsewhere, and investors withdraw first from high-connectedness countries. Second, it might be a salience effect: A disaster in the originating country might lead to increased questions about physical risk in the destination country, for which social connections can be leveraged.

One salient concern with these results is that the risk from physical disasters might simply proxy for other institutions discussed in the previous section. Countries that may have good institutions may be better prepared to deal with oncoming natural disasters, thus reducing the frictions they create. To address this concern, we include interactions of the SCI with the Worldwide Governance Indicators (WGI) as a further control in Table A1.15. The results show that controlling for institutions does not significantly alter the finding that social connectedness is particularly strongly associated with FDI when physical climate change risk is high. In fact, when controlling for institutions, social connectedness becomes statistically more relevant when physical risk is measured through earnings calls (Sautner et al., 2023).

1.4.3 Climate Risk: Transition

Countries' exposure to climate change risk does not only stem from exposure to physical risk, but also from the risks associated with the transition to a carbon free economy. As countries have committed to reducing emissions in various international climate change agreements, e.g. in Paris in 2015, firms and industries emitting large amounts of CO2 are at risk of having to pay for their emissions through taxes, emission certificates or other direct regulatory interventions. Further, industries might also be exposed to this risk through changes in customers' preferences or changing supply chains more generally. Notably, firms may also be affected indirectly by transition risk if other firms in their supply chain are subject to additional costs from the green transition. On the other hand, one might argue that not transitioning poses similar or larger frictions related to future economic development. As a result, the role of social connectedness in this context is not clear ex-ante. It is also important to note that while physical risk is typically higher in developing countries, transition risk is higher in developed countries with large emissions. We demonstrate that this is the case in Figure A1.2, where we plot yearly averages of the number of floods (Panel A) and per capita CO2 emissions (Panel B).

To investigate the role of social connectedness in the context of transition risk, we start by interacting the social connectedness measure with the logarithmized CO2 emissions in the originating and destination countries. The results are displayed in Column (1) of Table 1.6. They indicate that the importance of social connectedness for FDI wanes in destination countries which emit more CO2. We confirm that the coefficient of this interaction is not easily explained by the inclusion of interactions with institutional indicators (Table A1.16) or with an extensive set of interactions with uni- and bilateral control variables (Table A1.17). In other words, it is unlikely that the interaction with carbon emissions is simply a proxy for other institutional frictions or other cultural and economic differences.

We find similar results when we look directly at the renewable energy supply available in the destination country (Column (2) of Table 1.6). The importance of social connectedness increases significantly, when the supply of renewable energy supply is large. The elasticity quadruples when moving from the median to the 75^{th} percentile of the distribution of renewable energy supply. Again, this is unlikely to be driven by being a rough proxy for other macroeconomic indicators (Tables A1.16 and A1.17). These results might serve as one indication that renewable energy supply in particular might be a cause of uncertainty or information frictions that make investors hesitant to act, with social connectedness helping to overcome such frictions.

We next investigate if technological advancement related to the green transition poses a similar friction that is bridged by social connectedness and interact the SCI with a country-level indicator describing the relative share of green technologies being developed in the originating and destination countries. The results in Column (3) do not indicate that social connectedness is relevant in overcoming such frictions. Yet, when we control for institutions and further factors, we do indeed find that social connectedness is relevant (Column (3) of Tables A1.16 and A1.17).

Another friction may be related to economic policies around the green transition. Countries moving away from fossil fuels might be subject to significant regulatory changes now and in the future, potentially raising the uncertainty around investment opportunities. To investigate whether social connectedness is relevant in overcoming such investment frictions, we start by interacting social connectedness with the number of changes in emission related policies from Gu and Hale (2023) and existing energy and carbon taxation to proxy for regulatory intensity. We do not find statistically significant evidence that the change in emission related policy increases the relevance of social connectedness (Column (4)) and also the volume of tax revenues raised appears not to do so (Column (5)). Both variables do not indicate policy changes in the past or today, yet future (expected) policy changes might be very relevant for FDI as long as it is not controlled for the institutional quality. This result transforms in case of changes in emission related policies. When controlling for institutional quality (Table A1.16), the elasticity of the SCI matches our baseline magnitude for countries at the 75$^{\text{th}}$ percentile of the distribution of changes in emission related policies.

To complement the analysis, we draw on a measure from Sautner et al. (2023), who measure regulatory risk for firms through text in firms' earnings calls. We aggregate this firm-level measure on the country-year level and interact it with the Social Connectedness Index. The interaction is positive and statistically significant (Column (6)). Significance increases once more controls are added (Column (6) of Tables A1.16 and A1.17). The fact that social connectedness becomes more relevant for FDI when regulatory risk in (listed) firms' earnings calls is more frequent would be consistent with the idea that multinational firms might leverage social networks in making FDI decisions when faced by transition risk, although more evidence would be needed to pin down a precise mechanism.

1.4.4 Macroeconomic Uncertainty

Global events over the last two decades have highlighted the importance of macroeconomic uncertainty for investment and growth (see Cascaldi-Garcia et al. (2023) for an overview). At the same time measurement of uncertainty has significantly widened and improved, mostly based on textual analysis of various kinds (Jurado et al., 2015; Segal et al., 2015; Baker et al., 2016; Bekaert et al., 2022). Uncertainty is likely to have a strong negative effect on foreign direct investment since investors are likely to withdraw to their home markets when uncertainty is high (Jardet et al., 2023). In this section, we ask the simple question of whether social connectedness can help to overcome this investment barrier created by increasing uncertainty. We investigate it by interacting SCI in Equation 1.3 with various measures of uncertainty. If social connectedness is more important in times of high uncertainty, we would consequently expect positive interaction coefficients. Note that the direct effects of uncertainty on FDI are taken out by country-time fixed effects in all regressions.

World Uncertainty Index: We begin by interacting social connectedness with the World Uncertainty Index (WUI), which has been shown to measure uncertainty that negatively affects FDI (Jardet et al., 2023). The index reflects the relative frequencies of the word "uncertainty" in EIU country reports in each country (Ahir et al., 2022). Column (1) of Table 1.7 demonstrates the familiar importance of social connectedness for FDI. Social connectedness becomes more important as the destination country's uncertainty increases. A one standard deviation increase in the WUI increases the effect of social ties on FDI by about 13% of the baseline, statistically significant at the 10% level. The importance of social connections also increases with uncertainty in the originating country, but the interaction is not statistically significant. This finding is in line with literature highlighting the importance of uncertainty in the destination country (Azzimonti, 2019), although the effect is not large and only marginally significant. More important than country-specific uncertainty is global uncertainty. The idea that global uncertainty plays an out-sized role for foreign direct investment similarly arises in Jardet et al. (2023). Column (2) demonstrates that social connections matter significantly more when global uncertainty is high. A one standard deviation increase in global uncertainty is connected to an increase in the importance of SCI for FDI by roughly 38% of the baseline effect.

To illustrate the economic importance of social connections on FDI, we consider the difference between periods of low and high uncertainty. Global uncertainty was lowest in 2010. In this low-uncertainty year, the elasticity of FDI with respect to the SCI amounts to 0.13, slightly lower than our baseline estimate. Compared to that, social connectedness is much more important with respect to FDI when uncertainty is large. This is the case for 2016, the year in which the UK voted for Brexit and the US for Donald Trump. Under this heightened uncertainty, the elasticity increases to 0.42.

One might be concerned that the global WUI is driven by a number of small countries which contribute disproportionately to global uncertainty. Consequently, Column (3) repeats the interaction using the GDP-weighted global uncertainty index with similar findings. In this case, the elasticity of social connectedness in low uncertainty 2010 is 0.22 but increases to 0.33 in high uncertainty 2016.

Geopolitical Risk Index: One potential channel through which global uncertainty can affect investment is geopolitical risk. Caldara and Iacoviello (2022) produce an index of geopolitical risk and show that higher values of their GPR measure predict the probability of economic disaster, lower investment, and employment. We consequently investigate whether social connections can help to mitigate investment frictions arising from geopolitical risk. Column (4) of Table 1.7 provides evidence that this is indeed the case. We interact the Geopolitical Risk Index (in logs) with the social connectedness between countries and find that in times of high geopolitical risk, social connections can mitigate the negative effect on FDI.

Non-text-based Uncertainty: In addition to text-based measures of uncertainty, other measures of time-varying uncertainty have been developed. Prominently, Jurado et al. (2015) argue for a prediction-based uncertainty measure that is available for a large set of countries. Their measure is based on the idea that uncertainty is high when economic predictability (or forecast-ability) is low. Ludvigson et al. (2021) build on this concept and further develop a measure of uncertainty that is generated only from real-economic variables such as output or unemployment. Both measures are less reliant on the perception of uncertainty and focus instead on realized uncertainty. We interact social connectedness with these macroeconomic uncertainty indices in Columns (5) and (6) of Table 1.7. Similar to other measures of uncertainty, social connectedness becomes significantly more important in predicting FDI in times of high uncertainty. The elasticity of SCI at the median of the Macroeconomic Uncertainty Index is 0.34, and it increases to 0.39 at the 75$^{\text{th}}$ percentile. At the median of the Real Uncertainty Index, the elasticity of SCI on FDI is about 0.32, while it increases to 0.39 at the 75$^{\text{th}}$ percentile of the distribution (Column (6)).

1.5 Conclusion

This paper shows that real-world social networks, proxied for by Facebook's Social Connectedness Index, are an important determinant of foreign direct investment. We employ both gravity and OLS estimations with a large set of distance, economic, and cultural covariates, in conjunction with restrictive fixed effects using a panel data set of FDI flows by the UNCTAD. The main result shows a statistically and economically significant effect of social connections on FDI. A 1% higher social connectedness is associated with an increase in FDI by 0.12%. This effect often statistically eliminates previously important determinants such as the distance between countries. To improve identification we conduct a test to rule out omitted variable bias, use data on genetic ancestry across the globe as an instrument and show that social connectedness has no effect for "phantom FDI". Combined, the results suggests a potential causal effect of social connections on FDI. Our results are in line with an emerging literature highlighting the economic effects of social relationships and suggest that social connections are a highly important determinant for foreign direct investment.

We emphasize the importance of social connectedness particularly in the presence of several frictions that often arise in the context of investment. First, we show that the importance of social connectedness is larger when cultural and institutional frictions persist. If country pairs don't trade frequently, if they have a different language and different legal origins, social connectedness is more important. Similarly, social connectedness helps to overcome the lack of good institutions and government in the destination country. Second, we show that social connectedness is more important if climate risks are high: In cases where both originating and destination countries are subject to higher physical risk from natural disasters, the importance of social connectedness for FDI is higher. Interestingly, this is also the case when countries transition toward a green economy. We hypothesize that when indicators of transition are present, social connections may help in overcoming uncertainty about the effects of transitioning towards a green economy, for example by helping to pin down the effects of transition-related regulation. Finally, measuring macroeconomic uncertainty through state-of-the-art text-based and prediction-based methods, we show that social connections are more important if uncertainty is high.

Our findings are not only important to better understand the cross-country determinants of FDI but may also be valuable for firms and governments attempting to promote investment in general and FDI in particular. Firms, for example, might strategically exploit social networks when making investment decisions. Our results suggest that there might be significant economic value in promoting cross-country exchange, especially when information frictions and rising global uncertainty prove challenging. Our results also open a new avenue for research trying to find factors mitigating the effects of climate risks and global uncertainty.

Bibliography

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., and Mitton, T. (2016). The value of connections in turbulent times: Evidence from the United States. Journal of Financial Economics, 121(2): 368–391.

Addoum, J. M., Ng, D. T., and Ortiz-Bobea, A. (2020). Temperature shocks and establishment sales. The Review of Financial Studies, 33(3): 1331–1366.

Agrawal, A., Kapur, D., and McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. Journal of Urban Economics, 64(2): 258–269.

Ahir, H., Bloom, N., and Furceri, D. (2022). The World Uncertainty Index. National Bureau of Economic Research Working Paper.

Alfaro, L. and Chen, M. X. (2018). Transportation cost and the geography of foreign investment. In Handbook of International Trade and Transportation. Edward Elgar Publishing.

Allcott, H., Braghieri, L., Eichmeyer, S., and Gentzkow, M. (2020). The welfare effects of social media. American Economic Review, 110(3): 629–76.

Avom, D., Njangang, H., Nawo, L., et al. (2020). World economic policy uncertainty and foreign direct investment. Economics Bulletin, 40(2): 1457–1464.

Azzimonti, M. (2019). Does partisan conflict deter FDI inflows to the US? Journal of International Economics, 120: 162–178.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. Journal of Economic Perspectives, 32(3): 259–280.

Bailey, M., Dávila, E., Kuchler, T., and Stroebel, J. (2019). House price beliefs and mortgage leverage choice. The Review of Economic Studies, 86(6): 2403–2452.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., and Stroebel, J. (2021). International trade and social connectedness. Journal of International Economics, 129: 103418.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. The Quarterly Journal of Economics, 131(4): 1593–1636.

Balakina, O., Bäckman, C., Hackethal, A., Hanspal, T., and Lammer, D. (2023). Personal recommendations and portfolio quality. SAFE Working Paper.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2022). The time variation in risk appetite and uncertainty. Management Science, 68(6): 3975–4004.

Berger, A., Busse, M., Nunnenkamp, P., and Roy, M. (2013). Do trade and investment agreements lead to more FDI? Accounting for key provisions inside the black box. International Economics and Economic Policy, 10(2): 247–275.

Bergstrand, J. H. and Egger, P. (2007). A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises. Journal of International Economics, 73(2): 278–308.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. The Quarterly Journal of Economics, 98(1): 85–106.

Biswas, R. (2002). Determinants of foreign direct investment. Review of Development Economics, 6(3): 492–504.

Blonigen, B. A. and Piger, J. (2014). Determinants of foreign direct investment. Canadian Journal of Economics, 47(3): 775–812.

Bloom, N. (2009). The impact of uncertainty shocks. Econometrica, 77(3): 623–685.

Bloom, N. (2014). Fluctuations in uncertainty. Journal of Economic Perspectives, 28(2): 153–176.

Bloom, N., Bond, S., and Van Reenen, J. (2007). Uncertainty and investment dynamics. The Review of Economic Studies, 74(2): 391–415.

Brown, J. R., Ivković, Z., Smith, P. A., and Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. The Journal of Finance, 63(3): 1509–1531.

Burchardi, K. B., Chaney, T., and Hassan, T. A. (2019). Migrants, ancestors, and foreign investments. The Review of Economic Studies, 86(4): 1448–1486.

Burchardi, K. B. and Hassan, T. A. (2013). The economic impact of social ties: Evidence from German reunification. The Quarterly Journal of Economics, 128(3): 1219–1271.

Büthe, T. and Milner, H. V. (2008). The politics of foreign direct investment into developing countries: Increasing FDI through international trade agreements? American Journal of Political Science, 52(4): 741–762.

Cai, X., Lu, Y., Wu, M., and Yu, L. (2016). Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. Journal of Development Economics, 123: 73–85.

Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. American Economic Review, 112(4): 1194–1225.

Canh, N. P., Binh, N. T., Thanh, S. D., and Schinckus, C. (2020). Determinants of foreign direct investment inflows: The role of economic policy uncertainty. International Economics, 161: 159–172.

Cascaldi-Garcia, D., Sarisoy, C., Londono, J. M., Sun, B., Datta, D. D., Ferreira, T., Grishchenko, O., Jahan-Parvar, M. R., Loria, F., Ma, S., et al. (2023). What is certain about uncertainty? Journal of Economic Literature, 61(2): 624–654.

Choi, S., Furceri, D., and Yoon, C. (2021). Policy uncertainty and foreign direct investment. Review of International Economics, 29(2): 195–227.

Chung, S. (2014). Environmental regulation and foreign direct investment: Evidence from South Korea. Journal of Development Economics, 108: 222–236.

Cohen, L., Gurun, U. G., and Malloy, C. (2017). Resident networks and corporate connections: Evidence from world war ii internment camps. The Journal of Finance, 72(1):207–248.

Cole, M. A., Elliott, R. J., and Zhang, L. (2017). Foreign direct investment and the environment. Annual Review of Environment and Resources, 42(1): 465–487.

Combes, P.-P., Lafourcade, M., and Mayer, T. (2005). The trade-creating effects of business and social networks: Evidence from France. Journal of International Economics, 66(1): 1–29.

Corcoran, A. and Gillanders, R. (2015). Foreign direct investment and the ease of doing business. Review of World Economics, 151: 103–126.

Damgaard, J., Elkjaer, T., and Johannesen, N. (2024). What is real and what is not in the global FDI network? Journal of International Money and Finance, 140: 102971.

Desbordes, R. and Wei, S.-J. (2017). The effects of financial development on foreign direct investment. Journal of Development Economics, 127: 153–168.

Di Maggio, M., Kermani, A., and Song, Z. (2017). The value of trading relations in turbulent times. Journal of Financial Economics, 124(2): 266–284.

Eckstein, D., Künzel, V., and Schäfer, L. (2021). The global climate risk index 2021. Bonn: Germanwatch.

Egger, P. and Pfaffermayr, M. (2004). Distance, trade and FDI: A Hausman–Taylor SUR approach. Journal of Applied Econometrics, 19(2): 227–246.

Facebook, I. (2020). Facebook reports third quarter 2020 results. https://s21.q4cdn.com/399680738/files/doc_financials/2020/q3/FB-09.30.2020-Exhibit-99.1.pdf.

Fisman, R. (2001). Estimating the value of political connections. American Economic Review, 91(4): 1095–1102.

Fourné, M. and Li, X. (2023). Climate policy and international capital reallocation. Available at SSRN 4741752.

Georgarakos, D., Haliassos, M., and Pasini, G. (2014). Household debt and social interactions. The Review of Financial Studies, 27(5): 1404–1433.

Giglio, S., Kelly, B., and Stroebel, J. (2021). Climate finance. Annual Review of Financial Economics, 13: 15–36.

Ginglinger, E. and Moreau, Q. (2023). Climate risk and capital structure. Management Science, 69(12): 7492–7516.

Giroud, A. and Ivarsson, I. (2020). World Investment Report 2020: International production beyond the pandemic: United Nations Conference on Trade and Development, Geneva and New York, 2020, 247 pp. isbn: 978-9211129854.

Greif, A. (1989). Reputation and coalitions in medieval trade: Evidence on the Maghribi traders. The Journal of Economic History, 49(4): 857–882.

Gu, G. W. and Hale, G. (2023). Climate risks and FDI. Journal of International Economics, 146: 103731.

Gulen, H. and Ion, M. (2016). Policy uncertainty and corporate investment. The Review of Financial Studies, 29(3): 523–564.

Hanna, R. (2010). US environmental regulation and FDI: Evidence from a panel of US-based multinational firms. American Economic Journal: Applied Economics, 2(3): 158–189.

Head, K. and Ries, J. (2008). FDI as an outcome of the market for corporate control: Theory and evidence. Journal of International Economics, 74(1): 2–20.

Hellmanzik, C. and Schmitz, M. (2017). Taking gravity online: The role of virtual proximity in international finance. Journal of International Money and Finance, 77: 164–179.

Hong, H., Kubik, J. D., and Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. The Journal of Finance, 60(6): 2801–2824.

Hsieh, H.-C., Boarelli, S., and Vu, T. H. C. (2019). The effects of economic policy uncertainty on outward foreign direct investment. International Review of Economics & Finance, 64: 377–392.

Jardet, C., Jude, C., and Chinn, M. (2023). Foreign direct investment under uncertainty evidence from a large panel of countries. Review of International Economics, 31(3): 854–885.

Jayasuriya, D. (2011). Improvements in the world bank's ease of doing business rankings: Do they translate into greater foreign direct investment inflows? World Bank Policy Research Working Paper, (5787).

Julio, B. and Yook, Y. (2016). Policy uncertainty, irreversibility, and cross-border flows of capital. Journal of International Economics, 103: 13–26.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring Uncertainty. American Economic Review, 105(3): 1177–1216.

Kox, H. L. and Rojas-Romagosa, H. (2020). How trade and investment agreements affect bilateral foreign direct investment: Results from a structural gravity model. The World Economy, 43(12): 3203–3242.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). Social proximity to capital: Implications for investors and firms. The Review of Financial Studies, 35(6): 2743–2789.

Kuchler, T., Russel, D., and Stroebel, J. (2022b). JUE insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook. Journal of Urban Economics, 127: 103314.

Kuchler, T. and Stroebel, J. (2021). Social finance. Annual Review of Financial Economics, 13: 37–55.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. The Journal of Law, Economics, and Organization, 15(1): 222–279.

Li, X. and Gallagher, K. P. (2022). Assessing the climate change exposure of foreign direct investment. Nature Communications, 13(1): 1451.

Liu, J. (2024). What works in taiwan doesn't always in arizona, a chipmaking giant learns. The New York Times. Accessed: 2024-09-21.

Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and business cycles: Exogenous impulse or endogenous response? American Economic Journal: Macroeconomics, 13(4): 369–410.

Ly, A., Esperança, J., and Davcik, N. S. (2018). What drives foreign direct investment: The role of language, geographical distance, information flows and technological similarity. Journal of Business Research, 88: 111–122.

Nguyen, C. P. and Lee, G. S. (2021). Uncertainty, financial development, and FDI inflows: Global evidence. Economic Modelling, 99: 105473.

Ni, L., Li, L., Zhang, X., and Wen, H. (2022). Climate policy and foreign direct investment: Evidence from a quasi-experiment in Chinese cities. Sustainability, 14(24): 16469.

Novy, D. and Taylor, A. M. (2020). Trade and uncertainty. Review of Economics and Statistics, 102(4): 749–765.

Okawa, Y. and Van Wincoop, E. (2012). Gravity in international finance. Journal of International Economics, 87(2): 205–215.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics, 37(2): 187–204.

Ouimet, P. and Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. The Journal of Finance, 75(1): 133–172.

Paniagua, J., Korzynski, P., and Mas-Tur, A. (2017). Crossing borders with social media: Online social networks and FDI. European Management Journal, 35(3): 314–326.

Paul, J. and Feliciano-Cestero, M. M. (2021). Five decades of research on foreign direct investment by MNEs: An overview and research agenda. Journal of Business Research, 124: 800–812.

Pool, V. K., Stoffman, N., and Yonker, S. E. (2015). The people in your neighborhood: Social interactions and mutual fund portfolios. The Journal of Finance, 70(6): 2679–2732.

Portes, R. and Rey, H. (2005). The determinants of cross-border equity flows. Journal of International Economics, 65(2): 269–296.

Raff, H., Ryan, M., and Stähler, F. (2018). Financial frictions and foreign direct investment: Evidence from Japanese microdata. Journal of International Economics, 112: 109–122.

Rauch, J. E. (2001). Business and social networks in international trade. Journal of Economic Literature, 39(4): 1177–1203.

Rehbein, O. and Rother, S. (2024). Social connectedness in bank lending. Working Paper.

Sautner, Z., Van Lent, L., Vilkov, G., and Zhang, R. (2023). Firm-level climate change exposure. The Journal of Finance, 78(3): 1449–1498.

Schiavo, S. (2007). Common currencies and FDI flows. Oxford Economic Papers, 59(3): 536–560.

Segal, G., Shaliastovich, I., and Yaron, A. (2015). Good and bad uncertainty: Macroeconomic and financial market implications. Journal of Financial Economics, 117(2): 369–397.

Spolaore, E. and Wacziarg, R. (2013). How deep are the roots of economic development? Journal of Economic Literature, 51(2): 325–369.

Spolaore, E. and Wacziarg, R. (2018). Ancestry and development: New evidence. Journal of Applied Econometrics, 33(5): 748–762.

Wei, S.-J. (2000). How taxing is corruption on international investors? Review of Economics and Statistics, 82(1): 1–11.

Figures

Figure 1.1: Social Connectedness Between the U.S. and Other Countries Across the Globe

This figure presents the connectedness of the United States with other countries across the world. Connectedness is based on Facebook's Social Connectedness Index (SCI). The data is split into quintiles for the purposes of this figure. Quintiles are shaded from dark (high connectedness) to light (low connectedness). White Countries are countries without access to Facebook.

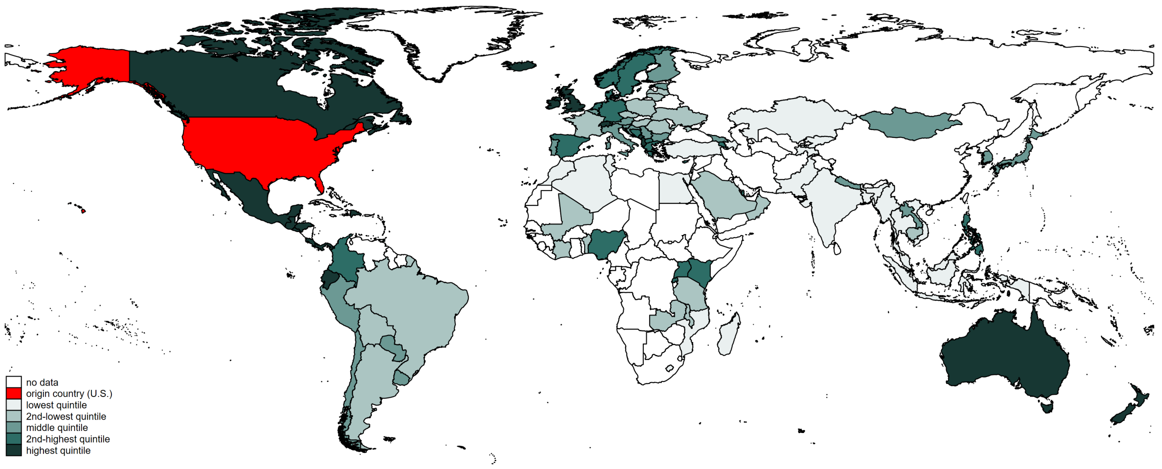


Figure 1.2: Social Connectedness Between Austria and Other European Countries

This figure presents the connectedness of Austria with other countries in Europe. Connectedness is based on Facebook's Social Connectedness Index (SCI). The data is split into deciles for the purposes of this figure. Deciles are shaded from dark (high connectedness) to light (low connectedness). Only the 5 most connectedness deciles are represented in European countries, so they are displayed. White countries are countries without access to Facebook.

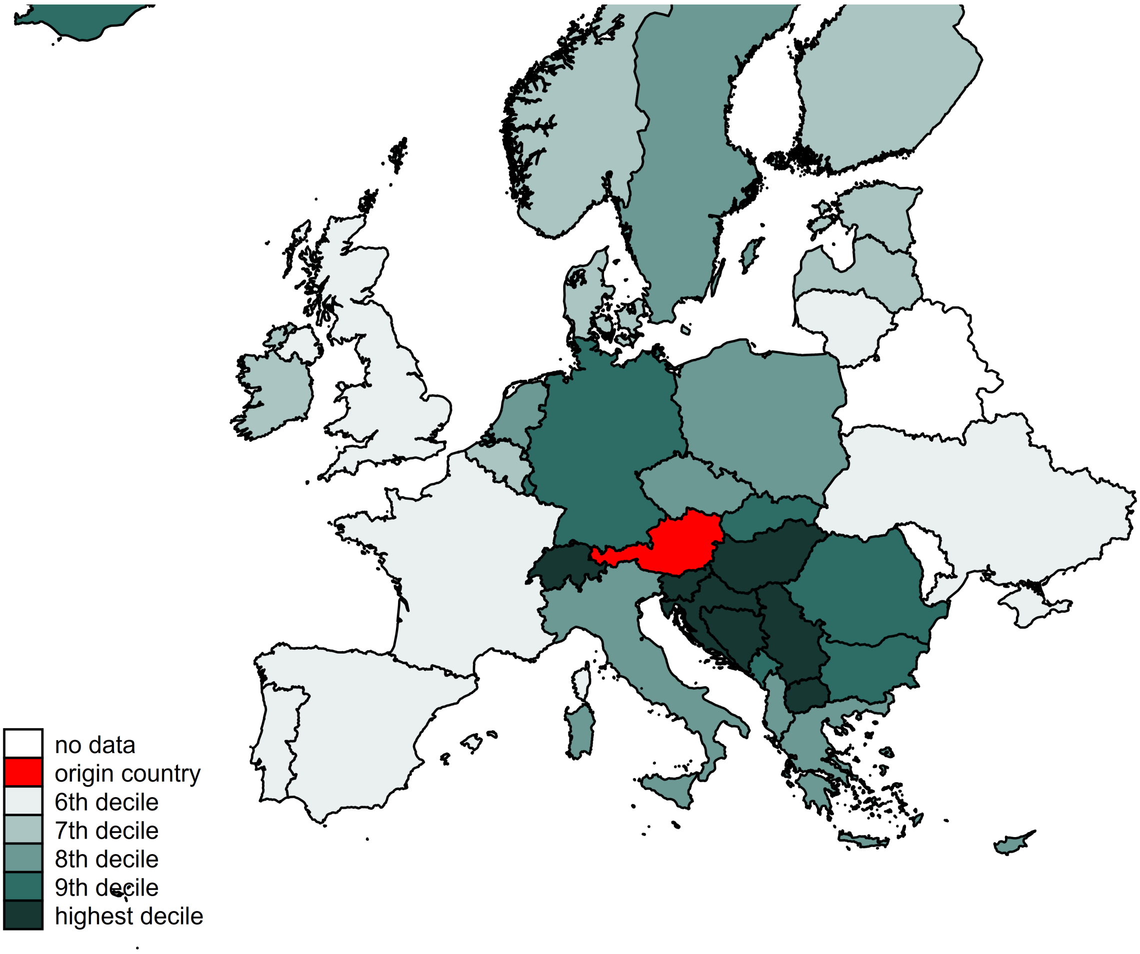
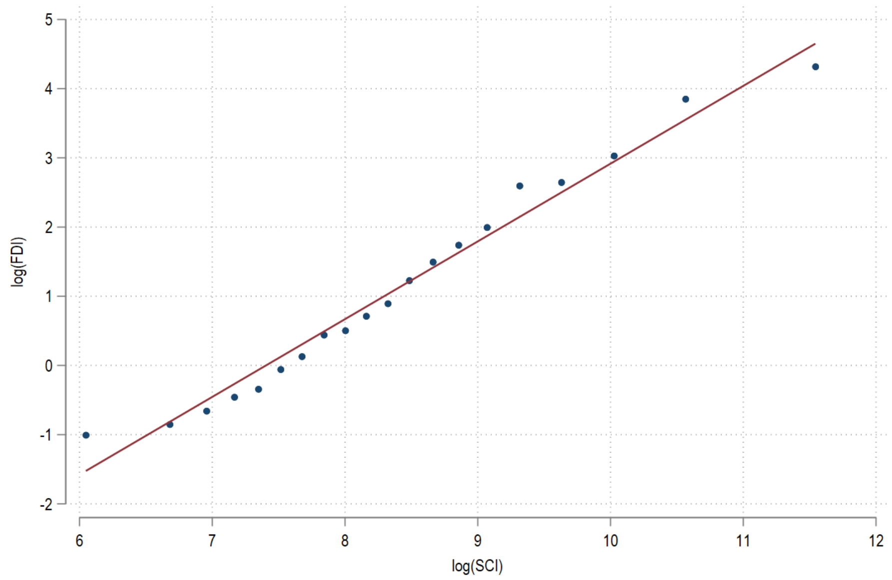


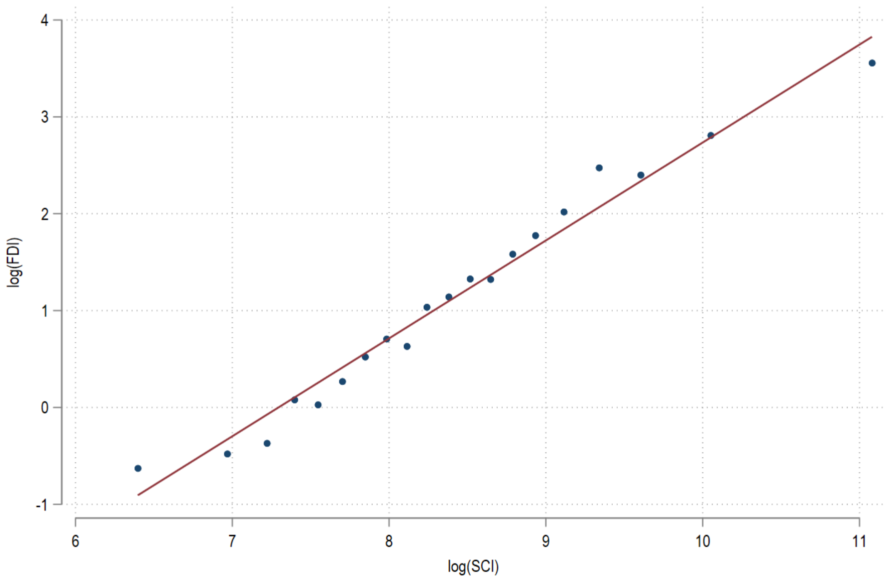
Figure 1.3: Foreign Direct Investment vs. Social Connectedness

This figure shows a binned scatterplot of inward FDI and the Social Connectedness Index (SCI). Panel A regresses the logarithm of FDI on the logarithm of SCI, Panel B shows the same specification including all gravity variables as control variables. Both panels incorporate country and year fixed effects. By taking the logarithm of FDI all country pairs with zero FDI are dropped. The remaining data set contains 31,442 observations and represents the intensive margin of FDI.

Panel A: Country and Year Fixed Effects



Panel B: Gravity Controls, Country and Year Fixed Effects



Tables

Table 1.1: Descriptive Statistics

This table displays descriptive statistics for the main data set. All variables are at the country pair level. For example, inward FDI in million USD captures FDI of investors from the origin countries to the destination countries. The data includes 157 origin countries and 114 destination countries as an unbalanced panel from 2010 until 2019 with a total number of country pairs equal to 6,552. A detailed description of all variables can be found in Table A1.1.

	Mean	Median	SD	Min	Max	Obs
Main Variables						
Inward FDI in Million USD	339.25	0.80	2,674	0.00	172,740	45,345
log(SCI)	8.34	8.20	1.65	4.19	12.89	45,345
Control Variables						
log(Distance)	8.25	8.45	1.00	4.88	9.89	45,345
log(Trade Volume)	19.36	19.55	2.80	5.78	27.11	45,345
log(GDP Difference)	26.68	26.76	1.94	14.39	30.61	45,345
log(GDP Growth Difference)	0.62	0.85	1.16	-8.51	3.48	45,345
Regional Trade Agreement	0.43	0.00	0.49	0.00	1.00	45,345
Common Currency	0.08	0.00	0.27	0.00	1.00	45,345
Shared Religion Index	0.44	0.50	0.29	0.00	0.99	45,281
Political Distance	0.19	0.11	0.25	0.00	1.00	45,345
Common Legal Origin	0.37	0.00	0.48	0.00	1.00	45,345
Common Ethno. Language	0.14	0.00	0.34	0.00	1.00	45,345
Common Border	0.15	0.00	0.36	0.00	1.00	45,345
Common Official Language	0.13	0.00	0.33	0.00	1.00	45,345
Common Colonizer	0.05	0.00	0.21	0.00	1.00	45,345
Colonial Relationship after 1945	0.02	0.00	0.12	0.00	1.00	45,345
Instrumental Variable						
Genetic Distance	0.02	0.02	0.02	0.00	0.08	41,415
Information Frictions						
Dest. Regulatory Quality	0.64	0.65	0.81	-2.24	2.26	45,345
Orig. Regulatory Quality	0.65	0.71	0.93	-2.35	2.26	45,281
Dest. Government Effectiveness	0.58	0.53	0.84	-1.64	2.24	45,345
Orig. Government Effectiveness	0.65	0.65	0.95	-2.08	2.24	45,281
Cost of Doing Business						
Dest. Days to Start a Business	15.90	11.00	16.49	0.50	174.00	45,306
Orig. Days to Start a Business	16.45	11.00	21.15	0.50	690.50	45,246
Dest. Procedures to Start a Business	6.60	6.00	2.86	1.00	16.00	45,306
Orig. Procedures to Start a Business	6.58	6.00	3.02	1.00	18.00	45,246
Dest. Costs to Start a Business (GNI share)	0.10	0.05	0.16	0.00	1.58	45,306

Table 1.1: Descriptive Statistics (continued)

	Mean	Median	SD	Min	Max	Obs
Orig. Costs to Start a Business (GNI share)	0.12	0.05	0.25	0.00	7.35	45,246
Dest. Financial Market Development	0.38	0.34	0.29	0.00	0.97	44,733
Orig. Financial Market Development	0.42	0.43	0.30	0.00	0.97	45,010
Physical Risk						
Dest. Global Climate Risk Index	0.69	0.68	0.31	0.02	1.26	42,995
Orig. Global Climate Risk Index	0.71	0.70	0.31	0.02	1.26	41,741
Dest. Number of Floods	1.40	1.00	1.88	0.00	15.00	33,486
Orig. Number of Floods	1.30	1.00	1.83	0.00	15.00	32,421
Dummy for Flood in Destination Country	0.64	1.00	0.48	0.00	1.00	33,486
Dummy for Flood in Origin Country	0.60	1.00	0.49	0.00	1.00	32,421
Dest. % Population Exposed to Wildfires	2.88	0.00	7.08	0.00	49.56	23,222
Orig. % Population Exposed to Wildfires	2.85	0.00	7.38	0.00	49.56	21,842
Dest. Climate Change Physical Risk Sautner et al. (2023)	0.05	0.00	0.25	0.00	3.77	28,144
Orig. Climate Change Physical Risk Sautner et al. (2023)	0.07	0.00	0.30	0.00	3.77	29,965
Transition Risk						
Dest. log(CO2 Emissions)	1.41	1.58	1.02	-2.70	3.68	45,292
Orig. log(CO2 Emissions)	1.44	1.67	1.12	-3.68	3.68	44,639
Dest. Renewable Energy Supply	0.21	0.13	0.22	0.00	1.50	45,000
Orig. Renewable Energy Supply	0.20	0.13	0.21	0.00	1.50	42,832
Dest. Development of Env.-related Technologies	0.13	0.12	0.09	0.00	1.00	43,839
Orig. Development of Env.-related Technologies	0.13	0.12	0.10	0.00	1.00	42,659
Dest. Changes in Emission Related Policies	0.05	0.04	0.06	0.00	0.30	40,800
Orig. Changes in Emission Related Policies	0.05	0.04	0.05	0.00	0.30	38,096
Dest. Energy Related Tax Revenue	0.66	0.74	0.28	-0.74	2.03	38,664
Orig. Energy Related Tax Revenue	0.65	0.69	0.26	-0.74	2.03	35,280
Dest. Climate Change Regulatory Risk Sautner et al. (2023)	0.03	0.00	0.16	0.00	2.89	28,144
Orig. Climate Change Regulatory Risk Sautner et al. (2023)	0.03	0.00	0.14	0.00	2.89	29,965
Macroeconomic Uncertainty						
Dest. WUI	0.22	0.21	0.13	0.03	0.50	40,325
Orig. WUI	0.22	0.20	0.14	0.03	0.54	40,621
Global WUI	0.22	0.23	0.04	0.15	0.28	45,345
Global GDP-Weighted WUI	0.24	0.23	0.06	0.17	0.41	45,345
log(Geopolitical Risk Index)	4.55	4.59	0.10	4.40	4.68	45,345
Macro Uncertainty Index Jurado et al. (2015)	0.88	0.87	0.02	0.85	0.91	45,345
Real Uncertainty Index Ludvigson et al. (2021)	0.86	0.86	0.01	0.84	0.90	45,345

Table 1.2: Social Connectedness and FDI: Baseline Results

This table shows the baseline results from estimating Equations 1.2 (PPML) and 1.3 (OLS). The dependent variable is inward FDI from country i to country j in year t. The main explanatory variable is the log(SCI). The model controls for gravity covariates (log of distance, a common border dummy, a common official language dummy, a dummy for a common colonizer post 1945, and a dummy for a colonial relationship post 1945), economic covariates (log of the total trade volume, log of the GDP difference, log of the GDP growth rate difference, a dummy for a regional trade agreement in place and a dummy indicating a shared currency) and cultural covariates (a shared religion index, a measure for political distance, a dummy for a common legal origin and a dummy indicating whether both countries share a common ethnological language). Columns (1) and (2) include fixed effects only, Columns (3) and (4) include the full set of controls. In Columns (5)-(6) genetic distance is used as an instrument for the SCI. First stage coefficients, p-value and F-value for the instruments are displayed at the bottom of the table, first stage results in Table A1.6. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI	log(FDI)	FDI	log(FDI)	FDI	log(FDI)
Model	PPML	OLS	PPML	OLS	PPML	OLS
Instrumental Variable					Genetic Distance	
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.513***	1.098***	0.121**	0.355***	0.528**	0.459***
	(0.036)	(0.025)	(0.052)	(0.043)	(0.268)	(0.144)
Residuals					-0.425	
					(0.273)	
log(Distance)			0.085	-0.856***	0.358*	-0.755***
			(0.085)	(0.088)	(0.211)	(0.119)
log(Trade Volume)			0.485***	0.367***	0.386***	0.335***
			(0.056)	(0.036)	(0.081)	(0.046)
log(GDP Difference)			-0.089***	-0.220***	-0.089***	-0.219***
			(0.031)	(0.033)	(0.032)	(0.035)
log(GDP Growth Difference)			-0.088***	0.021	-0.091***	0.015
			(0.023)	(0.026)	(0.024)	(0.029)
Regional Trade Agreement			0.160	-0.123	0.088	-0.152*
			(0.113)	(0.111)	(0.126)	(0.083)
Common Currency			-0.142	-0.040	-0.084	0.101
			(0.162)	(0.197)	(0.165)	(0.172)
Shared Religion Index			1.344***	0.954***	0.730	0.804***
			(0.338)	(0.184)	(0.512)	(0.212)
Political Distance			-0.297	-0.486**	-0.248	-0.412**
			(0.279)	(0.205)	(0.284)	(0.162)
Common Legal Origin			0.114	-0.057	0.037	-0.010
			(0.075)	(0.086)	(0.086)	(0.065)
Common Ethno. Language			0.136	0.076	-0.022	0.058
			(0.140)	(0.212)	(0.160)	(0.140)
Common Border			-0.408***	-0.292*	-0.442***	-0.240**
			(0.148)	(0.165)	(0.148)	(0.097)
Common Official Language			-0.011	0.616***	-0.212	0.479***
			(0.151)	(0.226)	(0.207)	(0.151)
Common Colonizer			0.287	0.280	0.027	0.264
			(0.367)	(0.212)	(0.451)	(0.161)
Colonial Relationship after 1945			0.223	0.691**	-0.103	0.532**
			(0.205)	(0.301)	(0.316)	(0.243)
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.762		0.780		0.776	
R^2		0.614		0.635		0.213
Instrument (1st stage)					-17.371	-17.371
p-value (1st stage)					0.000	0.000
F-value (1st stage)					127.5	200.6
Observations	45,280	31,364	45,216	31,332	41,351	28,584

Table 1.3: Social Connectedness and FDI: Trade, Culture and Institutions

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with the bilateral trade volume, dummies for regional trade agreements, a common official language, a common legal origin, a measure for political distance and two institutional factors between two countries. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. At the bottom, the table reports the elasticity of SCI for case that the interacted variable takes different values; Column (1) displays the elasticities for the trade volume between country i and country j at the median, the 25^{th} and 75^{th} percentile. In Columns (2), (3) and (4) the elasticity of the SCI in case the respective dummy variables equals 1 is presented. Column (5) represents the elasticities for the political distance between country i and country j at the median, the 25^{th} and 75^{th} percentile. Columns (6) and (7) include the interaction of the SCI with the destination and origin countries' value for regulatory quality and government effectiveness, taken from the Worldwide Governance Indicators. Elasticities are reported under the assumption that the value of regulatory quality and government effectiveness in the origin country equals the median. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(SCI)	0.647**	0.283***	0.195***	0.222***	0.069	0.288***	0.290***
	(0.303)	(0.065)	(0.053)	(0.052)	(0.056)	(0.067)	(0.067)
log(SCI) × log(Trade Volume)	-0.025*						
	(0.014)						
log(SCI) × Regional Trade Agreement		-0.272***					
		(0.066)					
log(SCI) × Common Legal Origin			-0.119***				
			(0.043)				
log(SCI) × Common Official Language				-0.318***			
				(0.071)			
log(SCI) × Political Distance					0.351***		
					(0.123)		
log(SCI) × Dest. Regulatory Quality						-0.197***	
						(0.036)	
log(SCI) × Orig. Regulatory Quality						-0.040	
						(0.040)	
log(SCI) × Dest. Government Effectiveness							-0.179***
							(0.036)
log(SCI) × Orig. Government Effectiveness							-0.054
							(0.038)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms							
Origin Country at Median and Destination at:							
Dummy=1		0.011	0.076	-0.096			
Median	0.158				0.108	0.132	0.160
25^{th} Percentile	0.204				0.073	0.266	0.264
75^{th} Percentile	0.115				0.178	0.007	0.004
Pseudo R^2	0.780	0.781	0.781	0.782	0.780	0.783	0.782
Observations	45,216	45,216	45,216	45,216	45,216	45,216	45,216

Table 1.4: Social Connectedness and FDI: Regulation and Financing

This table presents the results of Equation 1.2, estimating the influence of the SCI on FDI by interacting the SCI with variables for the time, procedures and costs that are required to start a business in the destination country as well as the destination country's financial market development. To check for robustness, interaction terms for the origin countries are also included. Time and procedures are simple count data, cost is originally measured as percentage of gross national income and rescaled by dividing by 100. The index for the financial market development ranges between 0 and 1 with higher values corresponding to more developed financial markets. Variables for Columns (1) to (3) are winsorized at the 99th percentile to account for outliers, data is taken from the CEPII database. Data for the financial market development is taken from the IMF Financial Development Index Database. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. At the bottom, the table reports the elasticity of SCI for different values of the interacted variable for the destination country of the investment. These are are the median, the 25^{th} and 75^{th} percentile, while the value of the interacted variable for the origin country is held at the median for all three cases. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI			
	(1)	(2)	(3)	(4)
log(SCI)	-0.008	-0.206**	0.035	0.386***
	(0.066)	(0.099)	(0.059)	(0.100)
log(SCI) × Dest. Days to Start a Business	0.005***			
	(0.001)			
log(SCI) × Orig. Days to Start a Business	0.003*			
	(0.002)			
log(SCI) × Dest. Procedures to Start a Business		0.043***		
		(0.008)		
log(SCI) × Orig. Procedures to Start a Business		0.006		
		(0.010)		
log(SCI) × Dest. Costs to Start a Business (GNI share)			0.739***	
			(0.184)	
log(SCI) × Orig. Costs to Start a Business (GNI share)			0.174	
			(0.208)	
log(SCI) × Dest. Financial Market Development				-0.454***
				(0.129)
log(SCI) × Orig. Financial Market Development				-0.124
				(0.128)
Full Set of Controls	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms				
Origin Country at Median and Destination at:				
Median	0.080	0.088	0.081	0.178
25^{th} Percentile	0.055	0.045	0.051	0.305
75^{th} Percentile	0.123	0.174	0.140	0.047
Pseudo R^2	0.781	0.782	0.781	0.780
Observations	45,079	45,079	45,079	44,333

Table 1.5: Social Connectedness and Climate Risk: Physical Risk

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with indicators for physical climate risk and natural disasters. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Column (1) refers to the Global Climate Risk Index taken from Germanwatch. Higher values correspond to lower risk. Column (2) uses the number of floods per country and year, taken from the International Disaster Database (EM-DAT). Column (3) translates these data into a dummy, that is equal to one if there has been at least one flooding event in the respective country and year. Column (4) refers to the percentage of population exposed to wildfires per country and year; the respective data is taken from the Green Growth Indicators. In Column (5) firm-level data from Sautner et al. (2023) on physical risk arising from climate change is used; we aggregate this data to the country-year level. Higher values correspond to higher risk. At the bottom, the table reports the elasticity of SCI for different values of the interacted variable. In Columns (1) and (2), elasticity for the median, the 25^{th} and the 75^{th} percentile are presented for the Global Climate Risk Index and the Number of flooding events. The value of the interacted variable for the origin country is held at the median in all three cases. In Column (3), the elasticity of the SCI is shown for the case that there has been at least one flooding event in the destination county as well as in the origin country. Columns (4) and (5) display the elasticity of the SCI for the case that the percentage of population exposed to wildfire is at the 90^{th} percentile in the destination country while is is held at the median for the origin country because of a large fraction (>50%) of zeros in the interacted variables. The same logic applies to Column (5) for physical risk arising from climate change as measured by Sautner et al. (2023). Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI				
	(1)	(2)	(3)	(4)	(5)
log(SCI)	0.513***	-0.036	-0.099	-0.021	0.110*
	(0.086)	(0.060)	(0.068)	(0.084)	(0.062)
log(SCI) × Dest. Global Climate Risk Index	-0.444***				
	(0.081)				
log(SCI) × Orig. Global Climate Risk Index	-0.179**				
	(0.071)				
log(SCI) × Dest. Number of Floods		0.054***			
		(0.013)			
log(SCI) × Orig. Number of Floods		0.056***			
		(0.014)			
log(SCI) × Dummy for Flood in Destination Country			0.204***		
			(0.051)		
log(SCI) × Dummy for Flood in Origin Country			0.128***		
			(0.040)		
log(SCI) × Dest. % Population Exposed to Wildfires				0.013**	
				(0.006)	
log(SCI) × Orig. % Population Exposed to Wildfires				0.014**	
				(0.005)	
log(SCI) × Dest. Climate Change Physical Risk Sautner et al. (2023)					0.142
					(0.088)
log(SCI) × Orig. Climate Change Physical Risk Sautner et al. (2023)					0.190***
					(0.060)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms					
Origin Country at Median and Destination at:					
Median	0.086	0.074			
25^{th} Percentile	0.188	0.020			
75^{th} Percentile	-0.030	0.128			
Dummy=1			0.233		
90^{th} Percentile				0.098	0.119
Pseudo R^2	0.785	0.810	0.809	0.754	0.739
Observations	39,491	24,134	24,134	10,237	17,817

Table 1.6: Social Connectedness and Climate Risk: Transition Risk

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with indicators for transition risk. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Column (1) includes an interaction term of the SCI and the log of CO2-emissions in tons per capita taken from the Worldbank. Columns (2), (3) and (5) use measures for the renewable energy supply, the development of environmental related policies and energy related tax revenue from the Green Growth Indicators. Column (4) interacts the SCI with the changes in emission related policies as used by Gu and Hale (2023). Column (6) refers to climate change regulatory risk as measured by Sautner et al. (2023). At the bottom, the table reports the elasticity of SCI for different values of the interacted variable for the destination country of the investment. These are are the median, the 25^{th} and 75^{th} percentile, while the value of the interacted variable for the origin country is held at the median for all three cases. Column (6) and displays the elasticity of the SCI for the case that the climate change regulatory risk measure is at the 90^{th} percentile in the destination country while is is held at the median for the origin country because of a large fraction (>50%) of zeros in the interacted variable. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.453***	-0.119*	0.130	0.057	0.103	0.129**
	(0.091)	(0.063)	(0.083)	(0.058)	(0.124)	(0.062)
log(SCI) × Dest. log(CO2 Emissions)	-0.199***					
	(0.033)					
log(SCI) × Orig. log(CO2 Emissions)	-0.040					
	(0.037)					
log(SCI) × Dest. Renewable Energy Supply		0.953***				
		(0.163)				
log(SCI) × Orig. Renewable Energy Supply		0.278				
		(0.183)				
log(SCI) × Dest. Development of Env.-related Technologies			-0.210			
			(0.322)			
log(SCI) × Orig. Development of Env.-related Technologies			-0.052			
			(0.357)			
log(SCI) × Dest. Changes in Emission Related Policies				0.558		
				(0.404)		
log(SCI) × Orig. Changes in Emission Related Policies				-0.129		
				(0.404)		
log(SCI) × Dest. Energy Related Tax Revenue					0.001	
					(0.098)	
log(SCI) × Orig. Energy Related Tax Revenue					0.027	
					(0.133)	
log(SCI) × Dest. Climate Change Regulatory Risk Sautner et al. (2023)						1.115**
						(0.444)
log(SCI) × Orig. Climate Change Regulatory Risk Sautner et al. (2023)						-0.094
						(0.244)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms						
Origin Country at Median and Destination at:						
Median	0.072	0.041	0.099	0.074	0.122	
25^{th} Percentile	0.193	-0.007	0.105	0.057	0.122	
75^{th} Percentile	-0.024	0.165	0.092	0.096	0.122	
90^{th} Percentile						0.174
Pseudo R^2	0.787	0.787	0.783	0.782	0.776	0.739
Observations	44,522	42,461	41,250	34,191	29,975	17,817

Table 1.7: Social Connectedness and FDI: Macroeconomic Uncertainty

This table shows the results for Equation 1.3, estimating the influence of the SCI interacted with different measures for uncertainty. These include local and global versions of the World Uncertainty Index (Ahir et al., 2022), the Geopolitical Risk Index (Caldara and Iacoviello, 2022) the Real Uncertainty Index (Ludvigson et al., 2021) and the Macro Uncertainty Index (Jurado et al., 2015). Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. The table reports the elasticity of SCI at the median, the 25^{th} and the 75^{th} percentile of the interacted (destination country) indicators. For the case of the country-specific World Uncertainty Index in Column (1) it reports the elasticity of SCI at the median of the origin country and the median, the 25^{th} and the 75^{th} percentile of the interacted (destination country) value. At the bottom, the table reports the elasticity of SCI for different values of the interacted variable for the destination country of the investment. These are are the median, the 25^{th} and 75^{th} percentile, while the value of the interacted variable for the origin country is held at the median for all three cases. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	log(FDI)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.217***	-0.225***	0.084	-3.471***	-1.002**	-5.563***
	(0.058)	(0.083)	(0.063)	(0.580)	(0.447)	(0.843)
log(SCI) × Dest. WUI	0.218*					
	(0.116)					
log(SCI) × Orig. WUI	0.131					
	(0.126)					
log(SCI) × Global WUI		2.327***				
		(0.340)				
log(SCI) × Global GDP-Weighted WUI			0.878***			
			(0.202)			
log(SCI) × log(Geopolitical Risk Index)				0.841***		
				(0.128)		
log(SCI) × Macro Uncertainty Index Jurado et al. (2015)					1.544***	
					(0.506)	
log(SCI) × Real Uncertainty Index Ludvigson et al. (2021)						6.845***
						(0.979)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms						
Origin Country at Median and Destination at:						
Median	0.244	0.310	0.286	0.389	0.341	0.324
25^{th} Percentile	0.222	0.240	0.251	0.255	0.326	0.324
75^{th} Percentile	0.268	0.333	0.312	0.406	0.388	0.392
R^2	0.641	0.642	0.641	0.636	0.635	0.636
Observations	25,144	25,144	25,144	31,332	31,332	31,332

CHAPTER 2

Close By or Closed By? Bank Branches and the Rise of Fintech Mortgages[1]

Abstract

Online mortgage providers have flourished over the last decade while banks have downsized branch networks. This paper shows that rising market shares of fintech lenders in the US residential mortgage market account for a significant share of the decrease in bank branches. On average the rise of fintechs can explain 45% of the overall losses of branches per county. The relationship is more pronounced in metro areas and among small banks and leads to lower local deposits. To strengthen identification, I exploit the timing of the surge in fintech penetration and an instrumental variable based on Facebook's Social Connectedness Index to Wayne County, the headquarter location of Quicken Loans, the largest fintech lender.

[1]Unpublished

2.1 Introduction

In 2019, Rocket Mortgage was the largest volume lender in the US residential mortgage market, originating about 541,000 loans (Barriere et al., 2020). This is just one example of an emerging number of financial technology ("fintech") companies that have extensively started offering financial products over the last decade (see Berg et al. (2022) for a comprehensive overview). Especially in residential mortgages, a segment that had previously been occupied by traditional brick and mortar banks, fintech lenders increased their market share from 2.8% in 2010 to 10.4% in 2017. At the same time, many banks consolidated their branch network and local offline access to finance decreased in some parts of the country (Richardson et al., 2017; Dumont et al., 2019). The literature paid large attention to the effects of regulation, supervision and monetary policy on banks (Favara and Imbs, 2015; Baker and Wurgler, 2015). In addition, research on the role of technology used in the fintech sector and the implications for banks is growing (e.g. Anagnostopoulos, 2018; Buchak et al., 2024; Jakšič and Marinč, 2019; Stulz, 2019; Fu and Mishra, 2022; Zhou, 2022). However, only little attention has been paid to the branch network of banks (see e.g. He and Yeung, 2011; Barbieri et al., 2021; Galardo et al., 2021). To partially fill this gap in the literature, this paper answers the following question: Does a growing fintech market share affect the number and distribution of bank branches?

Understanding fintech mortgage lending and its repercussions on traditional banks is crucial for several reasons: First, facilitating home ownership for large groups of the population has constantly been a political goal and subject to great governmental effort.[2] Closely keeping track of the development of the mortgage market helps to monitor and adjust policies accordingly. Second, understanding the dynamics of branch networks not only helps to identify banking deserts (Richardson et al., 2017) but also to locate regions at risk of becoming such. Physical distance to bank branches is important in terms of financial inclusion. Especially the facilitation of small business lending (Petersen and Rajan, 1994, 2002; Bellucci et al., 2013, 2019; Amberg and Becker, 2024), access to credit (Nguyen, 2019; Kärnä et al., 2021) and deposit-taking (Drechsler et al., 2017) are heavily depending on local branches and the exchange of soft information (Agarwal and Hauswald, 2010). In addition, access to finance is crucial for supporting upward economic mobility (Demirgüç-Kunt and Levine, 2009; Banerjee et al., 2015; Demirgüç-Kunt and Singer, 2017) and wealth accumulation (Célerier and Matray, 2019). According to several studies, minorities and low-income communities are less likely to be fintech customers (Ergungor, 2010; Fuster et al., 2019; Friedline et al., 2020). Finally, the research question touches two channels of monetary policy: Fintech lenders operate on large spatial scales and increase the propensity to refinance

[2] A large number of acts and laws have been established to reach and sustain these goals, for example the Dodd-Frank Act, the Community Reinvestment Act (CRA), the Home Mortgage Disclosure Act (HMDA) or the Federal Deposit Insurance Act (FDIC).

(Uluc and Wieladek, 2018; Beraja et al., 2019; Fuster et al., 2019). Thereby they potentially change the transmission of shocks in monetary policy to households (Zhou, 2022). Moreover, branch networks give banks market power in retail deposits, which allows them to borrow at low rates, insensitive to rate shocks. This attenuates the interest rate risk resulting from their large maturity mismatch (Drechsler et al., 2021).

This paper uses a panel data set of all US-mortgages under the Home Mortgage Disclosure Act (HMDA) between 2010 and 2017 combined with the locations of all bank branches reported to the Federal Deposit Insurance Corporation (FDIC) between 2012 and 2019. After describing the distribution and development of physical bank branches and fintech market shares in the US, I estimate a long-difference regression as well as a panel data model to establish a relationship between rising fintech shares and the decreasing number of bank branches. I control for economic, demographic and housing-market-specific covariates.

Several papers identify technology as the main driver behind the rise of the fintech industry. However, in times of tightened regulation, increasing fintech market shares might not be the only reason for banks to restructure their branch network. To discuss reverse correlation, I implement two strategies. First, I estimate structural breaks in the fintech market share on county level. The development of bank branches does not reveal any shocks right before the structural break points in the fintech market shares. Second, I use a geographically purged version of Facebook's Social Connectedness Index to Wayne County, headquarter location of Rocket Mortgage, as instrumental variable for the fintech market share. The main argument is that due to the size and economic importance of Rocket Mortgage for Detroit, social connectedness is positively correlated with and hence relevant for a county's rise in the fintech market share. At the same time it should not have any influence on the existence of physical bank branches.

The results show that a surge in fintech penetration is associated with a decrease in the number of brick and mortar bank branches. The panel data approach and the instrumental variable approach confirm this significantly negative relationship. At the mean, a county's fintech share increases by about 7.7% from 2010 to 2017 which results in the loss of one local bank branch. For counties in the top 90^{th} percentile of the change in fintech market share the loss ranges between two and six physical branches, which exceeds two standard deviations of the change in branches for these counties. The paper shows that not only regulatory pressure but to a large part a lack of innovation in technology causes banks to lose market share and to downsize their branch networks. At the mean, around 45% of all branch closures can be attributed to the rising fintech segment. Mortgages are highly standardized products. In principle, fintech lenders offer the same products as banks at similar or even slightly higher prices using superior technology (Fuster et al., 2019; Buchak et al., 2018). Therefore, the main reason for consumers preferring online mortgages is most likely convenience. Results are confirmed by a Spatial Durbin Model (SDM) that tailors the level of the analysis to the

operational area of a bank branch. This model takes spatial lags into account, which allow for endogenous spatial spillovers of the dependent as well as the independent variables.
The banking sector in the USA is heterogeneous. Closures in response to the rise of fintech lenders are more pronounced in metro areas. I also show that they are driven by a decrease of branches of small banks (< 10bn USD in assets). When banks shut down branches, they do not only exit from the local mortgage market but also stop providing basic bank services like deposit taking, provision of checking accounts, small loans and transactional services. These implications for the local population are measurable: Deposits are significantly lower in the aftermath of rising fintech shares.

The paper contributes to two strands of the literature. It adds to the growing body of research on fintech lenders and their economic impact. The literature categorizes lenders into traditional banks, non-banks (or shadow banks), and fintechs. Traditional banks accept deposits and keep mortgages on their balance sheets, while non-banks do not take deposits but also retain mortgages. Fintech lenders, on the other hand, do not accept deposits and typically follow an originate-to-distribute model, selling most of their mortgages to government-sponsored enterprises (GSEs) like Fannie Mae and Freddie Mac. Two key studies examine the growth of fintech lenders using HMDA data: Buchak et al. (2018) require a fintech lender to have "a strong online presence" and the loan process to take place "online with no human involvement from the lender". Fuster et al. (2019) define a lender as a fintech "if they enable a mortgage applicant to obtain a pre-approval online".
Buchak et al. (2018) identify two factors driving the growth of non-banks and fintech lenders: regulatory pressure on traditional banks and fintechs' technological advantage. While non-banks often substitute for traditional banks in areas with increased regulatory burdens, fintechs operate differently. They originate more refinancing loans and serve fewer first-time buyers, often charging slightly higher rates. This suggests that they attract customers through better technology and a more convenient process rather than targeting riskier or undeserved segments. Their quantitative model indicates that 90% of the fintech market share growth is driven by technology. Stylianou et al. (2023) further confirm that over half of fintech growth results from technology and the superior quality of their mortgage offerings.
Fuster et al. (2019) focus on the differences in processing speed of fintech mortgages as a primary channel of fintech market expansion. They find a reduction of ten days (20%) of the average processing time, while fintechs do not serve riskier borrowers. In addition, low access to finance or the presence of highly constrained borrowers are no determinants of the rising fintech market share in mortgages. This channel is very similar to the effect of technology emphasized by Buchak et al. (2018).
In contrast, Jagtiani et al. (2021) examine mortgages from 2016 and 2017, finding that fintech lenders are concentrated in areas with higher rates of previous mortgage denials and lower credit scores, suggesting fintechs are gaining market share by serving underbanked regions,

including non-metropolitan areas. Other studies explore fintechs' impact on traditional banking. Le et al. (2021) find that increased fintech lending boosts traditional banks' efficiency, while Boualam and Yoo (2022) theorize that fintechs can crowd out traditional banks by reducing their profitability. Similarly, Jiang et al. (2022) link 3G rollouts to branch closures, though they cannot fully attribute this to fintechs. Erel and Liebersohn (2020) and Gopal and Schnabl (2022) find no crowding-out effect, instead noting an expansion of financial services through fintechs in programs like the Paycheck Protection Program and small business lending.

The paper also enlarges the literature on the determinants of bank branches and their reaction to increasing competition. It combines the rise of fintech lenders in the mortgage market and the number of branches on county level. Literature mostly focuses on banks only, which neglects the local dimension of branches. Even in the digital age, bank branches continue to play a significant role, particularly in lending decisions, where they help convey soft information about borrowers (Petersen and Rajan, 2002; Agarwal and Hauswald, 2010; Bellucci et al., 2019). Proximity to a bank can also influence loan amounts and interest rates (Drechsler et al., 2017; Bonfim et al., 2021; Kärnä et al., 2021; Rehbein and Rother, 2024). While some research explores the location of banks in general (e.g. He and Yeung, 2011), fewer studies focus specifically on branch locations, often examining single countries in both developed and developing countries (Ansong et al., 2015; Barbieri et al., 2021; Galardo et al., 2021; Zhang et al., 2021). A common finding is that opening branches in rural and underserved areas could improve financial inclusion, whereas closures in developed countries often lead to financial exclusion. For instance, Aversa et al. (2022) find that bank closures in Toronto primarily affect neighborhoods with high debt risk, minority populations, and low equity scores.

Local market competition is a key factor in branch location decisions. Studies frequently discuss the impact of competition on banking and small business lending. Petersen and Rajan (1995) highlight that lending relationships are more prominent in concentrated markets, while Boot and Thakor (2000) observe more relationship loans with increased interbank competition.[3] Complementing this business-focused literature, research also shows that bank branches positively affect household wealth. For example, Célerier and Matray (2019) demonstrate that an increase in the number of branches leads to greater household welfare, supporting evidence from studies on financial inclusion in developing countries (e.g. Schaner, 2018; Ndlovu and Toerien, 2020).

Closest to this paper is a study Yuan et al. (2023) who also explore the link between fintech

[3]More detailed, Degryse and Ongena (2007) observe no unambiguously detrimental effect of increased competition. They show a non-monotonic u-shaped influence of market concentration on relationship lending which Presbitero and Zazzaro (2011) further describe as dependent on the type of competition that is prevalent in the market. Their results point towards a mitigating effect of new large and external competitors on relationship lending, while the entry of non-hierarchical banks can increase an already prevailing focus on relational lending.

expansion and branch closures. However, they focus on Chinese bank branches, using data from the largest Chinese fintech company. They find significant debranching effects for non-state-owned banks in segments with product overlaps between banks and fintech lenders. Due to their business-model, lending activity largely focuses on loans qualified to be secured by GSEs, e.g. requiring not to exceed a certain county-time specific threshold of the lending volume. Zhang (2020) leverages the density of borrowers around this threshold to instrument for non-bank growth in the mortgage market, highlighting how the shift towards fintech lending affects bank branch networks. Further Zhang (2020) shows spillover effects on small business loans and on deposits, focusing on the largest 100 banks by national deposits.

2.2 Data and Descriptive Statistics

2.2.1 Data

Mortgage data from the HMDA, categorized into banks and fintech lenders, is used to compute the fintech market share in a county's mortgage market. In addition, bank branches in the US are identified by location on the county level and combined with a number of economic, housing market related and demographic control variables. Data on the fintech market shares spans the time period from 2010-2017, as identified by Buchak et al. (2018) and Fuster et al. (2019), who classify all lenders in the HMDA data. As banks need time to adjust their branch network, I observe branches between 2012 and 2019 (after the Great Financial Crisis and before the pandemic). I exclude the counties located in the state of Delaware since it is known as a tax haven and shows an exceptionally high level of deposits. Further I drop Alaska and Hawaii due to their remote location and a lack of controls, as well as some counties due to boundary changes between 2010 and 2017. The main specification covers $2,907$ US counties from 48 US states.

Fintech Market Shares: The fintech market share in the mortgage market of a county is calculated using data from the HMDA reports for nationwide mortgage originations in the US, available at the Consumer Financial Protection Bureau (CFPB). The data includes the loan amount, loan purpose (purchase, refinance, improvement), an identifier for the lender as well as some borrower-specific information such as gender and race. It also categorizes loans as conventional, FHA-insured or guaranteed by the Office of Veterans Affairs or Rural Housing Services. While the data also includes pre-approvals, I concentrate on loans that are denoted as finally originated only.

Bank Branch Locations: The Federal Deposit Insurance Corporation (FDIC) publishes an Institution Directory with identifiers and exact locations of all FDIC insured bank branches and their deposits. This directory also matches the branch to the headquarter of the institution and its assets. I exclude all branches that offer limited services (e.g. withdrawals or deposit taking only), merge the data to the HMDA data set and combine the categorization of Buchak et al. (2018) and Fuster et al. (2019) to distinguish traditional banks from fintech lenders. This allows to compute the market share for fintech lenders on county level and the number of bank branches, institutes and deposits per county.

Control Variables: The Social Determinants of Health Database (SOD) provides a large number of socioeconomic variables on county level. Mostly stemming from the American Community Survey (ACS), I use information on income, education, gender, race and age composition of the counties as well as the median home value. To control for unemployment on the county level, I gather data from the U.S. Department of Agriculture. The same source is used to distinguish between metropolitan and non-metropolitan areas of different population size using 2013 Rural-Urban Continuum Codes. The data is complemented by population per county from the United states Census Bureau and local GDP from the Bureau of Economic Analysis.[4] Building permits are from the Building Permits Survey (BPS). Definitions for all variables can be found in Table A2.1.

2.2.2 Descriptive Statistics

Descriptive statistics are presented in Table 2.1. At the median, there are 10 branches in a county. Since counties are very heterogeneous in population, size and many other aspects, standard deviation is large (74.16). 564 observations from 97 unique counties show one bank branch only. Los Angeles County has by far the largest number of bank branches (1785 in 2014).[5]

Figure 2.1 illustrates the number of branches by year. The number of bank branches declines year by year with an increasing magnitude. In 2010, there are 93,310 branches. Over the entire panel, this number drops by 11.8% to 82,299 in 2019. The loss is almost equally distributed between metropolitan and non-metro areas. The general trend in the sample is comparable to Richardson et al. (2017), who explore bank branch closure in the US from 2008-2016[6] and a trend which has been observed longer than that.

Panel A of Figure 2.2 shows the average number of bank branches between 2012 and 2019. Light shaded regions have a low number of branches while darker shades represent a high number of bank branches. Densely populated areas such as California, Arizona, Florida and the Northeast show high numbers of bank branches, while the rural states such as Montana,

[4] Due to missing GDP data on county level, some counties in Virginia are excluded.

[5] Three other counties stand out: Cook County with its county seat Chicago (1563 branches in 2010), Harris County in Texas, hosting the city of Houston (1024 branches in 2010) an Maricopa County in Arizona including Phoenix (878 branches in 2010).

[6] In their report, they use ZIP codes to locate the branches, while I refer to state and county FIPS codes. Since some observations are excluded as described in the Data section, my numbers are slightly lower than those of Richardson et al. (2017).

Wyoming or Kansas have a low number of branches (see Figure A2.1 for the US population density). Panel B presents the total change in bank branches from 2012 to 2019. Patterns match each other, which shows that densely populated metropolitan areas with the largest absolute number of bank branches experienced the largest absolute loss.

Development of Fintech Market Shares: At the median, a county's fintech market share equals 5% with a standard deviation of 4%. The maximum market share is as large as 78%, while some counties also exhibit zero loans by fintech lenders. The median share is much higher for refinancings (8%) than for purchases (3%). By construction, descriptive figures of the fintech market share by Fuster et al. (2019) are almost identical. Figure 2.3 shows the development of the fintech market share over time. Starting off with 2.8% in 2010 the fintech market share increases to 10.4% in 2017.[7] It increases very stably over time with only a little slow down in growth in 2014. The decomposition into purchases and refinancings shows that this can be attributed to a small decrease in the fintech share of refinancings in that year. Despite that, refinancings show the largest increase over the whole sample. In 2010, roughly 4% of refinancing loans have been originated by fintech lenders. This share quadruples until 2017 which makes refinancings the most strongly growing market segment for fintech lenders. Purchases account for a lower proportion of the fintech market share, starting with less than 1% in 2010, the share increased to 8.1% in 2017. This is in line with findings of Buchak et al. (2018) and Jagtiani et al. (2021). In addition, Fuster et al. (2019) show that propensity to refinance is higher in counties with high fintech shares. They point out that one channel through which fintech lenders are expanding is by picking up unexploited potential in the refinancing sector. According to Jagtiani et al. (2021), refinancing loans are only slightly smaller in terms of the loan amount. At the same time they require less screening while being more suitable for automation. A comparable development has already been identified by Bennett et al. (2001) in the early 1990s. Figure 2.4 shows the change in fintech market share between 2010 and 2017. Darker shading corresponds to a higher change. Less than 5% of counties exhibit a decrease in the fintech market share, the mean change amounts to an increase of 8% with a standard deviation of 5%.

Bank Branches and Fintech Market Shares: Figure 2.5 displays the relationship of the number of bank branches and the fintech market share in the mortgage market. Panel A presents the full picture, using all observations, while Panel B focuses on the counties with less than 100 branches and bins the observations. Both Panels reveal that a higher fintech market share is associated with less bank branches per county. In Panel A one can clearly identify four outstanding counties in terms of the number of bank branches. Los Angeles County (orange), Cook County (green), Harris County (purple) and Maricopa County (red)

[7] The share from Fuster et al. (2019) is lower in 2017, most likely due to the fact that there data set already ends in mid 2017.

represent the four single groups to the right of the x-axis. Panel B zooms into counties with less than 100 branches and bins the observations from these counties, confirming that counties with more bank branches tend to have a lower fintech share.

2.3 Empirical Analysis

2.3.1 Baseline Estimation

This paper explores the link between the rising fintech share in the US residential mortgage market and the number of bank branches across the US. To do so, the following long-difference specification is estimated:

$$\Delta log(Branches)_{i,[2012,2019]} = \beta_0 + \beta_1 \times \Delta FtShare_{i,[2010,2017]} + \beta_2 \times \Delta G_{i,[2010,2017]} + \epsilon_i \quad (2.1)$$

For all counties i, $\Delta log(Branches)_{i,[2012,2019]}$ describes the change in the logarithmized number of bank branches from 2012 to 2019. I assume that banks need time to adjust their branch network when the market changes. Therefore a lag of two years is taken into account. The main explanatory variable of interest $\Delta FtShare_{i,[2010,2017]}$ represents the change in the fintech share between 2010 and 2017. The vector $\Delta G_{i,[2010,2017]}$ includes a set of control variables that also enter in changes between 2010 and 2017. The log of the local GDP, ratio of unemployment, control for economic factors. Demographic factors such as the log of total population and the log of household income, the share of people older than 64 years, the share of people holding a bachelor degree, the share of people reporting being white race, the share of people black or of African American race and the share of females might also influence bank branch locations. The log of the median home value and the log of building permits capture housing market characteristics. Standard errors are clustered on metro level. Descriptive statistics for all controls are presented in Section 2.2 and in Table 2.1.

Table 2.2 presents the results using the overall fintech market share in Column (1), the fintech market share in purchases in Column (2), for refinancings in Column (3) and for a comparison the fintech market share as computed by Fuster et al. (2019) in Column (4). All specifications show a negative relationship between the change in the fintech market share and the change in bank branches. This suggests that increasing fintech market shares are correlated with a decrease in local bank branches. The relationship is statistically significant with the exception of the specification that uses mortgages from purchases only. Referring to the overall fintech market share in Column (1), a rise of one standard deviation (5.38%) in a county's fintech share results in the loss of approximately 1 local bank branch. What sounds negligible at first has an economically significant impact in two respects: First, the share of branch closures resulting from the rise of fintech lenders is substantial compared to the total number of closures. The average county records a loss of 3.2 bank branches between 2012 and 2019. On average, almost 45% of this change can be attributed to the rise of fintech lenders in the mortgage market. Second, the sample covers the rise of fintech lenders until 2017 only. Figure 2.3 indicates an exponential growth of fintech lenders. Significantly higher market shares do not seem unrealistic or a long way off. For counties in the top 90^{th} percentile of the change in fintech market share, the results already imply a loss of at least 2.5 branches which equals almost 10%. The relationship is statistically insignificant and much smaller for the fintech share in purchases in Column (2). However, when using the fintech share in refinancings in Column (3), the coefficient doubles and is highly statistically significant. This is in line with Fuster et al. (2019) who discover a higher propensity to refinance in counties with higher fintech share. Column (4) confirms the results by using their data. There is no research on the effects of fintech lenders on the number of bank branches to compare this number with. Nevertheless, the results seem plausible, as Stylianou et al. (2023) estimate that changes in consumer preferences and technological margin explain more than half of the rise in fintech market shares and 40% of the decline in the bank market share. As a second approach, a panel model is estimated to leverage both cross-sectional and time-series variations. It controls for unobserved heterogeneity by including county and year fixed effects.

$$log(Branches_{i,t+2}) = \beta_0 + \beta_1 \times FtShare_{i,t} + \beta_2 \times G_{i,t} + \delta_i + \delta_t + \epsilon_{i,t} \tag{2.2}$$

Equation 2.2 is estimated using an OLS fixed effects panel-model estimator. The set of controls remains unchanged. To account for the fact that banks need a certain amount of time to react to changing market structures and adjust their branch network, the dependent variables again enters the model with a lag of two years. δ_i and δ_t capture county and year fixed effects, absorbing all county-specific as well as year-specific common shocks. Since the general time trend from 2010 to 2017 shows a strong decrease in branches, fixed effects are expected to already explain a large fraction of variation in the data. Standard errors are clustered at metro level.

Table 2.3 shows the results. Column (1) confirms a negative and significant relationship between the fintech market share and the number of bank branches. Column (2) replicates this finding using the fintech market share data from Fuster et al. (2019). Columns (3) and (4) use a Pseudo Poisson Maximum Likelihood (PPML) estimator based on the fact that the data on bank branches is count data with a highly right-skewed distribution. The results are again comparable.

2.3.2 Robustness

Compared to the operating area of a bank branch, US counties are relatively large spatial units. Some large cities, e.g. St. Louis (Missouri), Cincinnati (Ohio), Kansas City (Missouri) or Boston (Massachusetts) are divided in multiple counties. This can be a threat to the results, in case branches move or operate in different counties. To ensure that the results are robust across the spatial dimensions of counties, I estimate a a Spatial Durbin Model (SDM) on census tract level for the state of Illinois. A census tract in the US is defined as an area with around $4,000$ inhabitants and a relatively homogeneous distribution regarding the socio-economic characteristics. It therefore more likely captures the operating area of a bank branch. Still, on the demand side, many tracts that exhibit fintech loans do not host a bank branch. Borrowers might be located in neighboring tracts to those hosting a physical bank branch. People also locally move without changing their bank branch. It is reasonable to assume that the fintech shares of neighboring tracts (as well as all other covariates) affect the number of bank branches in a tract. On the supply side, operating a branch in one tract might be sufficient to cover demand of neighboring tracts as well. The number of branches in a tract is also dependent on the number of neighboring branches. The SDM incorporates spatial lags of both the dependent variable and the independent variables into the regression model (see Section A2.1 for details). The results are presented in Table A2.4 and confirm the baseline findings. Increase of the fintech market share in the own tract as well as the adjacent tracts are associated with a decrease in the number of branches.

Figure 2.5 shows that four counties stand out with an exceptionally high number of branches. At the same time some counties exhibit very large fintech shares. To rule out that these counties drive the results, I exclude all counties with more than 700 branches or a fintech market share larger than 30%. Table A2.5 presents the results for estimating Equation 2.1, Table A2.6 the results for estimating Equation 2.2. Both tables almost match the respective baseline findings. In addition, Table A2.7 confirms the results by re-estimating Regression 2.1 with the log of bank branches per $10,000$ people as dependent variable.

Stylianou et al. (2023) present a model of imperfect competition where banks, non-banks and fintechs compete in the loan market. Next to increasing market concentration within traditional banks, they observe a decline in banks' quality which they attribute to less people preferring branch-based interaction and turning to more convenient services. According to the Annual Retail E-Commerce Sales Statistics from the US Census Bureau, the share of e-commerce has developed very similarly to the fintech mortgage market. Dolfen et al. (2023) assess the benefits of e-commerce to consumers and distinguish between gains from an increased variety and gains from convenience. Mortgages are highly standardized products with little variation in product characteristics between providers, especially in the case of conforming loans that dominate the fintech market. Therefore product variety can be ruled out as the main driver for the rise of fintech lenders. Thus, one can think about the changes in the fintech market share as a realization of changes in customer preferences. People value the convenience offered by the fintech lenders' advanced technology. Following Buchak et al. (2018), this leads to customers accepting marginally higher rates. To underline this interpretation, I look at other business segments that are expected to profit from this change in customer preference. Equation 2.2 is estimated with the number of McDonald's branches from 2010 to 2017 in Column (1) and the change in overall fast food restaurants[8] in Column (2) of Table A2.8. Fast food restaurants, particularly McDonald's, are known for offering highly standardized menus and fast processing times. They generate a large part of their sales through take-away and drive-thru options which save people time. Jekanowski et al. (2001) describe the important role of convenience in the business strategy of fast food restaurants (explicitly McDonald's) and explain its significance in the context of consumer preference theory. The results in Table A2.8 show a positive and significant relationship of rising fintech shares and the number of restaurants. Since restaurants are not subject to strict regulation such as banks and fintechs and fast food chains are entirely different industries, there are less concerns regarding reverse correlation.

2.4 Identification

The consolidation of traditional banks' brick and mortar branches is not solely driven by the rise of the fintech industry. The results from the previous section suggest two interpretations. First, fintech lenders may fill the gap when traditional banks exit the market. This negative relationship between tightened regulation and the number of banks and branches is not new; it dates back to the late 19$^{\text{th}}$ century, as shown by Jaremski (2013). Similarly, Houston et al. (2012) find that banks often shift assets and establish subsidiaries or branches in countries with more lenient regulations. Alternatively, fintech lenders may gain market share, leading to a crowding-out effect on traditional branches. If fintechs offer better quality services, such as more convenient loan origination processes, customers may increasingly prefer them over traditional banks, resulting in decreased lending volumes for bank branches. Buchak et al. (2018) also examine these two different mechanisms in a structural model. They show that the two effects complement each other and that there is no simple one-directional story. However, their model also shows that "technology alone is responsible for approximately 90% of gains of fintech firms". Fuster et al. (2019) reinforce this by showing that slow processing speed of traditional banks is one of the drivers for the rise of fintech lenders. To this regard differences in processing speed are clearly a technological factor. Fuster et al. (2019) do not find a relationship between low access to finance and fintech borrowing either but report that fintech market share is increasing in bank branch density. The literature therefore hints towards a crowding-out of traditional bank branches by fintech lenders.

[8] Data taken from Food Environment Atlas. Data is available for 2011 and 2016 only. All covariates are adjusted to this time period. No time lag is applied in this specification.

2.4.1 Structural Break Points

This subsection presents descriptive evidence that suggests a potential causal effect of rising fintech market shares on the number of bank branches. Nevertheless, I remain cautious in using causal interpretations throughout the paper. Charles et al. (2019) use sudden changes in housing prices as an instrumental variable for speculative activity in housing markets. Their assumption is that these sharp surges are independent of the underlying market fundamentals. I follow Zhou (2022) and use a similar strategy by estimating the following regression for each county:

$$FtShare_{i,t} = \gamma_c + \tau_c t + \lambda_c \left(t - t_c^*\right) \mathbb{1}\left(t \geq t_c^*\right) + \varepsilon_{c,t} \tag{2.3}$$

The goal is to find a single structural break point t_c^* that maximizes the R^2 of the regression. This structural break point can occur between 2011 and 2016. λ_c represents the size of the structural break and τ_c is a linear time trend of each county's time series of fintech market shares. This procedure identifies a statistically significant (p-value $\leq$ 0.05) structural break point for $1,075$ counties. Figure 2.6 plots the fintech market share and the number of bank branches around the structural break points. Note that the graph is not based on a constant sample due to the aggregation. Particularly at the margins there may be a small sample bias. I will therefore limit the interpretation to the general trend. The blue line shows a moderate upward trend in the fintech share up until the structural break points. At this point a sharp increase from 5.5% to 9.3% occurs. Afterwards the fintech market share is relatively stable. The dashed red line represents the number of bank branches. It shows an almost linear and slightly negative trend, from four years before to three years after the structural break in the fintech market shares. This indicates that no sudden drops in local bank branches triggered the rise in fintech penetration, at least not in those counties that exhibit a structural break in fintech market shares and are therefore most likely affected by reverse causality.
Equation 2.3 is re-estimated for each county separately using the number of bank branches as dependent variable and allowing for structural breaks between 2011 and 2018. A statistically significant structural break point is identified for 1,196 counties. For those 441 counties which exhibit break points for both metrics, I deduct the break point year of bank branches from the break point year of the fintech market share. Panel B of Figure 2.6 plots the frequencies of this difference. Overall, the distribution is shifted to the right. Almost 20% of county pairs show a structural break in the number of bank branches one year after the break in the fintech market share. While for roughly 15% of all counties with breaks in variables there is no time difference between the structural breaks, in almost 50% of the counties the break occurs first in the fintech market and second in the branch network. This analysis again lacks evidence that reverse correlation is driving the results.

2.4.2 Instrumental Variable Approach

While the previous section aims to rule out reverse correlation, the following instrumental variable approach carefully attempts to establish a causal relationship. To do so, I estimate Equation 2.1 using a geographically purged version of Facebook's Social Connectedness Index (SCI) to Wayne County (Michigan) as an instrument for the fintech market share. Rocket Mortgage, the largest fintech lender in the residential mortgage market, has its headquarter located in Detroit, Wayne County. The fintech company, formerly known as Quicken Loans, is a big player in Wayne County's economic landscape, employing over $17,000$ people in the city (King, 2017). This amounts to more than 2.5% of Detroit's population in 2017 and makes Rocket Mortgage the largest employer and taxpayer of Detroit. Even outside the business world, the fintech lender is permanently present in the city. Between 2011 and 2016, the company bought 62 properties in downtown Detroit, spending over 450 million USD according to Aguilar (2016). The local news outlet further states that Rocket Mortgage's founder "[Dan] Gilbert is so closely identified with the downtown business district that the phrase 'Gilbertville' has been used". Other sources describe Wayne County as an emerging center for technology-driven sectors and attribute a pioneering role to Rocket Mortgage in the last decade (Hirsch, 2021; World Economic Forum, 2023).

The Social Connectedness Index uses friendship data from Facebook and describes real-world social ties between counties in the US (Kuchler et al., 2022b). A detailed description of this data including important correlations at the US county level can be found in Bailey et al. (2018). I argue that Rocket Mortgage's significant economic and social footprint in Detroit as well as Detroit's development as hub for technology-driven sectors propagates via social ties. Several studies show similar effects for product adoption (Bailey et al., 2022), investment decisions (Kuchler et al., 2022a) and most importantly mortgage lender choices (Zhou, 2022) and house price beliefs (Bailey et al., 2019). The SCI to Wayne County is therefore expected to show a positive relationship with local fintech market shares. Higher exposure to Wayne County and thereby the largest fintech lender in the US as well as a vibrant high-tech sector should be positively correlated with the connected counties' fintech adoption. At the same time, it should not have any influence on the number of bank branches aside from the impact via fintech market shares. The SCI is highly correlated with distance and other geographic factors. To isolate pure social connectedness I purge the SCI, regressing it on the log of distance to Detroit, dummies for being in the same state, commuting-zone or sharing a border as well as a dummy that indicates whether a county is located in a metropolitan area.[9] The residuals of this regression represent the instrumental variable for the change in the fintech market share. Both the SCI and the geographically purged SCI are illustrated in Figure 2.7. Especially in southern counties along the Mississippi social connectedness to Wayne County is high. This is most likely driven by migratory movements

[9] Table A2.9 presents the results of this regression.

such as the Great Migration in the early last century (see Bailey et al. (2018) for a detailed discussion). Table 2.4 presents the results of the IV approach. Column (1) shows strong results for the first stage (F-values > 75), indicating a significant positive effect of the SCI on the change in the fintech market share. This confirms the assumption that the relevance condition holds. Counties with a larger social connectedness to Wayne County exhibit higher fintech shares. Column (2) presents the second stage results, showing a negative effect of the change in fintech market share on the number of bank branches.

2.5 Heterogeneity and Effects of Branch Closures

Banks and bank branch networks are very heterogeneous in size and structure. In addition, US counties vary to a great degree, especially between rural and urban areas. I use differences along these lines to explore how competition between banks and fintechs affect branch networks and bank fundamentals. The most intuitive way to start is by identifying counties within metropolitan statistical areas ("metro counties") and compare them to rural counties. The US comprises more than 900 metropolitan statistical areas that have a high population density and strong economic bonds. I interact the fintech market share with a dummy equal to one for metro counties. Table 2.5 presents the results from estimating Equation 2.1. Column (1) displays the results for the overall fintech share, Columns (2) and (3) use the fintech share in purchases and refinancings only. Column (4) uses the fintech market share as computed by Fuster et al. (2019). The baseline effect is attenuated and shrinks by 30% for non-metro counties. For the case of metro counties, it almost triples and the interaction term is highly statistically significant and almost constant throughout all columns. For non-metro counties there is almost no effect for purchases, but equally large coefficients for the overall fintech share and refinancings only. Spatial heterogeneity of the results therefore suggests that the relationship between rising fintech market shares and branch closures is not concentrated on rural areas where the last bank branch might close and people are forced to use online lenders. The likelihood that another branch, sometimes even of the same bank, is quickly accessible is much higher in metro counties compared to non-metro counties. Thus, the spatial heterogeneity points towards customers that prefer convenient fintech products which leads to a restructuring in the banks' branch network.

Banks are heterogeneous in size. These differences propagate to their business models: Small banks often concentrate on relationship lending with local branches covering small areas that offer an in-depth service in all areas of personal finance. Large banks engage in transactional lending, have lighter branch networks and outsource many activities to centralized service units. This suggests different results for small and large institutes. I define small banks as having less than 10bn USD in assets and large banks as having more than 50bn USD in assets. Table 2.6 shows the results for using the number of branches from small banks in Column (1) and the share of branches from small banks in Column (2). Columns (3) and (4) repeat the analysis for large banks. Since small banks are more active in rural areas, the interaction of the fintech share and the dummy for metro counties is again included. This ensures that the results are not driven by a correlation of the small bank share and rural characteristics. Column (1) shows that the decrease in branches is more pronounced for small banks. The interaction term is also negative and statistically significant again which means that the decline is even stronger in metro counties. The number of observations is only marginally smaller compared to the full sample. Only about 3% of the entire set of counties do not host a branch from a small bank. The findings are therefore not driven by the sample composition. The share of small branches in Column (2) is also negatively correlated to increasing fintech market shares. This ensures that the decline is not mechanical and mirrored by large banks. Columns (3) and (4) confirm this. Both specifications yield coefficients close to zero and suggest that the branch network of large banks is not changing in the context of rising fintechs. Small institutes rely heavily on relationship lending and their comparative advantages in using soft information which are hardly transferable between branches or locations (Berger et al., 2005; Berger and Black, 2011; Uchida et al., 2012; Cenni et al., 2015). On the contrary, large banks are more active in transactional lending, relying on hard information. Several studies suggest that branches leveraging long-term relationships between borrower and lender are more costly to operate (Kovner et al., 2014; Bolton et al., 2016). The branch networks of large banks are often already highly optimized and designed to cover many customers with locally based employees. Due to their size they might also smooth losses and can sustain branches at the margin of profitability longer than small banks. Residential mortgage lending also most likely accounts for a larger proportion of the overall assets for small banks what makes them more vulnerable towards competition from fintechs.

If bank branches are closed, this also has implications for other product segments than mortgage lending. Especially small banks offer a variety of products including bank transfers, insurances and deposit-taking. This leads to two potential spillover-effects from increased fintech lending. First, local access to finance might become harder and more costly. Several papers show that access to a bank branch increases financial inclusion and decreases redlining by leveraging soft information. The role of distance, social ties and relationships has been emphasized by many studies (Berger and Black, 2011; Bellucci et al., 2013; Rehbein and Rother, 2024; Amberg and Becker, 2024). Sakong and Zentefis (2022) show, by using geolocated mobile device data, that low-income and black households tend to visit branches less often than high-income and white households. This pattern is explained by larger distances to bank branches and goes as far as an entire drop-off in branch use. Ergungor (2010) show that mortgage originations in low- and moderate-income regions increase while interest rate spreads decrease if there exists a physical bank branch. Second and in light of the increased magnitude in metro counties more important, branch closures might impact the balance sheets of banks. Table 2.7 tests both aspects. Column (1) tests whether the number of institutes that operate in a county changes. The coefficient is statistically insignificant and relatively small. This suggests that banks do not exit the market and might migrate their customer relationships to different branches. Column (2) uses the average change in nationwide assets of all banks that operate a branch in the respective country as a dependent variable. The result is positive but statistically insignificant. Banks do not seem to lose in terms of assets. Since assets are not local, this might be to imprecise to find any effects. Finally, I test whether local deposits react by using the change in the local deposits per county. Deposits can be attributed to single branches and are therefore expected to change if branches close. An insignificant effect would mean that every customer migrates to another branch of the same or a different bank, but remains inside the county. In contrast to that, Column (3) shows a statistically significant and negative coefficient. Lower deposits in the aftermath of increased fintech lending can have several consequences. On the one hand, banks maintain branches to use their market power in retail deposits as a hedge against the interest rate risk on their balance sheet. Drechsler et al. (2021) show that the duration mismatch that a bank faces between short-term deposits they take in and long-term loans they supply makes them vulnerable against interest rate shocks. Banks substitute high costs from interest rate risk by the relatively stable and predictable costs of supplying a branch network. If competition on the residential mortgage market or changing consumer preferences increase the costs of branches since they become less profitable, banks might shift to other strategies to hedge their interest rate risk. On the other hand, monetary policy transmission changes by two channels. First, shocks are passed through differently by fintechs compared to banks. Fintechs increase the propensity to refinance, since these loans are associated with less screening and are easier to automate. This changes the bank lending channel of monetary policy transmission. Second, closing bank branches might have an impact on the deposit channel. Depending on where customers switch to if they choose not to stay at the local bank after a branch closure, the two effects might cancel or reinforce each other.

2.6 Conclusion

This paper explores the relationship between rising fintech market shares and the number of bank branches. Therefore, I combine US mortgage originations from the HMDA with bank branch locations from the FDIC Institution Directory on county level. Fintech lenders are identified following the categorization of Buchak et al. (2018) and Fuster et al. (2019). Using a long-difference specification, I show a negative relationship between the change in the fintech market share between 2010 and 2017 and the change in the number of bank branches between 2012 and 2019. The baseline results suggest that at the mean, a county loses about one bank branch due to the rise in the fintech market share. This accounts for almost 45% of the overall decline in the number of bank branches for the mean US county. Results are confirmed by a panel data approach. Due to the fact that the distribution of bank branches is extremely right-skewed and has large values for some metropolitan counties, I also show that the results remain stable using a Poisson Maximum Likelihood estimator. To tailor the analysis to the operating area of a bank branch I estimate a Spatial Durbin Model on census tract level for the state of Illinois, allowing for endogenous spatial spillovers between adjacent tracts in terms of the dependent and independent variables. Evaluating the results with a focus on the spatial spillovers of the fintech market share confirms the previous findings.

Previous studies argue that the largest part in the rise of fintechs can be attributed to their advanced lending technology (Buchak et al., 2024; Fuster et al., 2019; Stylianou et al., 2023). Although this points towards a crowding out-effect, the closure of traditional brick and mortar branches can also be triggered by increasing regulatory requirements. To discuss these concerns regarding reverse correlation, I exploit the timing between rising fintech market shares and branch closures. I estimate structural break points in the times series of each counties' fintech market share. Using the structural break points, I show that there are no shocks in the branch network right before the fintech market share increases. This is complemented by an instrumental variable approach. Facebook's Social Connectedness Index (SCI) to Detroit, home of the largest fintech company (Quicken Loan) is used as instrumental variable for the change in the fintech market share. I purge the SCI by geographic factors before using it as an IV to extract pure social connectedness. Rocket Mortgage's significant economic and social footprint in Detroit as well as Detroit's development as a hub for technology-driven sectors propagates via social ties which leads to a large correlation between the purged social connectedness to Wayne County and fintech penetration. The results show that a rising fintech market share is associated with a decline in traditional bank branches.

The effect is mostly driven by metro counties and changes in the branch networks of small banks. Small banks rely on long-term relationships between borrower and lender which makes the operation of branches more costly. Relocation of unproductive branches or reorganization of the branch network might be more difficult for smaller banks since they have an advantage in relationship-lending and rely more heavily on personal customer-relationships which makes their production capital more stationary. At the same time, large banks are better at smoothing losses over their large branch network. Higher market shares for fintech lenders are also associated with lower local deposits. This suggests that the closure of bank branches in the aftermath of rising fintech shares has spillover effects on the supply of other banking services as well as banks balance sheets. Next to an impact of the bank lending channel of monetary policy (Zhou, 2022), there seems to be an effect on on the deposit channel (Drechsler et al., 2017). Finally, branch closures can also influence the risk of banks due to reduced market power in retail deposits (Drechsler et al., 2021).

Bibliography

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. The Review of Financial Studies, 23(7): 2757–2788.

Aguilar, L. (2016). Putting a price tag on properties linked to Gilbert. The Detroit News. Accessed: 2024-09-12.

Amberg, N. and Becker, B. (2024). Banking without branches. Working Paper.

Anagnostopoulos, I. (2018). Fintech and regtech: Impact on regulators and banks. Journal of Economics and Business, 100: 7–25.

Ansong, D., Chowa, G., and Adjabeng, B. K. (2015). Spatial analysis of the distribution and determinants of bank branch presence in Ghana. International Journal of Bank Marketing, 33(3): 201–222.

Aversa, J., Shaker, R., Cleave, E., and Salooja, N. (2022). Determinants of bank closures: Exploring the relationship between neighborhood characteristics and bank branch locations. Papers in Applied Geography, 8(4): 393–413.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. Journal of Economic Perspectives, 32(3): 259–280.

Bailey, M., Dávila, E., Kuchler, T., and Stroebel, J. (2019). House price beliefs and mortgage leverage choice. The Review of Economic Studies, 86(6): 2403–2452.

Bailey, M., Johnston, D., Kuchler, T., Stroebel, J., and Wong, A. (2022). Peer effects in product adoption. American Economic Journal: Applied Economics, 14(3): 488–526.

Baker, M. and Wurgler, J. (2015). Do strict capital requirements raise the cost of capital? Bank regulation, capital structure, and the low-risk anomaly. American Economic Review, 105(5): 315–20.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. American Economic Journal: Applied Economics, 7(1): 22–53.

Barbieri, L., Borroni, M., Lippi, A., Piva, M., Rossi, S., et al. (2021). Determinants of bank branch presence in local areas: A comparison between north and south of Italy. International Journal of Economics and Finance, 13(9): 1–15.

Barriere, L., Liu, F., Jo, Y., and Skhirtladze, A. (2020). Data point: 2019 mortgage market activity and trends. Consumer Financial Protection Bureau Office of Research Reports Series, (20-05).

Bellucci, A., Borisov, A., Giombini, G., and Zazzaro, A. (2019). Collateralization and distance. Journal of Banking & Finance, 100: 205–217.

Bellucci, A., Borisov, A., and Zazzaro, A. (2013). Do banks price discriminate spatially? Evidence from small business lending in local credit markets. Journal of Banking & Finance, 37(11): 4183–4197.

Bennett, P., Peach, R., and Peristiani, S. (2001). Structural change in the mortgage market and the propensity to refinance. Journal of Money, Credit, and Banking, 33(4): 955–975.

Beraja, M., Fuster, A., Hurst, E., and Vavra, J. (2019). Regional heterogeneity and the refinancing channel of monetary policy. The Quarterly Journal of Economics, 134(1): 109–183.

Berg, T., Fuster, A., and Puri, M. (2022). Fintech lending. Annual Review of Financial Economics, 14: 187–207.

Berger, A. N. and Black, L. K. (2011). Bank size, lending technologies, and small business finance. Journal of Banking & Finance, 35(3): 724–735.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., and Stein, J. C. (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. Journal of Financial Economics, 76(2): 237–269.

Bolton, P., Freixas, X., Gambacorta, L., and Mistrulli, P. E. (2016). Relationship and transaction lending in a crisis. The Review of Financial Studies, 29(10): 2643–2676.

Bonfim, D., Nogueira, G., and Ongena, S. (2021). "Sorry, we're closed" Bank branch closures, loan pricing, and information asymmetries. Review of Finance, 25(4): 1211–1259.

Boot, A. W. and Thakor, A. V. (2000). Can relationship banking survive competition? The Journal of Finance, 55(2): 679–713.

Boualam, Y. and Yoo, P. (2022). Fintech disruption, banks, and credit (dis-)intermediation: When do foes become friends? Working Paper.

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. Journal of Financial Economics, 130(3): 453–483.

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2024). Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. Journal of Political Economy, 132(2): 616–693.

Célerier, C. and Matray, A. (2019). Bank-branch supply, financial inclusion, and wealth accumulation. The Review of Financial Studies, 32(12): 4767–4809.

Cenni, S., Monferrà, S., Salotti, V., Sangiorgi, M., and Torluccio, G. (2015). Credit rationing and relationship lending. Does firm size matter? Journal of Banking & Finance, 53: 249–265.

Charles, K. K., Hurst, E., and Notowidigdo, M. J. (2019). Housing booms, manufacturing decline and labour market outcomes. The Economic Journal, 129(617): 209–248.

Degryse, H. and Ongena, S. (2007). The impact of competition on bank orientation. Journal of Financial Intermediation, 16(3): 399–424.

Demirgüç-Kunt, A. and Levine, R. (2009). Finance and inequality: Theory and evidence. Annual Review of Financial Economics, 1(1): 287–318.

Demirgüç-Kunt, A. and Singer, D. (2017). Financial inclusion and inclusive growth: A review of recent empirical evidence. Working Paper.

Dolfen, P., Einav, L., Klenow, P. J., Klopack, B., Levin, J. D., Levin, L., and Best, W. (2023). Assessing the gains from e-commerce. American Economic Journal: Macroeconomics, 15(1): 342–370.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The deposits channel of monetary policy. The Quarterly Journal of Economics, 132(4): 1819–1876.

Drechsler, I., Savov, A., and Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk. The Journal of Finance, 76(3): 1091–1143.

Dumont, A. M., Roberts, A., et al. (2019). Perspectives from main street: Bank branch access in rural communities. Working Paper.

Erel, I. and Liebersohn, J. (2020). Does fintech substitute for banks? Evidence from the paycheck protection program. Working Paper.

Ergungor, O. E. (2010). Bank branch presence and access to credit in low-to moderate-income neighborhoods. Journal of Money, Credit and Banking, 42(7): 1321–1349.

Favara, G. and Imbs, J. (2015). Credit supply and the price of housing. American Economic Review, 105(3): 958–92.

Friedline, T., Naraharisetti, S., and Weaver, A. (2020). Digital redlining: Poor rural communities' access to fintech and implications for financial inclusion. Journal of Poverty, 24(5-6): 517–541.

Fu, J. and Mishra, M. (2022). Fintech in the time of covid- 19: Technological adoption during crises. Journal of Financial Intermediation, 50: 100945.

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The role of technology in mortgage lending. The Review of Financial Studies, 32(5): 1854–1899.

Galardo, M., Garrì, I., Mistrulli, P. E., and Revelli, D. (2021). The geography of banking: Evidence from branch closings. Economic Notes, 50(1): e12177.

Gopal, M. and Schnabl, P. (2022). The rise of finance companies and fintech lenders in small business lending. The Review of Financial Studies, 35(11): 4859–4901.

He, C. and Yeung, G. (2011). The locational distribution of foreign banks in China: A disaggregated analysis. Regional Studies, 45(6): 733–754.

Hirsch, L. (2021). Dan gilbert will invest $500 million to help revitalize Detroit. The New York Times. Accessed: 2024-09-12.

Houston, J. F., Lin, C., and Ma, Y. (2012). Regulatory arbitrage and international bank flows. The Journal of Finance, 67(5): 1845–1895.

Jagtiani, J., Lambie-Hanson, L., and Lambie-Hanson, T. (2021). Fintech lending and mortgage credit access. The Journal of FinTech, 1(01): 2050004.

Jakšič, M. and Marinč, M. (2019). Relationship banking and information technology: The role of artificial intelligence and fintech. Risk Management, 21: 1–18.

Jaremski, M. (2013). State banks and the national banking acts: Measuring the response to increased financial regulation, 1860–1870. Journal of Money, Credit and Banking, 45(2-3): 379–399.

Jekanowski, M. D., Binkley, J. K., and Eales, J. (2001). Convenience, accessibility, and the demand for fast food. Journal of Agricultural and Resource Economics, pages 58–74.

Jiang, E. X., Yu, G. Y., and Zhang, J. (2022). Bank competition amid digital disruption: Implications for financial inclusion. Working Paper.

Kärnä, A., Manduchi, A., and Stephan, A. (2021). Distance still matters: Local bank closures and credit availability. International Review of Finance, 21(4): 1503–1510.

King, R. (2017). Quicken loans family of companies surpasses 17k employees in detroit, more to come. DBusiness Magazine – Detroit's Premier Business Journal. Accessed: 2023-08-07.

Kovner, A., Vickery, J. I., and Zhou, L. (2014). Do big banks have lower operating costs? Economic Policy Review, 20(2): 1–27.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). Social proximity to capital: Implications for investors and firms. The Review of Financial Studies, 35(6): 2743–2789.

Kuchler, T., Russel, D., and Stroebel, J. (2022b). JUE insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook. Journal of Urban Economics, 127: 103314.

Le, T. D., Ho, T. H., Nguyen, D. T., and Ngo, T. (2021). Fintech credit and bank efficiency: International evidence. International Journal of Financial Studies, 9(3): 44.

Ndlovu, G. and Toerien, F. (2020). The distributional impact of access to finance on poverty: Evidence from selected countries in sub-Saharan Africa. Research in International Business and Finance, 52: 101190.

Nguyen, H.-L. Q. (2019). Are credit markets still local? Evidence from bank branch closings. American Economic Journal: Applied Economics, 11(1): 1–32.

Petersen, M. A. and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. The Journal of Finance, 49(1): 3–37.

Petersen, M. A. and Rajan, R. G. (1995). The effect of credit market competition on lending relationships. The Quarterly Journal of Economics, 110(2): 407–443.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. The Journal of Finance, 57(6): 2533–2570.

Presbitero, A. F. and Zazzaro, A. (2011). Competition and relationship lending: Friends or foes? Journal of Financial Intermediation, 20(3): 387–413.

Rehbein, O. and Rother, S. (2024). Social connectedness in bank lending. Working Paper.

Richardson, J., Mitchell, B., Franco, J., and Xu, Y. (2017). Bank branch closures from 2008–2016: Unequal impact in America's heartland. Technical report.

Sakong, J. and Zentefis, A. (2022). Bank access across america. Working Paper.

Schaner, S. (2018). The persistent power of behavioral change: Long-run impacts of temporary savings subsidies for the poor. American Economic Journal: Applied Economics, 10(3): 67–100.

Stulz, R. M. (2019). Fintech, bigtech, and the future of banks. Journal of Applied Corporate Finance, 31(4): 86–97.

Stylianou, K., Iacovides, M., and Lundqvist, B. (2023). Fintech Competition: Law, Policy, and Market Organisation, volume 17. Bloomsbury Publishing.

Uchida, H., Udell, G. F., and Yamori, N. (2012). Loan officers and relationship lending to SMEs. Journal of Financial Intermediation, 21(1): 97–122.

Uluc, A. and Wieladek, T. (2018). Capital requirements, monetary policy and risk shifting in the mortgage market. Journal of Financial Intermediation, 35: 3–16.

World Economic Forum (2023). Struggling cities should take notes on Detroit's revival. World Economic Forum. Accessed: 2024-09-12.

Yuan, K., Li, W., and Zhang, W. (2023). Your next bank is not necessarily a bank: Fintech expansion and bank branch closures. Economics Letters, 222: 110948.

Zhang, A. T. (2020). Estimation of high-dimensional dynamic games: Fintech lenders and bank branch closures. Working Paper.

Zhang, Q., Arora, R., and Colombage, S. (2021). The determinants of bank branch location in India: An empirical investigation. International Journal of Bank Marketing, 39(5): 856–870.

Zhou, X. (2022). Financial technology and the transmission of monetary policy: The role of social networks. Working Paper.

Figures

Figure 2.1: Number of Bank Branches

This figure illustrates the number of bank branches in the US from 2010 to 2019.

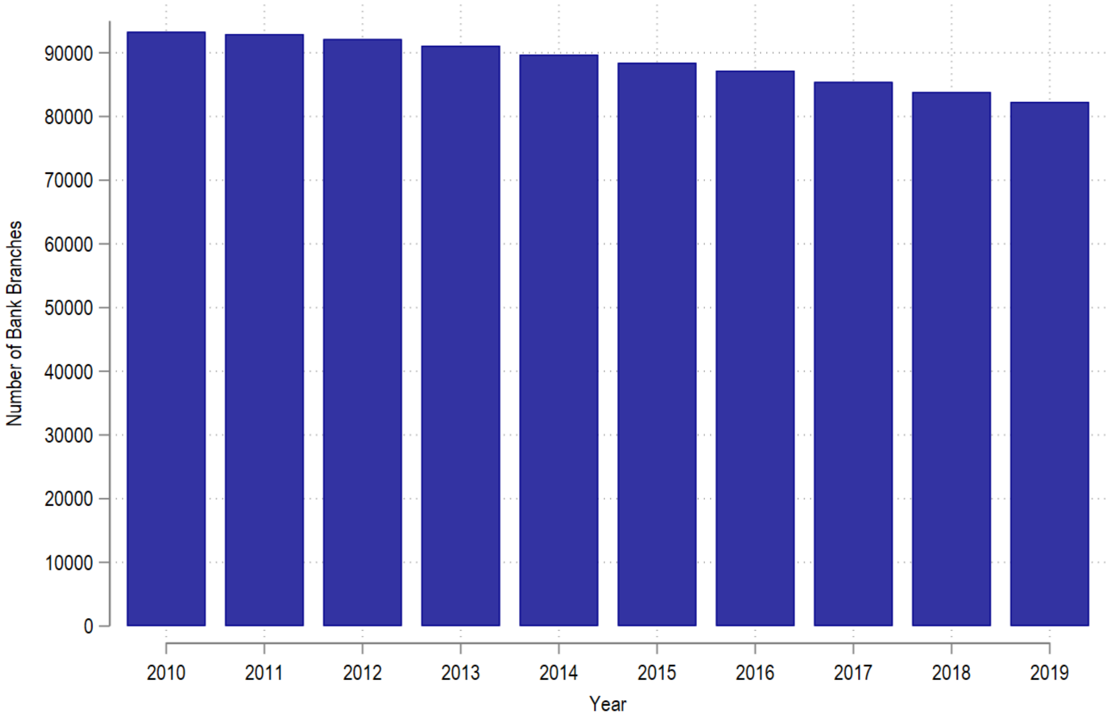
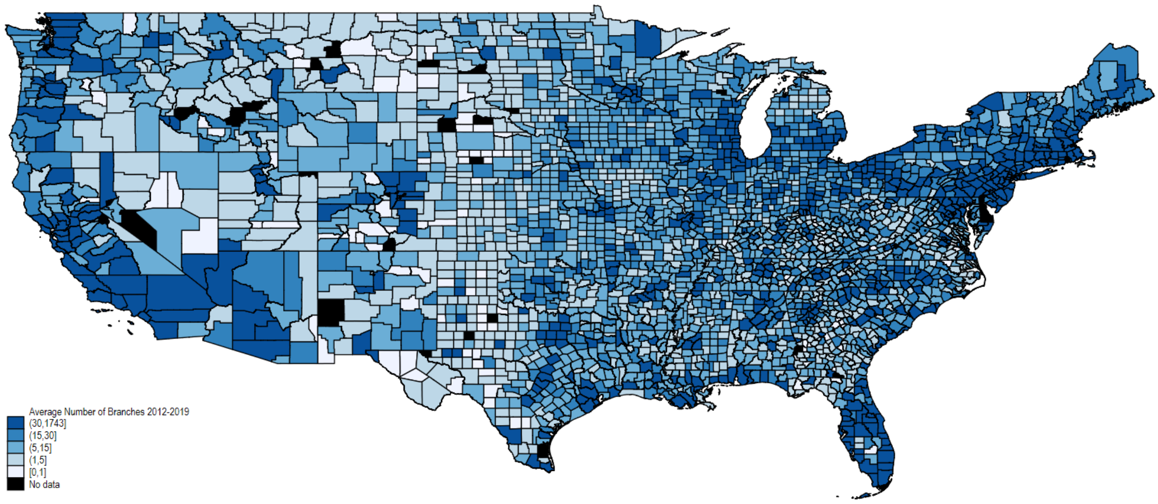


Figure 2.2: Distribution of Bank Branches

Panel A shows the average total number of bank branches from 2012-2017 per county by quintiles. Light shaded regions have a low number of branches while darker shades represent a high number of bank branches. As expected, densely populated ares such as California, Arizona, Florida and the Northeast show high numbers of bank branches, while the rural states such as Montana, Wyoming or Kansas have a low number of branches. Panel B shows the change in bank branches by county between 2012 and 2019. White shaded counties experience no losses in bank branches while the dark blue counties report the largest losses.

Panel A: Average Number of Bank Branches (2012-2019)



Panel B: Change in Number of Bank Branches (2012-2019)

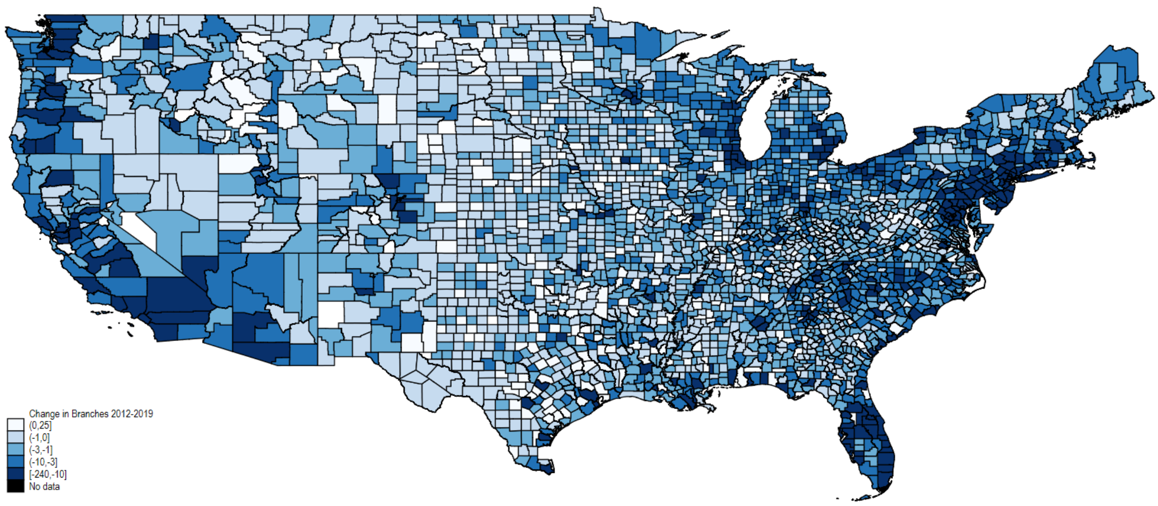


Figure 2.3: Fintech Market Shares

This figure illustrates the fintech market share for different samples and market segments from 2010-2017.

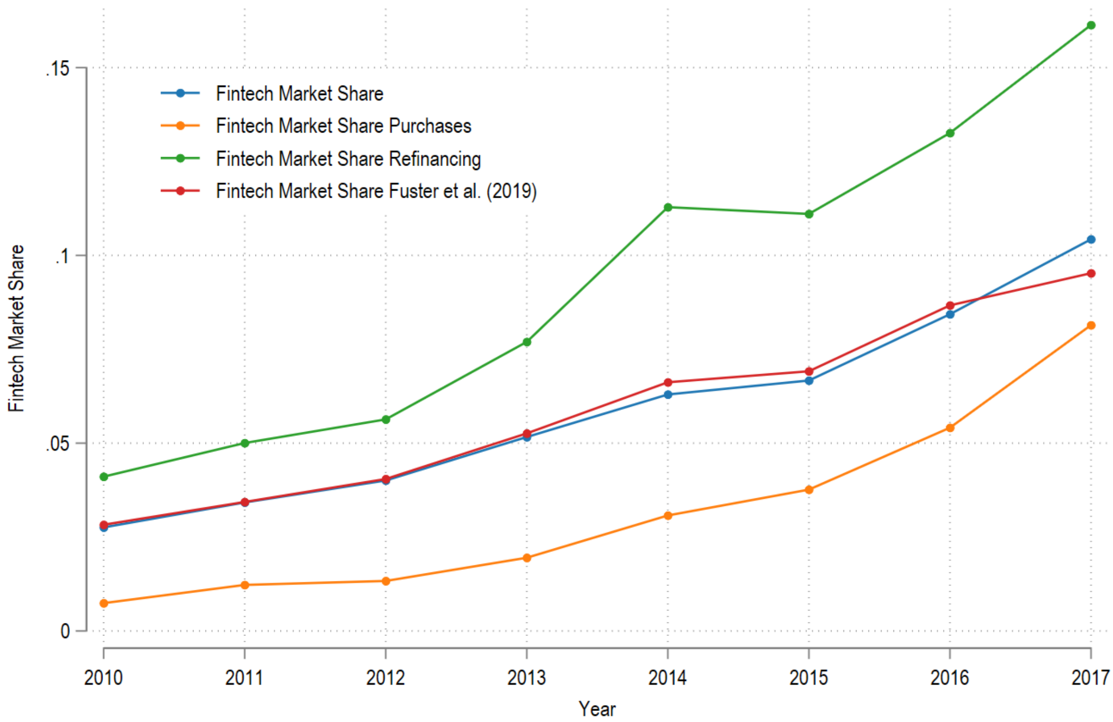


Figure 2.4: Change in Fintech Market Shares

This figure illustrates the change in fintech market share from 2010-2017 by quintiles. Darker areas exhibit a larger change in the fintech market share, no data is available for counties that are colored in black.

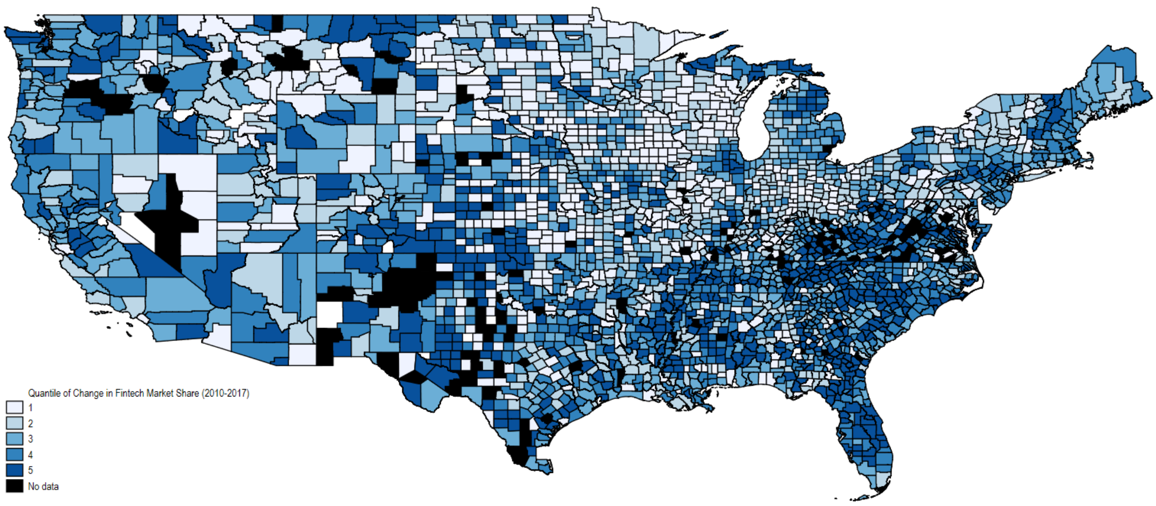
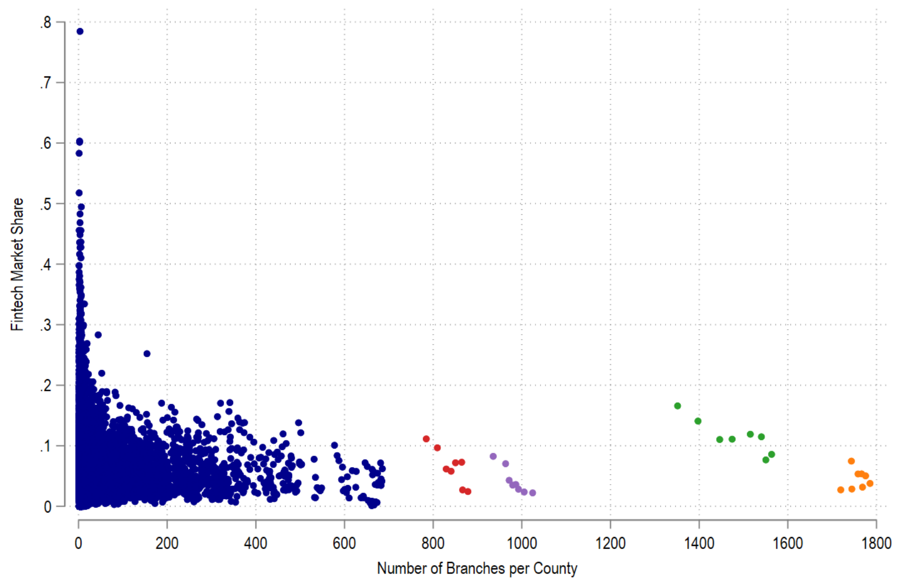


Figure 2.5: Fintech Market Share and the Number of Bank Branches per County

Panel A shows the relationship between the number of bank branches in US counties and the fintech market share in the mortgage market. Four counties with an outstanding number of bank branches are colored: Los Angeles County in orange, Cook County in green, Harris County in purple and Maricopa County in red. Panel B limits the data to all counties with less than 100 bank branches and plots the relationship between the number of branches and the fintech market share in the mortgage market.

Panel A: Full Sample



Panel B: 100 Bins from Counties with Less Than 100 Bank Branches

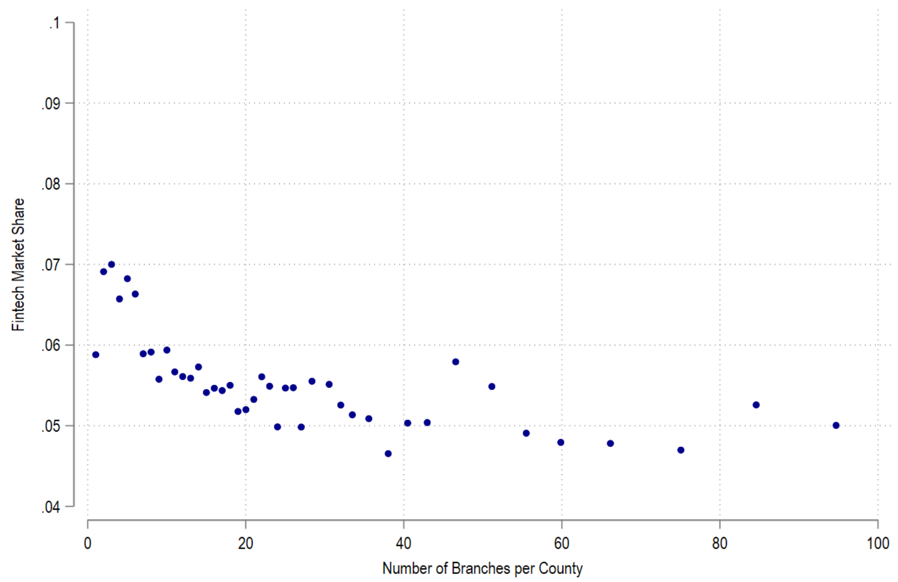
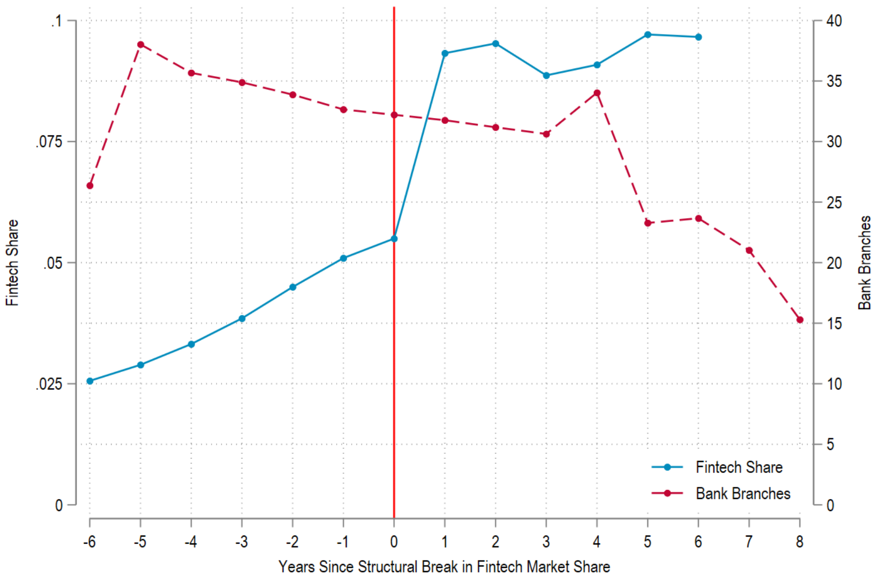


Figure 2.6: Structural Break Points

Panel A displays the fintech market share and the number of bank branches around the structural break points estimated for each county separately by Regression 2.3. Panel B plots the results of deducting the structural break year in a counties' time series of the number of bank branches from the structural break year for the time series of the fintech market share, pooled for all counties that exhibit a statistically significant structural break point in both time series.

Panel A: Structural Break Points in Fintech Market Shares



Panel B: Timing of Structural Break Points

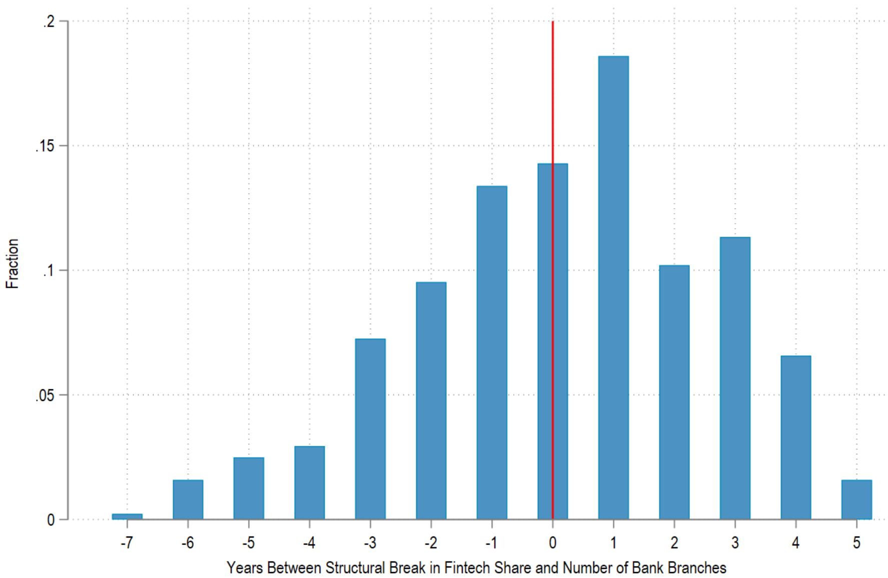
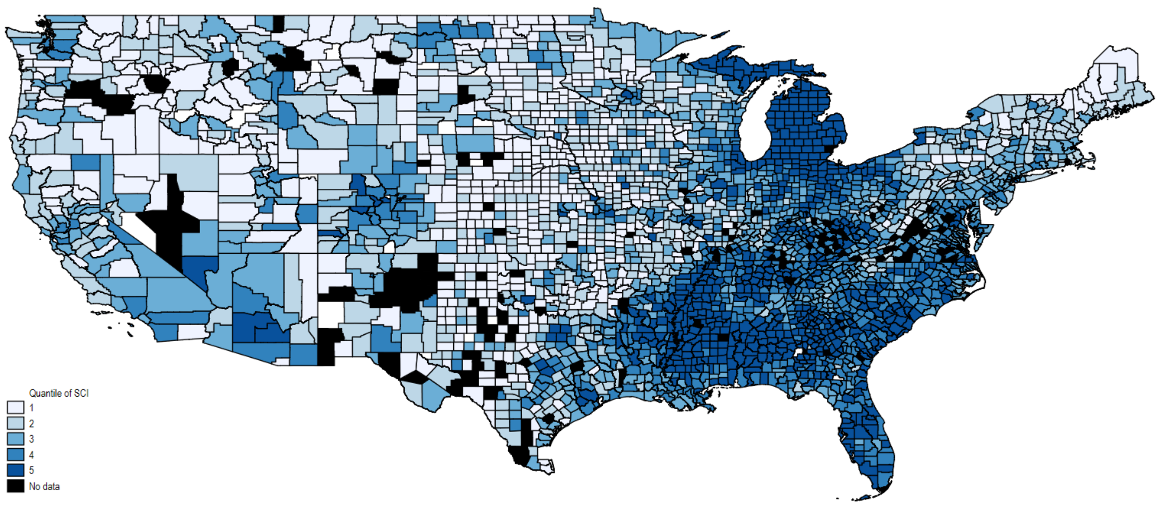


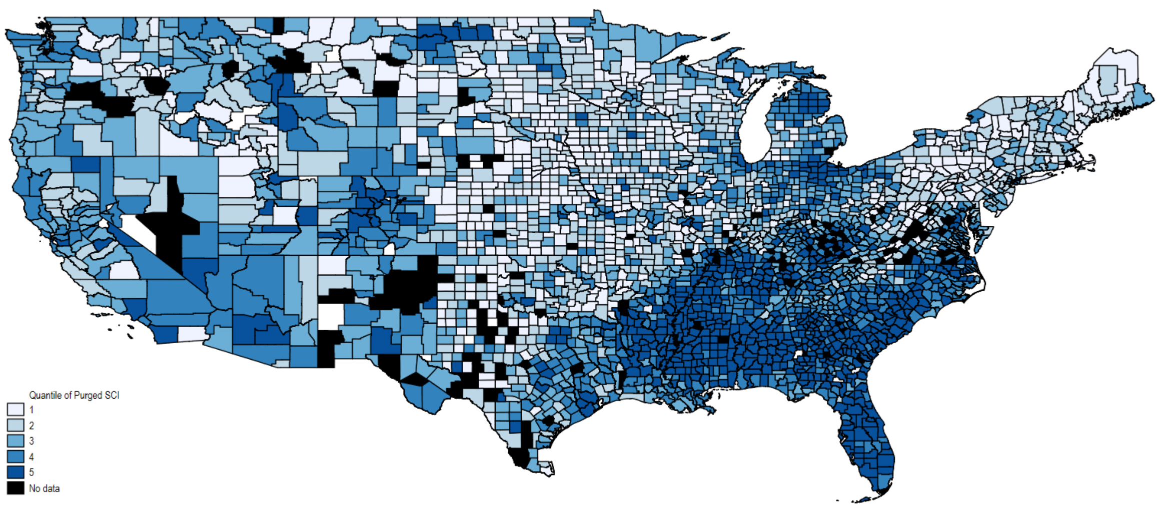
Figure 2.7: Social Connectedness to Wayne County (Michigan)

Panel A shows the social connectedness to Wayne County (Michigan) where the headquarter of the largest fintech lender Quicken Loans is located using Facebook´s Social Connectedness Index (SCI) by quintiles. Panel B shows the quintiles residuals from regressing the SCI on the log of distance to Detroit, dummies for being in the same state, commuting-zone or sharing a border as well as a dummy that indicates whether a county is located in a metropolitan area. This geographically purged version of the SCI is used as an instrumental variable for the change in the fintech market share from 2010 − 2017. Darker shading indicates a higher (purged) social connectedness.

Panel A: Social Connectedness to Wayne County, Michigan



Panel B: Purged Social Connectedness to Wayne County, Michigan



Tables

Table 2.1: Descriptive Statistics

This table displays descriptive statistics. All variables are at the county level. The data set is an unbalanced panel from 2010-2017, covering 2,907 counties from 48 states. A detailed description of all variables can be found in Table A2.1.

	Mean	Median	SD	Min	Max	Obs
Main Variables in Changes						
Δ log(Number of Bank Branches (2012-2019))	-0.11	-0.08	0.17	-1.61	0.69	2,907
Δ Fintech Market Share	0.08	0.07	0.05	-0.20	0.44	2,907
Δ Fintech Market Share in Purchases	0.07	0.06	0.07	-0.41	1.00	2,907
Δ Fintech Market Share in Refinancings	0.12	0.11	0.08	-0.54	0.74	2,907
Δ Fintech Market Share Fuster et al. (2019)	0.07	0.06	0.04	-0.20	0.43	2,693
Control Variables in Changes						
Δ log(Total Population)	0.00	-0.01	0.06	-0.27	0.68	2,907
Δ log(County GDP)	0.08	0.07	0.24	-1.61	3.05	2,907
Δ log(Median Home Value)	0.09	0.09	0.15	-0.72	0.99	2,907
Δ log(Building Permits)	0.28	0.29	0.95	-3.92	4.81	2,907
Δ log(Median Household Income)	0.12	0.11	0.09	-0.38	0.62	2,907
Δ % of Population Over 64	0.02	0.02	0.02	-0.10	0.14	2,907
Δ % of Unemployed Population	-0.03	-0.03	0.02	-0.13	0.07	2,907
Δ % of Population with Bachelor Degree	0.01	0.01	0.02	-0.09	0.17	2,907
Δ % of Population Reporting White Race	-0.01	-0.01	0.03	-0.44	0.28	2,907
Δ % of Population Reporting Black or African American Race	0.00	0.00	0.01	-0.18	0.16	2,907
Δ % of Female Population	-0.00	-0.00	0.01	-0.16	0.11	2,907
Alternative Dependent Variables in Changes						
Δ log(Small Bank Branches (2012-2019))	-0.23	-0.12	0.37	-3.43	0.69	2,819
Δ Small Bank Share (2012-2019)	-0.06	0.00	0.17	-1.00	1.00	2,907
Δ log(Large Bank Branches (2012-2019))	-0.19	-0.10	0.32	-1.95	1.10	1,887
Δ Large Bank Share (2012-2019)	-0.02	0.00	0.09	-1.00	1.00	2,907
Δ log(Institutes (2012-2019))	-0.06	0.00	0.17	-1.39	0.69	2,907
Δ log(Nationwide Assets (2012-2019))	0.05	0.30	1.25	-9.82	4.81	2,907
Δ log(Local Deposits (2012-2019))	0.16	0.15	0.20	-1.98	1.30	2,907
Main Variables						
Bank Branches (2012-2019)	29.42	10.00	74.16	1.00	1,785.00	23,426
log(Bank Branches (2012-2019))	2.47	2.30	1.21	0.00	7.49	23,426
Fintech Market Share	0.06	0.05	0.04	0.00	0.78	23,426
Fintech Market Share in Purchases	0.03	0.02	0.05	0.00	1.00	23,426
Fintech Market Share in Refinancings	0.09	0.08	0.07	0.00	1.00	23,426
Fintech Market Share Fuster et al. (2019)	0.06	0.05	0.04	0.00	0.78	22,576

Table 2.1: Descriptive Statistics (continued)

	Mean	Median	SD	Min	Max	Obs
Control Variables						
log(Total Population)	10.37	10.23	1.42	6.38	16.13	23,426
log(County GDP)	13.97	13.79	1.54	9.35	20.31	23,426
log(Median Home Value)	11.68	11.62	0.45	10.35	13.82	23,426
log(Building Permits)	3.64	3.64	2.14	0.00	10.60	23,426
log(Median Household Income)	10.72	10.71	0.24	9.85	11.77	23,426
% of Population Over 64	0.17	0.16	0.04	0.03	0.54	23,426
% of Unemployed Population	0.05	0.05	0.02	0.00	0.21	23,426
% of Population with Bachelor Degree	0.13	0.12	0.05	0.02	0.44	23,426
% of Population Reporting White Race	0.84	0.90	0.16	0.09	1.00	23,426
% of Population Reporting Black or African American Race	0.09	0.02	0.15	0.00	0.87	23,426
% of Female Population	0.50	0.50	0.02	0.19	0.58	23,426
Instrumental Variables						
log(Social Connectedness to Wayne County (Detroit))	6.93	6.77	0.88	4.74	13.34	23,426
Purged SCI to Wayne County (Detroit)	0.00	-0.11	0.66	-1.82	2.92	2,906

Table 2.2: Baseline Result

This table shows the results from estimating Equation 2.1. The dependent variable is change the number of bank branches in county i between 2012 and 2019. The main explanatory variables are the changes from 2010 to 2017 in fintech market share in Column (1), in the fintech market share in purchases in Column (2) and in refinancings in Column (3) and the change in the fintech market share from Fuster et al. (2019) in Column (4). The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017. Standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Bank Branches (2012-2019))			
	(1)	(2)	(3)	(4)
Δ Fintech Market Share	-0.186**			
	(0.073)			
Δ Fintech Market Share in Purchases		-0.079		
		(0.074)		
Δ Fintech Market Share in Refinancings			-0.149***	
			(0.044)	
Δ Fintech Market Share Fuster et al. (2019)				-0.252***
				(0.081)
Δ log(Total Population)	0.036	0.040	0.043	0.123**
	(0.061)	(0.060)	(0.061)	(0.062)
Δ log(County GDP)	0.033**	0.034**	0.030**	-0.001
	(0.015)	(0.015)	(0.015)	(0.016)
Δ log(Median Home Value)	0.189***	0.189***	0.187***	0.244***
	(0.030)	(0.030)	(0.030)	(0.029)
Δ log(Building Permits)	0.001	0.001	0.001	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Δ log(Median Household Income)	0.023	0.025	0.024	0.060
	(0.046)	(0.046)	(0.046)	(0.047)
Δ % of Population Over 64	-0.447*	-0.464*	-0.403	-0.623**
	(0.258)	(0.258)	(0.259)	(0.273)
Δ % of Unemployed Population	0.541***	0.550***	0.545***	0.530***
	(0.172)	(0.173)	(0.172)	(0.182)
Δ % of Population with Bachelor Degree	-0.126	-0.096	-0.131	-0.320
	(0.219)	(0.221)	(0.218)	(0.220)
Δ % of Population Reporting White Race	-0.061	-0.070	-0.070	-0.003
	(0.097)	(0.098)	(0.096)	(0.110)
Δ % of Population Reporting Black or African American Race	0.193	0.182	0.219	0.434
	(0.310)	(0.311)	(0.308)	(0.345)
Δ % of Female Population	-0.008	0.003	0.015	0.193
	(0.350)	(0.349)	(0.351)	(0.403)
R^2	0.058	0.056	0.059	0.079
Observations	2,907	2,907	2,907	2,693

Table 2.3: Panel Model

This table shows the results from estimating Equation 2.2 using an OLS panel data estimator in Column (1) and (2) and a PPML panel data estimator in Column (3) and (4). The dependent variable is the Number of Bank Branches in County i, logarithmized for the OLS case. The main explanatory variable is the Fintech market share. The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. Further δ_i and δ_t capture county and year fixed effects, absorbing all county-specific as well as year-specific common shocks. Standard errors are clustered on the metro level. Columns (2) and (4) serve as a robustness check by using the fintech market share data from Fuster et al. (2019). Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	log(Bank Branches (2012-2019))		Bank Branches (2012-2019)	
Model	OLS		PPML	
	(1)	(2)	(3)	(4)
Fintech Market Share	-0.065**		-0.117**	
	(0.030)		(0.048)	
Fintech Market Share Fuster et al. (2019)		-0.104***		-0.122***
		(0.033)		(0.039)
log(Total Population)	0.054	0.084	0.174***	0.182***
	(0.054)	(0.055)	(0.049)	(0.050)
log(County GDP)	0.021**	0.013	0.037***	0.037***
	(0.009)	(0.009)	(0.013)	(0.013)
log(Median Home Value)	0.141***	0.157***	0.116***	0.117***
	(0.019)	(0.019)	(0.022)	(0.023)
log(Building Permits)	0.000	0.001	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
log(Median Household Income)	0.025	0.044*	0.091***	0.098***
	(0.028)	(0.027)	(0.034)	(0.034)
% of Population Over 64	-0.409**	-0.494**	-0.657***	-0.692***
	(0.189)	(0.205)	(0.235)	(0.246)
% of Unemployed Population	0.491***	0.495***	0.783***	0.793***
	(0.132)	(0.137)	(0.145)	(0.149)
% of Population with Bachelor Degree	-0.230	-0.293**	-0.185	-0.193
	(0.144)	(0.134)	(0.146)	(0.149)
% of Population Reporting White Race	-0.103	-0.090	0.002	0.003
	(0.089)	(0.094)	(0.083)	(0.086)
% of Population Reporting Black African American Race	0.221	0.309	0.258	0.282
	(0.235)	(0.256)	(0.288)	(0.298)
% of Female Population	0.063	0.129	0.043	0.095
	(0.169)	(0.194)	(0.207)	(0.232)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.996	0.997		
Pseudo R^2			0.934	0.934
Observations	23,426	22,562	23,426	22,562

Table 2.4: Instrumental Variable Approach

This table shows the results from estimating Equation 2.1 with an instrumental variable for the change in the fintech market share. The dependent variable is change the Number of Bank Branches in County i between 2012 and 2019. The main explanatory variables are the changes from 2010 to 2017 in fintech market share. Facebook's Social Connectedness Index (SCI) to Wayne County, where the headquarter of Quicken Loans (the largest fintech company) is located in Detroit is used as instrumental variable. The SCI is purged by the following geographic factors to isolate pure social connectedness: Log of Distance to Wayne County, dummy variables for the same state, a common border, a common commuting zone (determined by labor market area) and a dummy that indicates whether the respective county is part of a metropolitan area. Column (1) displays the first stage, Column (2) the second stage results. Both stages control on county level for changes in the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, the dependent variable of the second stage enters the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017, standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ Fintech Market Share (2010-2017)	Δ log(Bank Branches (2010-2017))
	First Stage	Second Stage
	(1)	(2)
Purged SCI to Wayne County (Detroit)	0.015***	
	(0.002)	
Δ Fintech Market Share		-0.724**
		(0.347)
Δ log(Total Population)	-0.084***	0.001
	(0.019)	(0.063)
Δ log(County GDP)	0.011	0.038**
	(0.007)	(0.016)
Δ log(Median Home Value)	0.008	0.193***
	(0.010)	(0.031)
Δ log(Building Permits)	0.000	0.001
	(0.001)	(0.004)
Δ log(Median Household Income)	-0.005	0.011
	(0.015)	(0.047)
Δ % of Population Over 64	0.096	-0.374
	(0.076)	(0.261)
Δ % of Unemployed Population	0.013	0.494***
	(0.054)	(0.174)
Δ % of Population with Bachelor Degree	-0.172***	-0.235
	(0.064)	(0.224)
Δ % of Population Reporting White Race	0.099**	-0.011
	(0.040)	(0.103)
Δ % of Population Reporting Black or African American Race	0.013	0.195
	(0.083)	(0.315)
Δ % of Female Population	-0.124	-0.065
	(0.114)	(0.361)
F-value (1st stage)		75.676
Observations	2,906	2,906

Table 2.5: Spatial Heterogeneity

This table shows the results from estimating Equation 2.1 including an interaction term of the change in fintech market share and a dummy equal to one, if the county is located inside of major statistical area of the US. The dependent variable is the change in the number of bank branches between 2012 and 2019. The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017, standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Bank Branches (2012-2019)			
	(1)	(2)	(3)	(4)
Δ Fintech Market Share	-0.129*			
	(0.074)			
Δ Fintech Market Share in Purchases		-0.042		
		(0.074)		
Δ Fintech Market Share in Refinancings			-0.131***	
			(0.044)	
Δ Fintech Market Share Fuster et al. (2019)				-0.165**
				(0.080)
Δ Fintech Market Share× Metro Dummy	-0.382***	-0.404***	-0.391***	-0.378***
	(0.098)	(0.099)	(0.095)	(0.099)
Δ log(Total Population)	0.130**	0.140**	0.137**	0.215***
	(0.066)	(0.065)	(0.065)	(0.067)
Δ log(County GDP)	0.030**	0.030**	0.028*	-0.003
	(0.015)	(0.015)	(0.015)	(0.016)
Δ log(Median Home Value)	0.169***	0.167***	0.166***	0.224***
	(0.031)	(0.031)	(0.031)	(0.030)
Δ log(Building Permits)	0.002	0.002	0.002	0.004
	(0.004)	(0.004)	(0.004)	(0.004)
Δ log(Median Household Income)	0.021	0.022	0.021	0.055
	(0.046)	(0.047)	(0.046)	(0.047)
Δ % of Population Over 64	-0.424*	-0.435*	-0.379	-0.608**
	(0.256)	(0.256)	(0.257)	(0.271)
Δ % of Unemployed Population	0.538***	0.545***	0.538***	0.533***
	(0.171)	(0.172)	(0.171)	(0.181)
Δ % of Population with Bachelor Degree	-0.101	-0.078	-0.112	-0.283
	(0.218)	(0.220)	(0.217)	(0.218)
Δ % of Population Reporting White Race	-0.074	-0.082	-0.079	-0.024
	(0.097)	(0.098)	(0.096)	(0.110)
Δ % of Population Reporting Black or African American Race	0.183	0.176	0.206	0.440
	(0.310)	(0.311)	(0.308)	(0.341)
Δ % of Female Population	0.024	0.035	0.041	0.242
	(0.351)	(0.350)	(0.351)	(0.404)
R^2	0.064	0.063	0.066	0.086
Observations	2,907	2,907	2,907	2,693

Table 2.6: Heterogeneity by Bank Size

This table shows the results from estimating Equation 2.1 including an interaction term of the change in fintech market share and a dummy equal to one, if the county is located inside of major statistical area of the US. The dependent variables are the change in the log number of small bank (< 10bn USD) Column (1), the change in the share of small banks in Column (2), the change in the log number of branches from large banks (> 50bn USD) in Column (3) and the share of large banks in Column (4) between 2012 and 2019. The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017, standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Small Bank Branches (2012-2019))	Δ Share Small Bank Branches (2012-2019)	Δ log(Large Bank Branches (2012-2019))	Δ Share Large Bank Branches (2012-2019)
	(1)	(2)	(3)	(4)
Δ Fintech Market Share	-0.403***	-0.141**	0.005	-0.041
	(0.128)	(0.066)	(0.200)	(0.041)
Δ Fintech Market Share× Metro Dummy	-0.791***	0.039	-0.221	0.025
	(0.273)	(0.109)	(0.240)	(0.065)
Δ log(Total Population)	-0.990***	-0.376***	0.480***	0.064*
	(0.171)	(0.066)	(0.157)	(0.038)
Δ log(County GDP)	0.033	0.010	0.095***	0.015**
	(0.025)	(0.012)	(0.035)	(0.007)
Δ log(Median Home Value)	0.514***	0.123***	-0.178**	-0.055***
	(0.062)	(0.028)	(0.074)	(0.018)
Δ log(Building Permits)	-0.012*	0.000	-0.001	0.002
	(0.007)	(0.003)	(0.010)	(0.002)
Δ log(Median Household Income)	-0.041	-0.041	0.284**	0.005
	(0.078)	(0.041)	(0.128)	(0.026)
Δ % of Population Over 64	-1.314***	-1.000***	-0.702	-0.410***
	(0.493)	(0.259)	(0.651)	(0.133)
Δ % of Unemployed Population	1.535***	0.370**	0.840*	0.110
	(0.340)	(0.184)	(0.452)	(0.096)
Δ % of Population with Bachelor Degree	0.433	0.191	-0.065	0.100
	(0.355)	(0.203)	(0.516)	(0.092)
Δ % of Population Reporting White Race	0.068	0.075	-0.419	0.006
	(0.187)	(0.097)	(0.264)	(0.065)
Δ % of Population Reporting Black or African American Race	-0.919*	-0.600	0.280	0.144
	(0.540)	(0.374)	(0.618)	(0.129)
Δ % of Female Population	-0.361	0.339	-0.391	0.127
	(0.620)	(0.458)	(0.994)	(0.201)
R^2	0.132	0.053	0.026	0.015
Observations	2,819	2,907	1,887	2,907

Table 2.7: Effects on Banks

This table shows the results from estimating Equation 2.1. The dependent variables are the change in the log institutes that operate a bank in the respective county in Column (1), the change in the log of total deposits of all branches in the respective county in Column (2) and the change in the log total assets of institutes operating a branch in the respective county in Column (3). The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017, standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Institutes (2012-2019))	Δ log(Nationwide Assets (2012-2019))	Δ log(Local Deposits (2012-2019))
	(1)	(2)	(3)
Δ Fintech Market Share	-0.064	0.283	-0.174***
	(0.069)	(0.620)	(0.067)
Δ log(Total Population)	0.148**	1.053***	1.158***
	(0.061)	(0.359)	(0.068)
Δ log(County GDP)	0.012	0.023	0.098***
	(0.014)	(0.105)	(0.018)
Δ log(Median Home Value)	0.120***	-0.094	-0.025
	(0.029)	(0.215)	(0.030)
Δ log(Building Permits)	-0.000	0.037	0.009**
	(0.003)	(0.031)	(0.004)
Δ log(Median Household Income)	-0.042	-0.041	0.091*
	(0.045)	(0.381)	(0.047)
Δ % of Population Over 64	-0.466*	-3.655*	0.980***
	(0.249)	(1.866)	(0.260)
Δ % of Unemployed Population	0.273	1.284	0.139
	(0.171)	(1.194)	(0.184)
Δ % of Population with Bachelor Degree	-0.039	-0.609	0.140
	(0.211)	(1.635)	(0.214)
Δ % of Population Reporting White Race	-0.104	-0.057	-0.113
	(0.090)	(0.855)	(0.120)
Δ % of Population Reporting Black or African American Race	0.305	0.589	0.173
	(0.355)	(2.145)	(0.339)
Δ % of Female Population	-0.046	2.309	0.334
	(0.331)	(2.739)	(0.406)
R^2	0.025	0.007	0.196
Observations	2,907	2,907	2,907

CHAPTER 3

Flooded Friends - Peer Effects in Insurance Decisions[1]

Abstract

This paper explores how peer effects influence homeowners' decisions to insure against elemental damages. Using Facebook friendship links to flooded areas, it finds that stronger social ties increase insurance purchases in non-flooded regions: A doubling of social ties leads to a 1.2% rise in policies covering elemental damages. It highlights the moderating role of regional climate policy attitudes and social capital, showing that bridging social capital attenuates whereas bonding social capital amplifies the effect. This study contributes to the understanding of how social capital and climate attitudes influence household responses to climate risks.

[1] Unpublished

3.1 Introduction

Natural disasters are becoming more frequent, more visible and more destructive. Floods have been occurring with increasing frequency in recent years, devastating entire regions such as large parts of western Germany, Belgium and France in July 2021, the Emilia-Romagna region in Italy in May 2023 or parts of Bavaria in June 2024. As a result of climate change, extreme rainfall events are leading to increased flood magnitudes, raising the risk of significant flood damage to residential infrastructure (Wasko et al., 2021). Most recently, large parts of Austria, Poland, the Czech Republic and Romania have been flooded. Conditional on a specific location such events are relatively rare. However, the realization of such large-scale once-in-a-century floods poses a serious threat and results in substantial welfare loss for private households. Across countries, people misperceive the actual probability of these high-loss events which results in low rate of insurance coverage against these risks (Wagner, 2022; GDV, 2022).

According to Germany's National Meteorological Service, the flood in 2021 in parts of western Germany ("the flood") can be classified as a once-in-a-century event (Junghänel et al., 2021). This paper uses the flood as a shock to explore the following research question: Do households decide to insure against elemental risk based on their peers' flood experiences? This study utilizes a unique dataset provided by the German insurance company DEVK to address the research question. In Germany, standard residential building insurance covers common risks such as fire and storm damage. However, to include protection against natural hazards like flooding (referred to as "elemental damages"), homeowners must opt for an additional policy component. The dataset contains the number of residential building insurance policies, along with detailed information on the proportion of policies that include coverage for elemental damages. Additionally, it includes data on insurance premiums and coverage levels, enabling a thorough analysis of how peer effects shape insurance purchasing decisions.

To measure social ties into the flooded regions I use the Social Connectedness Index (SCI) from Bailey et al. (2018b) as a proxy for real-world connections. Combined with flood intensity, proxied by claim ratios from the German Insurance Association (GDV), I construct an intensity-weighted social connectedness for every non-flooded region to the flooded regions. This allows to estimate a difference-in-difference model, identifying the effect of social connectedness on the share of policies that include the additional coverage against elemental damages.

The results show that larger social ties are linked to a significantly higher share of policies covering elemental damages, while the overall number of policies does not increase. They indicate that, after the disaster in 2021, people decided to add protection against natural hazards to their existing policies. This effect remains stable when controlling for current risk exposure, past hazard experiences, future climate projections, and geographic as well as topological characteristics of regions. I explicitly show that the results are not driven by distance to the flooded regions, the definition of the flooded regions, or shocks correlated with social connectedness. People most likely think about their own insurance coverage in case of elemental damages when peers are affected. This exposure makes flood risks a salient topic after the flood with people reassessing their own flood risk (Hu, 2022; Gao et al., 2020). Put differently, the flood triggers an information update about the own flood risk. Results are robust when choosing a binary setting with a treatment and a control group instead of a continuous treatment.

This paper highlights the influence of regional disparities in climate policy attitudes and varying forms of social capital as key factors moderating the strength of peer effects on insurance decisions. I use heterogeneity in votes for the green party and the willingness to invest in or pay for renewable energy to approximate attitudes towards climate protection measures. Ex-ante, the direction of the effect is theoretically unclear. On the one hand, skepticism about climate change could lead to an inelastic response due to a drastically reduced perception of climate risks. On the other hand, first-hand experiences by social ties could cause awareness and a heightened sense of urgency to mitigate risks through insurance. Latent climate scepticism could be outweighed by information transmitted via social ties, making a very abstract risk more tangible. The results show that peer effects are stronger for regions with greater reluctance towards climate protection measures.

Additionally, the paper explores the role of social capital, distinguishing between bridging and bonding capital (Putnam, 2000). Bridging social capital refers to social connections that link individuals from diverse backgrounds, facilitating access to varied information and resources. For example, a local community that brings together residents with various occupations and educational backgrounds can facilitate knowledge exchange and collaborative problem-solving, helping to address community needs more effectively. Bridging capital is further linked to increased trust in institutions and the government. On the contrary, bonding capital refers to the strong social ties and support networks within a close-knit, homogeneous group. For example, in a neighborhood where everyone shares similar backgrounds and values, residents may provide each other with substantial personal support and resources, but have fewer connections with people outside their immediate community. Surprisingly, bridging capital attenuates the effect of social ties, while bonding capital enforces the peer effects. Both channels suggest that social connectedness can make up for failures in ex-ante risk adjustments.

The paper contributes to three strands of the literature. First, it adds to the large field of social effects in economic perceptions and decision-making, explaining how social interaction influences economic outcomes (Hirshleifer, 2020; Granovetter, 2011). Several papers show that social ties can mitigate trade and investment frictions using the Social Connectedness Index (Bailey et al., 2021; Kuchler et al., 2022a; Hirshleifer et al., 2024; Nguyen et al., 2023; Dornseifer and Rehbein, 2023). Social ties can reduce barriers in bank lending (Rehbein and Rother, 2024; Kariya and Shekhawat, 2024) and spur innovation (Diemer and Regan, 2022; Obschonka et al., 2023). With respect to the mortgage market, a large number of papers find peer effects in house price beliefs and financing decisions (Kuchler and Stroebel, 2021; Bailey et al., 2019; Maturana and Nickerson, 2019; McCartney and Shah, 2022). Besides that, personal finances are influenced by social networks. Ties to people of higher socioeconomic status increase economic mobility (Chetty et al., 2022a,b). They affect stock market participation (Cannon et al., 2024)[2], consumption (Agarwal et al., 2021; De Giorgi et al., 2020) and insurance uptake (Hu, 2022). Ratnadiwakara and Venugopal (2023) show that climate risk perceptions play a role for insurance uptake and find that, in areas with more people concerned about global warming, voluntary flood insurance coverage is higher.

The paper relates to the literature on natural disasters and green finance. Giglio et al. (2021) offer a comprehensive review of recent literature, covering both the integration of climate risk into macrofinancial models and empirical studies related to the pricing of climate risks across various asset classes. With respect to natural disasters, particularly floods, several papers study the implications for firms and their capital requirements (Noth and Rehbein, 2019; Segura and Villacorta, 2023; Koetter et al., 2020; Forte et al., 2024). The role of banks in financing the recovery after a disaster is emphasized by Brei et al. (2019) and Celil et al. (2022). Banks also cause spillover effects through their lending networks. Rehbein and Ongena (2022) explore indirect effects on non-affected firms via banks and identify real effects for the local economy. Alongside firm collateral, mortgages are the asset class most affected by climate risks and disasters. Implications for mortgage lending are studied by Ouazad and Kahn (2022); Wong and Ho (2023); Nguyen et al. (2022). Severe storms can result in significant asset losses for households. Gallagher and Hartley (2017) investigate how households manage the financial burden of reconstruction in the aftermath of such events. Similarly, You and Kousky (2024) examine household consumption patterns following hurricane landfalls. In addition, Billings et al. (2022) assess post-disaster credit outcomes and the effectiveness of federal disaster relief programs in addressing these financial challenges.

I contribute to the literature on decision-making in insurance uptake. Botzen and van den Bergh (2012) explore individual risk preferences and the willingness to pay for flood insurance. In the same direction, Wagner (2022) finds that high-risk homeowners' willingness to pay for flood insurance is below their expected payoff. Bakkensen and Barrage (2022) find overconfidence of coastal area residents with respect to flood risk by conducting a survey in Rhode Island. Potential explanations for these discrepancies include a mismatch between peoples' risk perception and the fundamentals in natural hazard models (Ratnadiwakara and Venugopal, 2023). Risk perceptions do not only rely on the underlying climate risks but also

[2] A large literature focuses on biases resulting from peer effects in stock market investments (Heimer, 2016; Ouimet and Tate, 2020; Balakina et al., 2022; Cookson et al., 2023) and retirement savings (Beshears et al., 2015; Bauer et al., 2022).

on perceived social expectations (Lo, 2013). Conceptualizing households' decision to insure against natural hazard risks, Gallagher (2014) models insurance behavior after a natural disaster as a Bayesian learning process that discounts past floods. Their model reveals a substantial gap between home values and fundamentals. They show that this misalignment can be solved by an enforced insurance mandate at actuarially fair rates. Their empirical finding that uptakes after natural hazards are a short-lived phenomenon is confirmed by Atreya et al. (2015). Finally, the experimental literature has identified the effect of social networks on insurance uptake (Jain, 2020; Cai et al., 2015).

Two papers, Ratnadiwakara (2021) and Hu (2022) combine social connectedness and flood incidents to identify peer effects in households' flood insurance decisions. However, the settings differ in several points. Ratnadiwakara (2021) interprets social connectedness as directly transmitting information via social media posts and links it to the perception of global warming, while this paper uses social connectedness as a proxy for real-world connections. Closer to this paper, Hu (2022) concentrates on flood events from the US east coast which are seasonal and relatively frequent. He compares counties that are at least 750 miles away from the natural hazards using a binary treatment definition by discretizing the social connectedness index. Both papers use social connectedness to regions that have a high risk of being hit by a hurricane (e.g. County Harris which was hit by Hurricane Harvey in 2017). In contrast to that, the flood in Germany in 2021 hit regions that have not been known and prepared for such events. Social ties to the region are therefore less likely to be confounded by prior flood experiences.

In the US, the National Flood Insurance Program (NFIP) offers subsidized contracts to people living in participating NFIP communities next to a private insurance market. In addition, flood policies are independent insurance contracts. In contrast to that, the German insurance market is an entirely private market with larger frictions than the US market. The coverage of natural hazard events such as flooding can be chosen as an extra component to a quasi-mandatory residential building insurance. Finally, this paper makes two key contributions by identifying factors that moderate peer effects in insurance adoption. First, it shows that the influence of peer effects is shaped by local attitudes toward climate protection policies. Second, it demonstrates that the degree of responsiveness to these peer effects is largely influenced by the level of local social capital within the community. These findings provide important insights into how regional characteristics with respect to climate risk and the ability to obtain information from various sources shape the effect of social connectedness on insurance uptake.

3.2 Background and Data

3.2.1 The Flood

Between July 12 and 15, an intense and prolonged period of heavy rainfall saturated the soil and overwhelmed rivers and tributaries across the western and southern parts of Germany. The flood lasted for several days, with torrential downpours and rising water levels causing extensive damage to residential buildings, infrastructure, and the environment.
Figure 3.1 illustrates the intensity of the flood in 2021. The flood hit the western as well as some eastern parts of Germany, close to the Czech border. In the north and south of Germany only a few regions reported damages. Two regions located south of the city Bonn, between the river Rhine and the French border, registered the highest claim ratios. In Euskirchen 23.9% of all insurance contracts filed a damage claim, the respective number for Ahrweiler equals 18.3%. Due to the high number of fatalities in the district of Ahrweiler and the particularly high water levels of the river Ahr, the catastrophe is known as the flood in the Ahr valley. Because of the geography of the affected areas, flash floods developed with masses of water flowing down steep, v-shaped valleys. Overall, amounts of rainfall that statistically occur at most once every 100 years in these regions were measured (Junghänel et al., 2021). The severity of the flood resulted in numerous evacuations, widespread disruption of essential services, and loss of lives. The event garnered significant attention both nationally and internationally, prompting discussions on the need for effective disaster preparedness and mitigation of natural hazards. The German Insurance Association records over 180 deaths, 2,187 insured losses to single-family homes with damage costs beyond 100,000, 12.6 billion euros of total damages in insured property and motor vehicles, and the most devastating year for the insurance industry in its existence (GDV, 2022). The third quarter of 2021 saw a surge in insurance demand due to natural catastrophes, coinciding with heightened media coverage in Germany. An estimated number of $400,000$ new contracts were acquired, with the typical range for new contracts falling between $50,000$ and $100,000$ per quarter (GDV, 2021).

3.2.2 Data

This section presents all data used throughout the paper. Descriptive statistics can be found in Table 3.1. Definition of all variables and data sources are presented in Table A3.1.

Insurance Data: Homeowners' insurance data is taken from a German insurance company (DEVK). DEVK Insurance, a German company established in 1886 and headquartered in Cologne, offers a wide range of insurance products, including property, life, and health coverage. With a strong national presence, the company serves both individual clients and businesses throughout Germany. The data set includes three different measures: First, the overall number of residential building insurance policies. Second, the premiums that are paid by the policyholder and third, the coverage, i.e. the overall value of the insured properties.[3] Residential building insurance in Germany is structured as follows: During the construction period, every residential building is covered by a fire shell policy. On completion, this policy is converted into a residential building insurance policy. In the event of a sale, the policy is transferred to the new owner. Residential building insurance is not mandatory by law, but it is included as a condition in the lending process of banks. Thus, it is indirectly obligatory for every building financed by bank loans and serves as collateral. By default, storm damage, hail damage, and damages resulting from lightning strikes are covered by this residential building insurance. However, to insure against damages caused by natural events, such as flooding, backflow, earthquakes, or snow pressure, elemental damage coverage is necessary. This type of coverage can be obtained as an additional component to the existing residential building insurance policy. The insurance data include information on which share of the existing policies include this additional component covering elemental damages. The percentage of all premiums paid that is attributable to policies that include elemental damages and the share of the overall coverage that is attributable to the policies including elemental damages are also reported. Data are available on postcode level and aggregated to county level ("Kreise und kreisfreie Städte"), which corresponds to the EU NUTS3 regions for Germany to match it with social connectedness data. Information on the individual policy level about inclusion or exclusion of elemental damages is not available.

Figure 3.2 illustrates the data aggregated for Germany with January 2017 normalized to 100. Panel A shows the total number of policies as well as the share of policies that includes elemental damages and the share that does not cover these risks. Several trends can be observed: The number of all policies is relatively stable, increases during the year and experiences a setback every January. This pattern is attributable to the contract design. Customers can enter into such insurance or add cover for natural hazards at any time. However, contracts can only be terminated at the end of the year or in case the property is sold. The green line shows a constant increase of policies including elemental damage coverage while the orange line decreases. While only 40% of policies included elemental damages in 2017, this number increases to over 60% towards the end of 2023. After the flood in 2021 one can see a rise that clearly diverges from the long-term trend. People seem to react to the flood and more customers than usual add coverage against elemental damages. Compared to the universe of residential building insurances in Germany, the distribution of policies with and without elemental damage coverage for the DEVK is comparable to the universe of policies in Germany (GDV, 2022). Panel B shows the overall premiums paid for all policies, those that are paid for all risks except elemental damages as well as those that specifically cover the additional component against elemental damages separately. As the contracts cover the

[3]To preserve confidentiality, postcodes with less than five contracts in any month of the panel are excluded from the analysis.

current replacement value of the property, the premiums are adjusted to a construction price index. The share attributable to the elemental damage coverage is negligible while the other categories show annual jumps that reflect the development of inflation. Panel C shows the coverage of policies and combines the described trends. Jumps at the end of the year are again caused by inflation adjustment.

Flood Data: To identify the intensity of the shock I also use data for the entire population of residential building insurance contracts published by the German Insurance Association as a yearly report of the most severe natural events. For this purpose, damage claims to all German insurance companies are summarizes on county level by each event (GDV, 2022). For the flood in 2021, this share varies between 0% and 23.93%.

Social Connectedness: To proxy for social connections between flooded and non-flooded regions, I use Facebook's Social Connectedness Index, introduced and described in detail for the US county level by Bailey et al. (2018b) and for Europe by Aref et al. (2020). It defines the connectedness between region i and region j as follows:

$$SCI_{i,j} = \frac{FacebookConnections_{i,j}}{FacebookUsers_i \times FacebookUsers_j} \tag{3.1}$$

The SCI can be interpreted as the relative probability of a Facebook user from region i being friends with a Facebook user from region j.[4] Since friendship connections and user numbers are relatively stable, the SCI is a time-invariant variable. I use data from August 2020, available at the Humanitarian Data Exchange.[5] To account for outliers in the distribution of the SCI, the data are winsorized at the 99$^{\text{th}}$ percentile. The network size of almost 40 million users in Germany around August 2020 (NapoleonCat, 2023) suggests that the usage of the social network is nationally widespread. The SCI therefore provides an unprecedented picture of social connectedness in Germany. Moreover, several studies show that Facebook friendships most likely map real-world connections between relatives, friends and colleagues (Bailey et al., 2018a; Kuchler et al., 2022b).

Figures 3.3 and 3.4 map the SCI for the counties Euskirchen and Ahrweiler. Both are highly connected to counties in western Germany. As the two areas are neighboring, the distribution looks very similar at first glance. However, one can clearly see that Euskirchen, which is part of the federal state of North Rhine-Westphalia (NRW), is more connected to the counties in the north of NRW. In contrast, Ahrweiler has stronger connections to Rhineland-Palatinate and counties in the south west. Both Euskirchen and Ahrweiler exhibit strong connections

[4]To preserve confidentiality, Facebook multiplies the result of Equation 3.1 by a random factor. Hence, no conclusions can be drawn on the actual user data, and the SCI takes values between one and 1 billion.

[5]The latest available data of the SCI is for October 2021 (see Humanitarian Data Exchange). However, I use data from August 2020 to rule out that connections formed after and because of the flood might bias the analysis.

to traditional tourist destinations, such as the North Sea coast, its islands, or the Alps. In terms of large cities, Euskirchen is best connected to Bonn and Cologne, while Ahrweiler has the strongest ties to Bonn and Koblenz.
Various factors can influence and strengthen the social connectedness between two counties, including shared economic interests, cultural ties, geographic proximity, and the movement of people between regions. Connections can also be shaped by historical relationships, infrastructure development, and collaborative governance or community initiatives. For instance, the district of Ahrweiler is far more socially connected to the Kyffhäuserkreis in Thuringia than the district of Euskirchen. Since 1990, the two regions have maintained a partnership that likely fostered personal connections. Following the flood, summer camps were organized in Thuringia for children from the affected areas, and local newspapers frequently reported on donations and other relief efforts. Notably, 50 emergency service members from Kyffhäuserkreis were awarded the Medal of Honor by the state of Rhineland-Palatinate for their commitment to help in the aftermath of the 2021 flood (Landratsamt Kyffhäuserkreis, 2023).

I complement the data by the population-weighted distance between counties, taken from TERCET Files from the European Commission, data on previous natural hazard events and claim ratios for the universe of policies against elemental damages of the entire German insurance industry from 2002-2021 (GDV, 2022) as well as climate risk measures from the European Climate Risk Typology (Hincks et al., 2023). To explore the underlying mechanisms, climate policy support measured on county level by Levi et al. (2022), official election data and Covid-19 vaccination data is used (Bade et al., 2024). As proxies for social capital, I compute a dispersion measure using the SCI and educational equality using data from the 2022 census. Detailed descriptions for all variables can be found in Table A3.1.

3.3 Empirical Analysis

3.3.1 Construction of Social Connectedness to Flooded Areas

The flood in July 2021 is used as a shock to assess the effect of social connectedness between unflooded counties and flooded counties on insurance uptake in unflooded regions. I estimate a difference-in-difference model with a continuous treatment. Since all counties are socially connected, there are no untreated regions that can serve as a control group. The difference-in-difference model with a continuous treatment allows to evaluate the effect of social connectedness even though a binary difference-in-difference approach is not applicable (Callaway et al., 2024) and has already been used by Card (1992). The main focus of the analysis is on insurance uptake in those regions which are not hit by the flood. Data from the German Insurance Association is used to distinguish between flooded and non-flooded counties. They collect the insurance claims of the entire German insurance industry for all natural disasters of a certain size.[6] For the flood in 2021, 281 regions exhibit a claim ratio of 0%.[7] These regions enter the model as the treated regions and are indexed with i in Equation 3.2. As described in Section 3.2, two counties stand out in terms of the claim ratio: Euskirchen with a claim ratio of 23.9% and Ahrweiler with a claim ratio of 18.3%. I assume that these two regions significantly shaped the perception of the disaster and are therefore suitable for investigating the influence of social connectedness on insurance policies in unaffected areas. Section 3.4 shows that the results hold if this assumption is relaxed and more regions are used to define social connectedness to the flood.

For the construction of the social connectedness to the flood, flooded regions are indexed j. In the baseline estimation, all other regions with positive claim ratios are omitted and serve as a buffer category. This is the case for 118 counties, 84 of them having claim ratios below 1%. It is reasonable to assume that the flood had a minor impact there. Nevertheless, it cannot be completely ruled out that this minor impact has a direct effect on insurance uptake which is the reason for excluding these counties from the analysis. Figure 3.5 shows the flooded regions used to construct the social connectedness to the flood, the buffer category and the treated regions that enter the model. Based on the method imposed by Bailey et al. (2018a), I define the flood experience of the social network of country i by:

$$SCI\ Flood_i = \frac{1}{J} \times \sum_{j \in J} (SCI_{i,j} \times Claim\ Ratio_j) \tag{3.2}$$

It defines the average flood-weighted social connectedness of each non-flooded region i to all flooded regions j, weighted by the claim ratio in county j as a proxy for the intensity of the flood. Figure 3.6 illustrates the flood-weighted social connectedness by quintiles. Darker shaded areas show larger social ties into the flooded areas. White areas depict the buffer category, the two red colored regions are Euskirchen and Ahrweiler. Most of the highly connected areas are located close to the flooded areas in the states of Rhineland-Palatinate or Hesse. Some regions in the north and south of Germany also have strong social ties to the flooded areas. Large parts of North Rhine-Westphalia, the neighboring state to the north are part of the buffer category. The distribution is similar to the SCI for the two flooded regions illustrated in Figures 3.3 and 3.4.

3.3.2 Baseline Estimation

The goal of this paper is to find out whether people that have not been hit by the flood react on their peer's experience and adapt their insurance coverage regarding elemental damages. Using the constructed measure of social connectedness to the flooded areas in 2021 rests on

[6]Data are published in their annual reports (e.g. GDV (2022)) and via an online dashboard on their webpage.

[7]Regions with less than 10 claims are also assumed to have a claim ratio of 0%.

several assumptions. To begin with, peer groups are not formed randomly and dependencies between treatment composition and the distribution of the main dependent variable of interest might impose a threat to identification (Manski, 1993). In the context of the flood, regions which are connected more strongly to the flooded areas must not be fundamentally different from regions that are less connected. In a classical difference-in-difference setting one would plot trends of the dependent variables before the flood for the treatment and control group to show that the common trend assumption holds. In this setting, there is no control group since social connectedness is never zero. However, the nature of the flood helps to argue that the assumption holds. The flood was caused by heavy, persistent rainfall. The topography of the disaster areas lead to flash floods with masses of water flowing down steep, v-shaped valleys, resulting in unprecedented water levels for small rivers (Junghänel et al., 2021). Unlike regions at the major rivers such as Rhine, Elbe or Danube, Euskirchen and Ahrweiler were poorly prepared and had little experience in dealing with large water masses. Apart from a few hours before the disaster, it was not clear where and when these large amounts of rain would occur. This makes the flood a suitable natural experiment in which ex-ante, better connected people are unlikely to have learned something about the importance of insurance coverage against elemental damage through their social network. There is also no selection into treatment, as it was not clear beforehand which counties would be affected. Further, all regions are treated at the same time. Therefore, the following regression can identify the effect of the social connectedness to the flooded regions:

$$Y_{it} = \beta_0 + \beta_1 \times log(SCI\ Flood)_i \times PostFlood_t + \gamma_i + \delta_t + \epsilon_{i,t} \tag{3.3}$$

Y_{it} is the outcome variable for county i at time t. Figure 3.2 suggests that the flood did not trigger any changes in the overall number of policies. This is also unlikely due to the fact that contracts can be only terminated at the end of the year or if the building is sold. Potential effects most likely occur directly after the flood. At this time people can only choose to add coverage against elemental damages to their existing policy. To test whether this is the case, the share of policies including elemental damages enters as the main dependent variable. In later analyses, the share of premiums that can be specifically attributed to cover these risk is also used, as well as the share of coverage that can be attributed to policies including elemental damages. The independent variable is an interaction of the log of the flood-weighted SCI and a dummy *PostFlood* which switches on in July 2021, the month in which the flood hit. γ_i represents county-level fixed effects. These account for all time-invariant factors such as distance to the flooded regions, general level of insurance against elemental damages[8] and the Facebook usage which determines the general level of social connectedness. δ_t represents the

[8]The general level varies strongly across federal states. Historically, elemental damage has been a mandatory part of residential property insurances in Baden-Wurttemberg. Therefore even in 2020 coverage equals 94% on average. In contrast to that, coverage in Lower Saxony, Bremen, Hamburg and Mecklenburg Western Pomerania yields only around 30%.

time-fixed effects on year-month level, controlling for general, county-invariant time trends, e.g. the economic cycle or trends in the German insurance market.

Direct effects of disaster trigger local insurance uptake. In combination with an overall positive trend as shown by the green line in Panel A of Figure 3.2 this raises a concern: Regionally different time trends might represent time-varying confounders. Instead of directly turning towards a triple difference-in-difference design (Wing et al., 2018; Olden and Møen, 2022), more restrictive, region-specific time trends are introduced. In this case, δ_t is a state-year-month fixed effect, which is equal to including state-dummies. $\epsilon_{i,t}$ represents the error term. To minimize the influence of other natural hazard events, I restrict the panel to run from January 2019 to December 2023. This rules out potential effects from flooding events in 2018. Table 3.1 displays descriptive statistics for all variables.

3.3.3 Results

The flooded areas show claim ratios around 20% in the aftermath of the flood. Insurance coverage against natural hazards is therefore a dominant topic in these regions. This propagates via social connectedness and people in unaffected but connected counties might learn about the importance of insurance coverage in case of a natural disaster. The share of policies including elemental damages is therefore expected to increase with social connectedness to the flood. Table 3.2 presents the results. Columns (1) and (2) show a statistically highly significant effect of social ties to the flooded areas on the share of policies including elemental damages. The rate of elemental damage coverage varies strongly across Germany and is partly dependent on former state laws. Applying the more restrictive state-month-year fixed effects reduces the coefficient by 50% to 0.012. However, it is still highly statistically significant. It indicates an increase of 0.012% in the share of policies including elemental damages in case of 1% more social connectedness to the flooded areas. The results hold for the spatial aggregation on postcode level as shown in Columns (3) to (5). Higher social connectedness is associated with a statistically significant increase in the share of policies including elemental damage. This effect is driven by people that opt for the additional coverage rather than new contracts. The design of the residential building insurance supports this finding. Customers can terminate their contracts at the end of the year or if the entity is sold. The only way to react immediately after the disaster is to opt for the supplementary module against natural hazards. A complete change of the policy or the insurance provider is not possible. The effect is expected to be particularly strong right after the flood. By design, two potential concerns can also be ruled out: First, the effect does not seem to be driven by people moving from flooded to non-flooded areas. Second, the fact that the analysis uses data from one insurance company only does not impact the results.

3.3.4 Discussion of Assumptions

This section discusses the assumption on which the identification of the effect of social connectedness on insurance uptake rests. Crucial for the implemented difference-in-difference method is the assumption of parallel trends for units that receive a low dosage and units that receive a high dosage of treatment. Addressing this assumption, the previous section argued that regions which are stronger connected to the flooded areas should not be fundamentally different from regions that are less connected. Plotting pre-trends for a treatment and a control group is impractical since there is no untreated control group in a continuous difference-in-difference model. Nevertheless, one can check whether there is an effect of social connectedness on the share of policies including elemental damages before the flood. The flood was unexpected and apart from hours before the rain started it was entirely unknown at what time and where heavy rain would occur. Based on data for previous hazard events, there had not been a similar event hitting the exact same regions (GDV, 2022) before either. This makes the flood as good as randomly assigned and eliminates all plausible reasons for people to top up their insurances with additional coverage based on their peers' experiences prior to the flood.

Figure 3.7 plots the coefficients of flood-weighted social connectedness for each month. The underlying regression includes state-year-month fixed effects to absorb the positive time trend. The plotted estimates show a significant jump at the time of the flood and a stable but slightly negative trend before. The significantly higher number of policies after the flood is persistent. Figure 3.8 illustrates the coefficients of the same regression but instead of the share, the monthly change in the number of policies including elemental damages serves as dependent variable. The effect is expected to be the largest shortly after the event. This specification identifies the deviation from the average change caused by social ties to the flooded areas by month. As expected, periods before the flood do not show significant effects of social connectedness. In July 2021, when the flood occurred, the coefficient for social connectedness is positive and significantly larger than zero. The magnitude almost doubles in August 2021. After August the effect decreases again, gradually declining over the next months and stabilizing around zero in January. This shows that the effect lasts three to five months. Results are in line with Gallagher (2014) who shows that insurance uptake can be explained by a Bayesian learning model, discounting past floods. Concentrating on directly affected regions he also finds that uptakes spike right after the flood and decline steadily afterwards. Atreya et al. (2015) explore a similar effect for flood insurance purchases in the state of Georgia, USA. They find that the effect fades after three years. Since I look at non-flooded regions and concentrate on social ties only, the shorter period of three to five months seems plausible.

3.4 Alternative Explanations

The empirical results show that people living in regions that have not been hit by the flood complement their existing policy with additional insurance against natural hazards. The previous section already discussed the underlying assumptions of the difference-in-difference model. Nevertheless, several alternative explanations might still be possible: Social connectedness to the flooded regions might serve as a proxy for an omitted variable that drives insurance decisions. Distance, which is highly correlated with the SCI, could also be the major determinant of reactions to the flood, being picked up by the SCI. Social ties could also coincidentally be correlated with characteristics that trigger an increase in insurance uptake that are not at all related to social connectedness. This section discusses each of these potential threats and presents further robustness checks for the baseline result.

3.4.1 Correlation of Treatment and Omitted Variables

To see whether the SCI is picking up other factors that might drive insurance-uptake, I generate placebo floods and aggregate the social connectedness to these floods as an additional control variable. To do so, the distribution of claim ratios of the actual flood is randomly distributed across all counties in Germany. For this artificial flood, I generate the flood-weighted social connectedness for each unflooded region to the flooded regions. Repeating this procedure 10,000 times produces a measure of social connectedness based on these placebo floods which is added as a second difference-in-difference term to Equation 3.3. Table 3.3 presents the results. It also expands the analysis to the share of premiums paid for the additional natural hazard insurance and the share of the overall coverage that can be attributed to policies that include elemental damages. Columns (1) to (3) show the results for county and year-month fixed effects, Columns (4)-(6) apply county and the more restrictive state-dependent time fixed effects.
The coefficients for the SCI to the real flood remains stable compared to the baseline result for the share of policies including elemental damages in Columns (1) and (4). The same holds for the share of coverage in Columns (3) and (6). The effect for the share of premiums attributable to the additional module for natural hazards in Columns (2) and (5) is smaller and less significant. The reason is most likely that only a relatively small fraction of the premiums is paid for the additional natural hazard insurance. In addition, increases in premiums due to higher construction prices are included into the general part of the premiums and cause distortions. Coefficients for the artificially generated social connectedness to the randomly drawn floods are small and statistically insignificant. The distribution of the SCI across German counties does not seem to proxy for some otherwise omitted variables driving the general result.

3.4.2 The Role of Distance

Distance is a major determinant of social connectedness. Bailey et al. (2018b) find that NUTS2-regions which are 10% further apart have a 13% lower social connectedness. Hu (2022) rules out that the effect of social ties on insurance-uptake is entirely driven by distance by maintaining a 750 mile buffer around flooded areas. This approach cannot be replicated in the German context, due to the country's size. However, I present several analyses concerning distance to rule out that it drives the baseline findings. First of all, region fixed effects absorb distance effects because distance is time-invariant. It might, however, become a crucial factor through social connectedness after the flood. Table 3.3 presents the results when including the flood-weighted distance as a separate interaction term. The coefficients are small and statistically insignificant for the case of the restrictive set of fixed effects. Distance seems not to determine the positive effect of social ties. Table A3.2 confirms this by showing that there is also no heterogeneity of the effect with respect to distance by including a triple difference-in-difference term. In a second robustness check, social connectedness is adjusted for distance by purging the SCI by distance. Distance has a strong negative and significant effect on social ties and explains already more than 55% of variation in the data (see Table A3.3). The residuals of regressing the SCI on distance are used as a measure of pure social connectedness to check whether the results hold for a distance-purged version of flood-weighted social connectedness. Quintiles of the purged SCI to the flooded regions are illustrated by Figure A3.1. Counties in the north and east of Germany significantly increase their connectedness while it decreases for counties close to the flooded areas. Results using the purged SCI presented in Table A3.4 are almost identical to the baseline findings. All in all, distance does not significantly impact the results.

3.4.3 Past, Current and Future Flood Exposure

I complement the analysis by investigating the role of current risk of flood, past events, climate projections into the future and geographic characteristics, adding these as interaction terms to Equation 3.3. Table 3.4 presents the results. Columns (1) and (2) reveal that there is no significant effect of the current flood risk as measured by Hincks et al. (2023). In Column (1) fluvial hazard measures the percentage of land cover that is expected to be flooded in a 1 in 100 year fluvial flood. Column (2) includes the share of population living in settlements that are prone to flooding in such an event. Both interaction terms yield coefficients close to zero and do not significantly impact insurance uptake. To control for general natural hazard events and their average local costs between 2002 and 2022, data from the German Insurance Association (GDV, 2022) is used. Again, the effect is almost zero, as is depicted in Columns (3) and (4). The same holds true for the number of heavy rain events and their respective claim ratios over the 20 year preceding the flood in Columns (5) and (6). The average damage costs resulting from heavy rain events between 2002 and 2021 yields a marginally significant and positive effect on insurance uptake as shown in Column (7). However, this does not change the effect of social ties. As insurance uptake is equal to hedging against future risks, climate projections might influence the decision to buy insurance. Column (8) shows a small but insignificant coefficient, similar to Column (9) which controls for the length of rivers. The coefficient is mirrored in Column (10), accounting for the share of a region that is categorized as mountains. Although insignificant, rivers are positively associated with higher insurance uptake, while having a higher share of mountainous area decreases the demand for policies including elemental damages after the flood. The effect of social connectedness fluctuates between 0.11 and 0.13. It is remarkably stable over all specifications as well as in Column (11), controlling for all potential confounders. It seems not to be driven by differences in current hazard risk, past events, future projections or geographic factors.

3.4.4 Robustness

Social connectedness to the flooded areas is constructed by weighting the SCI with claim ratios that serve as proxies for the intensity of the flood in 2021. Table A3.5 uses an unweighted social connectedness to show that the results do not rely on this weighting scheme. The effect holds across both spatial aggregation levels and the different sets of fixed effects.

Because all county pairs have a positive value for the SCI, the baseline setting lacks a control group. It estimates a continuous treatment comparing regions that are more connected with those that are less connected to the flooded regions. Discretization of the continuous dependent variable is a way to construct a treatment and a control group. To see if the results hold in a classical two-by-two framework one needs two time periods, before and after the flood, and two groups, a treatment and a control group. While the timing is no issue, dividing all counties that did not experience the flood in a treatment and a control group requires a strong assumption. I split the sample at the median of social connectedness and define counties with above-median connectedness as the treated group and those with a below-median connectedness as the control group. The general results hold, although coefficients are smaller compared to the baseline (see Table A3.6). Note that the separation in the two groups introduces arbitrariness. It also leads to the loss of information compared to the continuous setting (D'Haultfœuille et al., 2023).

Many robustness checks and specifications are driven by the idea that the SCI might proxy for some omitted factors. Social ties to the flooded areas should only have an influence on insurance uptake after the flood. According to the GDV, no major natural hazard events occurred in 2019. To conduct a placebo-test, I define a dummy for a placebo-flood that turns one in July 2019. I regress the shares of the three dependent variables from July 2019 on the interaction term of the SCI, weighted and constructed using the real flood in 2021. In addition, the panel is shifted to run from January 2019 to June 2021. The results for this placebo test are presented in Table A3.7. One would expect to find no effects at all. However, when using county and year-month fixed effects, coefficients are positive and significant. This clearly shows that the SCI picks up some of the general time trend. As soon as one implements more restrictive state-dependent time fixed effects, the coefficients drop towards zero and turn insignificant.

The buffer zone consists of all counties with positive claim ratios below 15% to rule out direct effects. I extend this buffer zone in two different ways. First, all counties within the same state as Euskirchen and Ahrweiler are also added to the buffer zone and do not enter the sample. That also removes all counties that are located in the same local media market as the flooded regions from the sample. The results in the first three columns of Table A3.8 are comparable to the baseline effect. Coefficients increase in size but decrease in terms of statistical significance. Second, following Hu (2022), I exclude all counties that lie within 100km of the flooded regions. Columns (4) to (6) report the results which also substantiate the baseline effect. This approach ensures that there is no distortion due to spatial proximity. Excluding all regions within a 100km radius rules out that people are directly influenced e.g. because they commute through the flood zone without having social ties to locals.

For the main results, only Euskirchen and Ahrweiler with claim ratios of 23.9% and 18.3% are defined as flooded. The assumption that only these two regions were sufficiently affected by the flood has both advantages and disadvantages: Choosing a high threshold in terms of claim ratios ensures that the flood experience is strong enough that information can transmit via social ties. On the other hand, the omitted number of counties is fairly large. Regions might have a small connectedness to the two heavily affected regions but a high connectedness to regions just below the threshold. This would bias the results downwards. Additionally, if regions with strong ties to the two heavily affected regions are structurally different to other regions, this would bias the results upwards.

The following specifications use a threshold of 5% and 1% to define regions as flooded. Figure A3.2 presents the respective sample, Figure A3.3 plots the flood-weighted SCI by quintiles in case of the 5% threshold. Using this definition, all flooded regions are still located close to each other, located in the two neighboring states of North Rhine-Westphalia and Rhineland-Palatinate. Although, in one case, flooded and non-flooded counties even share a common border. Figure A3.4 presents the respective sample and Figure A3.5 plots the flood-weighted SCI by quintiles for the threshold of 1%. In this case many common borders exist and the flooded regions are not clustered anymore. Some regions in the north, east and south of Germany are also defined as flooded. The distribution of social ties to the flooded counties is therefore expected to be smoothed, which should dilute the results. Columns (1) to (3) of Table A3.9 refer to the 5% sample, Columns (4) to (6) present coefficients for the 1% sample. Results still hold in terms of size with reduced statistical significance for the 5% threshold. The effect vanishes when the sample gets more scattered across Germany. Social contacts seem to become blurred and the lower intensity of the disaster could lead to fewer signals

being transmitted to socially connected regions.
Table A3.10 tests these results by considering Ahrweiler and Euskirchen separately. Columns (1) to (3) assume only Ahrweiler to be flooded, Euskirchen enters the buffer zone and is not taken into account when the SCI to the flood is calculated. The opposite holds true for Columns (4) to (6). Compared to the baseline effect of the SCI on insurance uptake is smaller for Ahrweiler and larger for Euskirchen. Given the larger claim ratio for the latter, this result is in line with previous findings and suggests that the effect is monotonic with respect to the intensity of the flood. All in all, results hold across different definitions of the explanatory variable.

3.5 Channels

Higher social connectedness to flooded regions is linked to more insurance policies that include elemental damage coverage. The effect is driven by people deciding to include elemental hazards in their residential building insurance. This section explores the mechanism behind these peer effects. The vast majority of policies is active since the owner built or bought the property and remains mostly unchanged. The majority of property-related risks only change if the owner actively modifies the property. That includes conversions, extensions or the installation of solar systems. Natural hazard risk imposes an exception. The underlying change of risk without the homeowner's active involvement often leads to the insurance cover not being adjusted. Literature has identified insurance uptake after people have been directly affected by a disaster (Botzen and van den Bergh, 2012; Atreya et al., 2015; Gao et al., 2020). Linking insurance uptake to a Bayesian learning model, Gallagher (2014) explains the discrepancy between low insurance rates and the low cost to acquire information about the own flood risk. Following Hu (2022), indirect effects via flood-exposure can be explained by an information update. Using the staggered digitization of flood-risk maps and social connectedness to flooded areas more than 750 miles away, he shows two distinct channels: First but probably less intense, people receive information about flood insurance policies through their peers that are otherwise costly to obtain. Second and more important, experiences of peers trigger attention towards a neglected risk. Without explicitly exchanging information about insurance products, people feel the urge to review their own risk and coverage. Individual data for the connections between people is not available, which makes it impossible to disentangle the direct and indirect effect. The following analysis is in line with information updates via social learning as the predominant mechanism. This paper does not focus on the mechanism but explores two distinct characteristics of regions that moderate the responsiveness to the identified peer effects: First, it shows that the influence of peer effects is shaped by local attitudes toward climate protection policies. Second, it demonstrates that the degree of responsiveness to these peer effects is largely influenced by the level of local social capital within the community.

3.5.1 Climate Policy Attitudes

Approval of climate protection measures varies widely between regions. I use heterogeneity in climate policy attitudes to identify regions that are more responsive to peer effects in the aftermath of natural hazards. Ex-ante it is unclear that higher approval rates towards mitigation policies are associated with larger insurance coverage against elemental damages. Flood risk perception might be correlated with climate policy approval (Botzen et al., 2016; Frondel et al., 2017). However, especially for locations at the tail of the flood risk distribution, many other factors have been found influential such as past disaster experiences (Gallagher, 2014), affordability (Thomas and Leichenko, 2011; Tesselaar et al., 2020), or social norms (Lo, 2013; Wilson et al., 2020). There are two potential answers for how climate policy attitudes influence peer effects. The recognition of increasing natural hazards might lead to stronger reactions after a shock due to higher risk perceptions. On the contrary, higher awareness about climate risk might lead to insurance updates since risks are more salient and households evaluate them more frequently. Shocks, directly or indirectly, then have an attenuated effect. The crucial question is whether climate policy attitudes only have a significant effect on risk perception as a result of the shock transmitted via peer effects, or whether they exert a strong influence on insurance coverage even in the absence of such shocks.

To proxy climate policy attitudes, local and state-level election data and policy preferences on county level are used. Empirically, these moderating factors are tested by including a triple interaction term into Equation 3.3 (Gruber, 1994; Olden and Møen, 2022). Table 3.5 presents the results. The election data for political representatives used at district level in Column (1) and on state level in Column (2) paint a similar picture. Regions with higher voting shares for the green party show significantly lower peer effects. A one standard deviation change in green votes reduces the peer effects by 20%-35%. At the same time, regardless of social connectedness, higher green party votes are associated with a higher share of insurance policies including elemental damages after the flood. This suggests that positive attitudes towards climate policies shape hedging behavior against natural hazard risks in two ways. First, they are linked to larger insurance uptake in general. Second, they are related with a reduction of peer effects.

The results are consistent when using more tailored measures of climate policy attitudes. By evaluating specific climate policies such as people's willingness to pay for fossil fuels, intention to invest in renewable energy or willingness to protest against local wind power plants, sentiment with respect to climate policies is measured very precisely. Columns (3) to (5) present the estimates. The observed patterns remain the same. Willingness to pay more for fossil fuels and especially plans to invest in renewable energy are linked to higher insurance uptake in general but lower peer effects. The reverse is true for regions with higher willingness to protest against local wind power plants. The more opposition there is to locally visible climate protection, the smaller the general reaction to the flood. Yet, peer effects are significantly larger: For an increase of one standard deviation in willingness to protest against local wind power plants in the mean connected county, peer effects increase by 3.1%. All specifications control for GDP per capita and purchasing power to rule out that results are driven by income. Regional differences in climate policy attitudes lead to heterogeneity in peer effects after the flood. Thus, understanding and recognition of the need for more sustainable resources softens the reaction in the unaffected regions. On the contrary, places with more resistance to climate protection measures are experiencing a stronger increase in policies with natural hazards coverage. The results are in line with information updates via social learning being the predominant mechanism. People receive information about a disaster through different channels and reassess their own flood risk. The strength of this information update varies with overall climate change policy preferences.

3.5.2 Social Capital

The ability and openness to absorb information play a major role when it comes to the strength of peer effects in insurance decisions. One could rephrase the findings of the previous subsection and state: The further away my peers' experiences are from what I perceive as climate risks, the greater the information update. The perception of climate risk is dependent on people's participation in the public debate. Increasing weather extremes and climate protection measures are part of the public debate and an almost perennial topic in the media. The literature identifies a large number of factors that influence belief in anthropogenic climate change (see Hornsey et al. (2016) for an overview). Druckman and McGrath (2019) highlight the ability to filter information with respect to credible evidence as an important factor. This subsection uses the concept of social capital to identify heterogeneity in regions' responses to socially connected floods. Social capital can be categorized into two types: Bridging social capital (Putnam, 1994, 2000), sometimes also referred to as linking capital (Guiso et al., 2011), or civic capital (Lichter et al., 2021), has been shown to result in more socially responsible behavior in general (Bartscher et al., 2021) and with respect to climate-mitigation in particular (Hao et al., 2020; Kyne and Aldrich, 2020). Bridging social capital fosters trust in strangers and facilitates the exchange of information beyond one's immediate social circle. It is also linked to cross-group solidarity. For example, memberships in sports clubs, which are often used as a proxy for bridging social capital (Tacon, 2021), provide members with access to a broad range of resources and opportunities to gather information. In contrast, bonding social capital strengthens solidarity within a specific group, promoting close-knit relationships and mutual support among members. This type of social capital typically emerges in tightly connected communities or networks, where trust and cooperation are reinforced through shared experiences and strong emotional ties, supporting within-group solidarity.

Again, it is ex-ante unclear how social capital is connected to peer effects via social ties. However, the results from the previous subsection suggest that bridging social capital leads to better information about climate risk and mitigation behavior which results in weaker peer effects. Table 3.6 tests this by introducing a triple interaction term with voter turnout in 2017, the last federal election before the flood. In addition, Covid-19 vaccination rates per capita from June 2021 are used.[9] One reason to get vaccinated was to be allowed to participate in public life. Vaccinations were heavily promoted by the government and health care institutions. The active choice to get vaccinated can therefore be interpreted as trusting experts and following authorities which is a known characteristic of bridging capital (Myeong and Seo, 2016). The results show that, conditional on social connectedness, peer effects decline with larger bridging social capital. A the same time insurance uptake after the flood, independently of social connectedness, increases in regions with larger voter turnout and vaccination rates. Again, all specifications control for GDP per capita and purchasing power to rule out that results are driven by income.

To explore the moderating effect of bonding capital, a measure of within social connectedness is constructed:

$$SCI\ Dispersion_i = \frac{\frac{1}{J}\Big(\sum\limits_{j \in J}(SCI_{i,j}) + (SCI_{i,i})\Big)}{\frac{1}{K}\sum\limits_{k \in K}(SCI_{i,k})} \tag{3.4}$$

Equation 3.4 divides the average social connectedness of county i to the own and all adjacent counties j to the average social connectedness to all other counties k. This compares the probability to be connected with individuals living nearby to the probability to be connected to individuals in more distant counties. Adding a triple interaction term with this measure of bonding capital confirms prior findings. Table 3.6, Column (4) shows that less dispersed social connectedness is associated with larger peer effects. The same holds true for educational equality as a proxy for bonding capital in Column (4) (Goldin and Katz, 1999; Kyne and Aldrich, 2020). Less diverse educational attainments in local societies result in stronger bonding capital which is linked to stronger peer effects.

The findings are in line with the results regarding climate policy attitudes as well as the literature on bridging and bonding capital with respect to responses to climate change (Hao et al., 2020). The more limited people's abilities to resource information from outside of their own circles, the larger is the increase in policies with natural hazards coverage. Identifying social capital as a moderating factor is a contribution to the literature on disaster

[9] Data represents the number of vaccinations per capita for the population between 18 and 59 years as it is used in Bade et al. (2024). Case rates or infections are not chosen as a measure since there is no clear evidence on the relationship between social distancing and bridging or bonding capital (Gibbons et al., 2022).

preparedness. It shows that first-hand experiences by socially connected regions are able to convince people who are not persuaded by warnings from experts to take disaster prevention measures. A stronger bond with the people affected, and thus the confrontation with their situation, replaces trust in the government and institutions.

3.6 Conclusion

After major natural hazard events, the discussion regularly arises how disaster preparedness and prevention measures can be implemented and improved. One of the most important components are insurance policies for residential buildings that include natural hazard risks, especially since residential property makes up the majority of household assets. In addition, climate researchers agree that heavy rainfall events will increase in the future (GDV, 2022). This paper uses the flood in Germany in July 2021 to study peer effects in insurance uptake. Heavy rainfall has caused small rivers to overflow and led to a disaster that is classified as a once-in-a-century event (Junghänel et al., 2021). Two regions were hit particularly hard, Euskirchen and Ahrweiler, both of which had never experienced a flood of this magnitude before. I compute a measure of flood-weighted social connectedness to the flooded regions for each unflooded county. To do so, I use Facebook's Social Connectedness Index (Bailey et al., 2018b), weighted by the flood intensity as measured by the local insurance claim ratio (GDV, 2022). I estimate a difference-in-difference approach with the flood weighted connectedness as a continuous treatment. The results show a positive and highly significant effect of social ties on insurance uptake after the flood. While the overall number of insurance policies does not change, regions with stronger social connections to the flooded areas exhibit an increase in the share of policies that include elemental damages in their policies. The findings are robust across different definitions of the flooded regions and to alternative model choices. Distance and correlated shocks can be ruled out as alternative explanations. The results show that a 1% higher social connectedness is associated with an increase of 0.012% of residential building insurances that include the coverage of elemental damages.

Climate policy attitudes and social capital are exploited to identify heterogeneity in the responsiveness to peer effects. Firstly, regional differences in climate policy acceptance is used. Votes for the green party, willingness to pay higher cost for fossil fuels, planned investments into renewables and protests against local wind power plants show that peer effects are less pronounced in regions that show a stronger aversion to climate protection measures.

Social capital is known to determine how people access information. While bridging capital is associated with trust in the government and larger ability to use external resources, bonding capital strengthens the focus on the immediate environment and people who are similar to oneself. Voter turnout and Covid-19 vaccination rates are used to proxy bonding capital. The results indicate that peer effects diminish in the presence of bridging social capital. In contrast, peer effects intensify with bonding social capital, which is proxied by a dispersion measure from the SCI and educational equality. The more cohesive and homogeneous local networks are, the higher is the rate of insurance adoption following the flood. The paper identifies moderating factors for peer effects in disaster preparedness, showing that social ties can spur insurance uptake where the trust in the government is low and expert warnings fail.

Bibliography

Agarwal, S., Qian, W., and Zou, X. (2021). Thy neighbor's misfortune: Peer effect on consumption. American Economic Journal: Economic Policy, 13(2): 1–25.

Aref, S., Bontcheva, K., Braghieri, M., Dignum, F., Giannotti, F., Grisolia, F., and Pedreschi, D., editors (2020). Social Informatics: 12th International Conference, SocInfo 2020, Pisa, Italy, October 6–9, 2020, Proceedings, volume 12467 of Lecture Notes in Computer Science. Springer International Publishing, Cham.

Atreya, A., Ferreira, S., and Michel-Kerjan, E. (2015). What drives households to buy flood insurance? New evidence from Georgia. Ecological Economics, 117: 153–161.

Bade, V., Schmitz, H., and Tawiah, B. B. (2024). Regional variations in vaccination against COVID-19 in Germany. PLOS ONE, 19(4): e0296976.

Bailey, M., Cao, R., Kuchler, T., and Stroebel, J. (2018a). The economic effects of cocial networks: Evidence from the housing market. Journal of Political Economy, 126(6): 2224–2276.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018b). Social connectedness: Measurement, determinants, and effects. Journal of Economic Perspectives, 32(3): 259–280.

Bailey, M., Dávila, E., Kuchler, T., and Stroebel, J. (2019). House price beliefs and mortgage leverage choice. The Review of Economic Studies, 86(6): 2403–2452.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., and Stroebel, J. (2021). International trade and social connectedness. Journal of International Economics, 129: 103418.

Bakkensen, L. A. and Barrage, L. (2022). Going underwater? Flood risk belief heterogeneity and coastal home price dynamics. The Review of Financial Studies, 35(8): 3666–3709.

Balakina, O., Backman, C., Hackethal, A., Hanspal, T., and Lammer, D. M. (2022). Good peers, good apples: Peer effects lead to better financial outcomes. Working Paper.

Bartscher, A. K., Seitz, S., Siegloch, S., Slotwinski, M., and Wehrhöfer, N. (2021). Social capital and the spread of Covid-19: Insights from European countries. Journal of Health Economics, 80: 102531.

Bauer, R., Eberhardt, I., and Smeets, P. (2022). A fistful of dollars: Financial incentives, peer information, and retirement savings. The Review of Financial Studies, 35(6): 2981–3020.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Milkman, K. L. (2015). The effect of providing peer information on retirement savings decisions. The Journal of finance, 70(3):1161–1201.

Billings, S. B., Gallagher, E. A., and Ricketts, L. (2022). Let the rich be flooded: The distribution of financial aid and distress after Hurricane Harvey. Journal of Financial Economics, 146(2):797–819.

Botzen, W. J. W. and van den Bergh, J. C. J. M. (2012). Risk attitudes to low-probability climate change risks: WTP for flood insurance. Journal of Economic Behavior & Organization, 82(1): 151–166.

Botzen, W. J. W., Michel-Kerjan, E., Kunreuther, H., De Moel, H., and Aerts, J. C. J. H. (2016). Political affiliation affects adaptation to climate risks: Evidence from New York City. Climatic Change, 138(1-2): 353–360.

Brei, M., Mohan, P., and Strobl, E. (2019). The impact of natural disasters on the banking sector: Evidence from hurricane strikes in the Caribbean. The Quarterly Review of Economics and Finance, 72: 232–239.

Cai, J., De Janvry, A., and Sadoulet, E. (2015). Social networks and the decision to insure. American Economic Journal: Applied Economics, 7(2): 81–108.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. C. (2024). Difference-in-differences with a continuous treatment. Working Paper.

Cannon, B., Hirshleifer, D., and Thornton, J. (2024). Friends with benefits: Social capital and household financial behavior. Working Paper.

Card, D. (1992). Using regional variation in wages to measure the effects of the federal minimum wage. Industrial and Labor Relations Review, 46: 22.

Celil, H. S., Oh, S., and Selvam, S. (2022). Natural disasters and the role of regional lenders in economic recovery. Journal of Empirical Finance, 68: 116–132.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022a). Social capital I: Measurement and associations with economic mobility. Nature, 608(7921): 108–121.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022b). Social capital II: Determinants of economic connectedness. Nature, 608(7921): 122–134.

Cookson, J. A., Engelberg, J. E., and Mullins, W. (2023). Echo chambers. The Review of Financial Studies, 36(2): 450–500.

De Giorgi, G., Frederiksen, A., and Pistaferri, L. (2020). Consumption network effects. The Review of Economic Studies, 87(1): 130–163.

D'Haultfœuille, X., Hoderlein, S., and Sasaki, Y. (2023). Nonparametric difference-in-differences in repeated cross-sections with continuous treatments. Journal of Econometrics, 234(2): 664–690.

Diemer, A. and Regan, T. (2022). No inventor is an island: Social connectedness and the geography of knowledge flows in the US. Research Policy, 51(2): 104416.

Dornseifer, F. and Rehbein, O. (2023). Social connectedness and foreign direct investment. Working Paper.

Druckman, J. N. and McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. Nature Climate Change, 9(2): 111–119.

Forte, A., Sahan, S., and Silipo, D. B. (2024). Do natural disasters reduce loans to the more CO2-emitting sectors? Sustainability, 16(10): 3943.

Frondel, M., Simora, M., and Sommer, S. (2017). Risk perception of climate change: Empirical evidence for Germany. Ecological Economics, 137: 173–183.

Gallagher, J. (2014). Learning about an infrequent event: Evidence from flood insurance take-up in the United States. American Economic Journal: Applied Economics, 6(3): 206–233.

Gallagher, J. and Hartley, D. (2017). Household finance after a natural disaster: The case of Hurricane Katrina. American Economic Journal: Economic Policy, 9(3): 199–228.

Gao, M., Liu, Y.-J., and Shi, Y. (2020). Do people feel less at risk? Evidence from disaster experience. Journal of Financial Economics, 138(3): 866–888.

GDV (2021). Juli-Flut sorgt für starke Nachfrage nach Versicherungen. https://www.gdv.de/gdv/medien/medieninformationen/juli-flut-sorgt-fuer-starke-nachfrage-nach-versicherungen-73148.

GDV (2022). Naturgefahrenreport 2022. Technical report, Gesamtverband der Deutschen Versicherungswirtschaft e. V.

Gibbons, J., Yang, T.-C., and Oren, E. (2022). Community boosts immunity? Exploring the relationship between social capital and COVID-19 social distancing. Spatial Demography, 10(1): 75–105.

Giglio, S., Kelly, B., and Stroebel, J. (2021). Climate finance. Annual Review of Financial Economics, 13(1): 15–36.

Goldin, C. and Katz, L. F. (1999). Human capital and social capital: the rise of secondary schooling in America, 1910-1940. The Journal of Interdisciplinary History, 29(4): 683–687.

Granovetter, M. (2011). The impact of social structure on economic outcomes. In The Sociology of Economic Life. Routledge, third edition.

Gruber, J. (1994). The incidence of mandated maternity benefits. The American Economic Review, 84(3): 622–641.

Guiso, L., Sapienza, P., and Zingales, L. (2011). Chapter 10 - Civic capital as the missing link. In Benhabib, J., Bisin, A., and Jackson, M. O., editors, Handbook of Social Economics, volume 1, pages 417–480. North-Holland.

Hao, F., Liu, X., and Michaels, J. L. (2020). Social capital, carbon dependency, and public response to climate change in 22 European countries. Environmental Science & Policy, 114: 64–72.

Heimer, R. Z. (2016). Peer pressure: Social interaction and the disposition effect. The Review of Financial Studies, 29(11): 3177–3209.

Hincks, S., Carter, J., and Connelly, A. (2023). A new typology of climate change risk for European cities and regions: Principles and applications. Global Environmental Change, 83: 102767.

Hirshleifer, D. (2020). Presidential address: Social transmission bias in economics and finance. The Journal of Finance, 75(4): 1779–1831.

Hirshleifer, D., Peng, L., and Wang, Q. (2024). News diffusion in social networks and stock market reactions. The Review of Financial Studies, page hhae025.

Hornsey, M. J., Harris, E. A., Bain, P. G., and Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. Nature Climate Change, 6(6): 622–626.

Hu, Z. (2022). Social interactions and households' flood insurance decisions. Journal of Financial Economics, 144(2): 414–432.

Jain, P. (2020). Imperfect monitoring and informal insurance: The role of social ties. Journal of Economic Behavior & Organization, 180: 241–256.

Junghänel, T., Bissolli, P., Daßler, J., Fleckenstein, R., Imbery, F., Janssen, W., Lengfeld, K., Leppelt, T., Rauthe, M., Rauthe-Schöch, A., Rocek, M., Walawender, E., and Weigl, E. (2021). Hydro-klimatologische Einordnung der Stark- und Dauerniederschläge in Teilen Deutschlands im Zusammenhang mit dem Tiefdruckgebiet „Bernd“ vom 12. bis 19. Juli 2021. Technical report.

Kariya, A. and Shekhawat, C. (2024). Distance lending & social connectedness. Journal of Financial Stability, 72: 101249.

Koetter, M., Noth, F., and Rehbein, O. (2020). Borrowers under water! Rare disasters, regional banks, and recovery lending. Journal of Financial Intermediation, 43: 100811.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). Social proximity to capital: Implications for investors and firms. The Review of Financial Studies, 35(6): 2743–2789.

Kuchler, T., Russel, D., and Stroebel, J. (2022b). JUE Insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook. Journal of Urban Economics, 127: 103314.

Kuchler, T. and Stroebel, J. (2021). Social finance. Annual Review of Financial Economics, 13(1): 37–55.

Kyne, D. and Aldrich, D. P. (2020). Capturing bonding, bridging, and linking social capital through publicly available data. Risk, Hazards & Crisis in Public Policy, 11(1): 61–86.

Landratsamt Kyffhäuserkreis (2023). Dankeschön an die Fluthelfer im Ahrtal. https://www.kyffhaeuser.de/dankeschoenan-die-fluthelfer-im-ahrtal/. Accessed: 2024-05-27.

Levi, S., Wolf, I., Sommer, S., and Howe, P. (2022). Local support of climate change policies is driven by spatial diffusion. Working Paper.

Lichter, A., Löffler, M., and Siegloch, S. (2021). The long-term costs of government surveillance: Insights from stasi spying in East Germany. Journal of the European Economic Association, 19(2): 741–789.

Lo, A. Y. (2013). The role of social norms in climate adaptation: Mediating risk perception and flood insurance purchase. Global Environmental Change, 23(5): 1249–1257.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The Review of Economic Studies, 60(3): 531–542.

Maturana, G. and Nickerson, J. (2019). Teachers teaching teachers: The role of workplace peer effects in financial decisions. The Review of Financial Studies, 32(10): 3920–3957.

McCartney, W. B. and Shah, A. M. (2022). Household mortgage refinancing decisions are neighbor influenced, especially along racial lines. Journal of Urban Economics, 128: 103409.

Myeong, S. and Seo, H. (2016). Which type of social capital matters for building trust in government? Looking for a new type of social capital in the governance era. Sustainability, 8(4): 322.

NapoleonCat (2023). Number of Facebook users in Germany from September 2018 to September 2022. https://www.statista.com/statistics/1017402/facebook-users-germany/.

Nguyen, D. D., Ongena, S., Qi, S., and Sila, V. (2022). Climate change risk and the cost of mortgage credit. Review of Finance, 26(6): 1509–1549.

Nguyen, G., Nguyen, M., Pham, A. V., and Pham, M. D. M. (2023). Navigating investment decisions with social connectedness: Implications for venture capital. Journal of Banking & Finance, 155: 106979.

Noth, F. and Rehbein, O. (2019). Badly hurt? Natural disasters and direct firm effects. Finance Research Letters, 28: 254–258.

Obschonka, M., Tavassoli, S., Rentfrow, P. J., Potter, J., and Gosling, S. D. (2023). Innovation and inter-city knowledge spillovers: Social, geographical, and technological connectedness and psychological openness. Research Policy, 52(8): 104849.

Olden, A. and Møen, J. (2022). The triple difference estimator. The Econometrics Journal, 25(3): 531–553.

Ouazad, A. and Kahn, M. E. (2022). Mortgage finance and climate change: Securitization dynamic in the aftermath of natural disasters. The Review of Financial Studies, 35(8): 3617–3665.

Ouimet, P. and Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. The Journal of Finance, 75(1): 133–172.

Putnam, R. D. (1994). Making democracy work: Civic traditions in modern Italy. Princeton university press.

Putnam, R. D. (2000). Bowling Alone: The Collapse and Revival of American Community. Simon and Schuster.

Ratnadiwakara, D. (2021). Flooded social connections. The Quarterly Journal of Finance, 11(04): 2150018.

Ratnadiwakara, D. and Venugopal, B. (2023). Climate risk perceptions and demand for flood insurance. Financial Management, 52(2): 297–331.

Rehbein, O. and Ongena, S. (2022). Flooded through the back door: The role of bank capital in local shock spillovers. Journal of Financial and Quantitative Analysis, 57(7): 2627–2658.

Rehbein, O. and Rother, S. (2024). Social connectedness in bank lending. Working Paper.

Segura, A. and Villacorta, A. (2023). Firm-bank linkages and optimal policies after a rare disaster. Journal of Financial Economics, 149(2): 296–322.

Tacon, R. (2021). Social Capital and Sport Organisations. Routledge, London, first edition.

Tesselaar, M., Botzen, W. J. W., Haer, T., Hudson, P., Tiggeloven, T., and Aerts, J. C. J. H. (2020). Regional inequalities in flood insurance affordability and uptake under climate change. Sustainability, 12(20): 8734.

Thomas, A. and Leichenko, R. (2011). Adaptation through insurance: Lessons from the NFIP. International Journal of Climate Change Strategies and Management, 3(3): 250–263.

Wagner, K. R. H. (2022). Adaptation and adverse selection in markets for natural disaster insurance. American Economic Journal: Economic Policy, 14(3): 380–421.

Wasko, C., Nathan, R., Stein, L., and O'Shea, D. (2021). Evidence of shorter more extreme rainfalls and increased flood variability under climate change. Journal of Hydrology, 603:126994.

Wilson, R. S., Herziger, A., Hamilton, M., and Brooks, J. S. (2020). From incremental to transformative adaptation in individual responses to climate-exacerbated hazards. Nature Climate Change, 10(3): 200–208.

Wing, C., Simon, K., and Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. Annual Review of Public Health, 39(Volume 39, 2018): 453–469.

Wong, M. C. S. and Ho, H. M. (2023). A framework for integrating extreme weather risk, probability of default, and loss given default for residential mortgage loans. Sustainability, 15(15): 11808.

You, X. and Kousky, C. (2024). Improving household and community disaster recovery: Evidence on the role of insurance. Journal of Risk and Insurance, 91(2): 299–338.

Figures

Figure 3.1: Claim Ratios

This maps illustrates claim ratios for the universe of policies against elemental damages in the German insurance industry (GDV, 2022) for the flood in July 2021. White shaded regions have a claim ratio of 0% or less than 10 cases. Dark red shading indicates claim ratios above 12%. This applies to Euskirchen (23.9%) and Ahrweiler (18.3%) only.

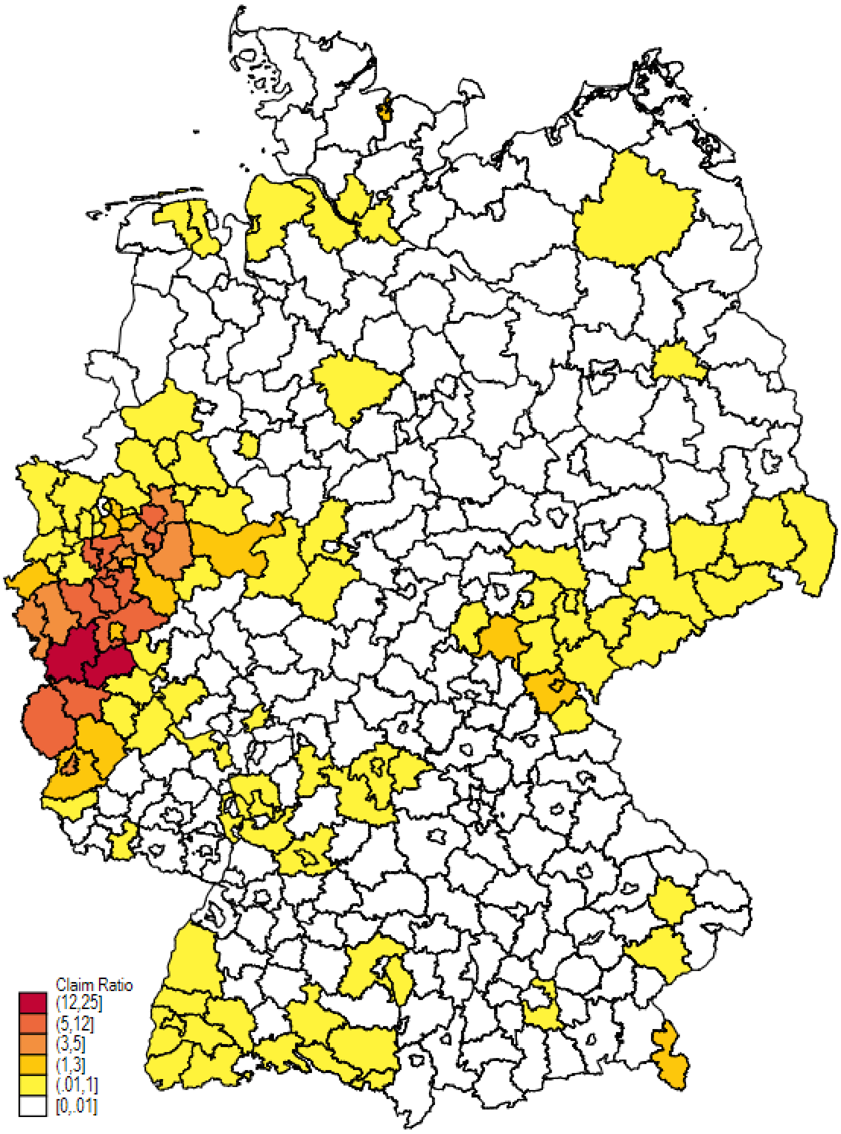
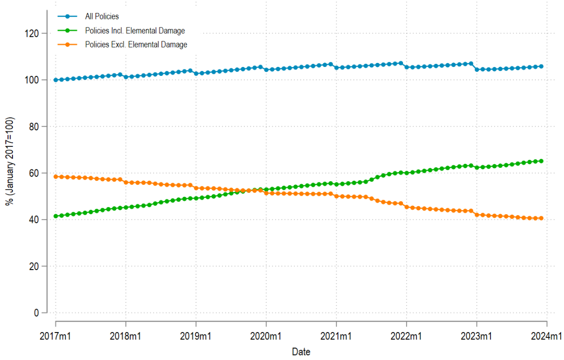


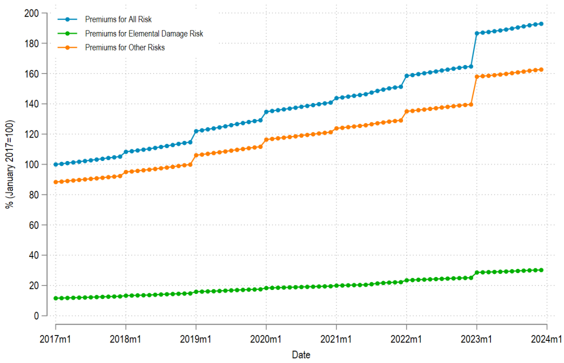
Figure 3.2: Insurance Data

This figure illustrates the insurance data obtained from a private German insurance company (DEVK) for Germany from 2017 to 2023. Panel A presents data on policies, Panel B on premiums and Panel C on the coverage with January 2017 normalized to 100.

Panel A: Share of Policies



Panel B: Share of Premiums



Panel C: Share of Coverage

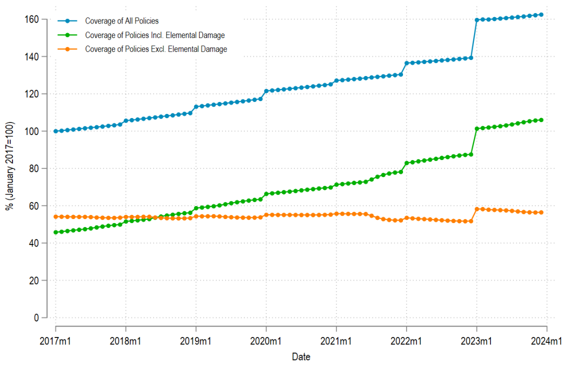


Figure 3.3: Social Connectedness Index Euskirchen

This maps illustrates the Social Connecedness Index for the county Euskirchen (white). Darker red shading indicates a higher connectedness.

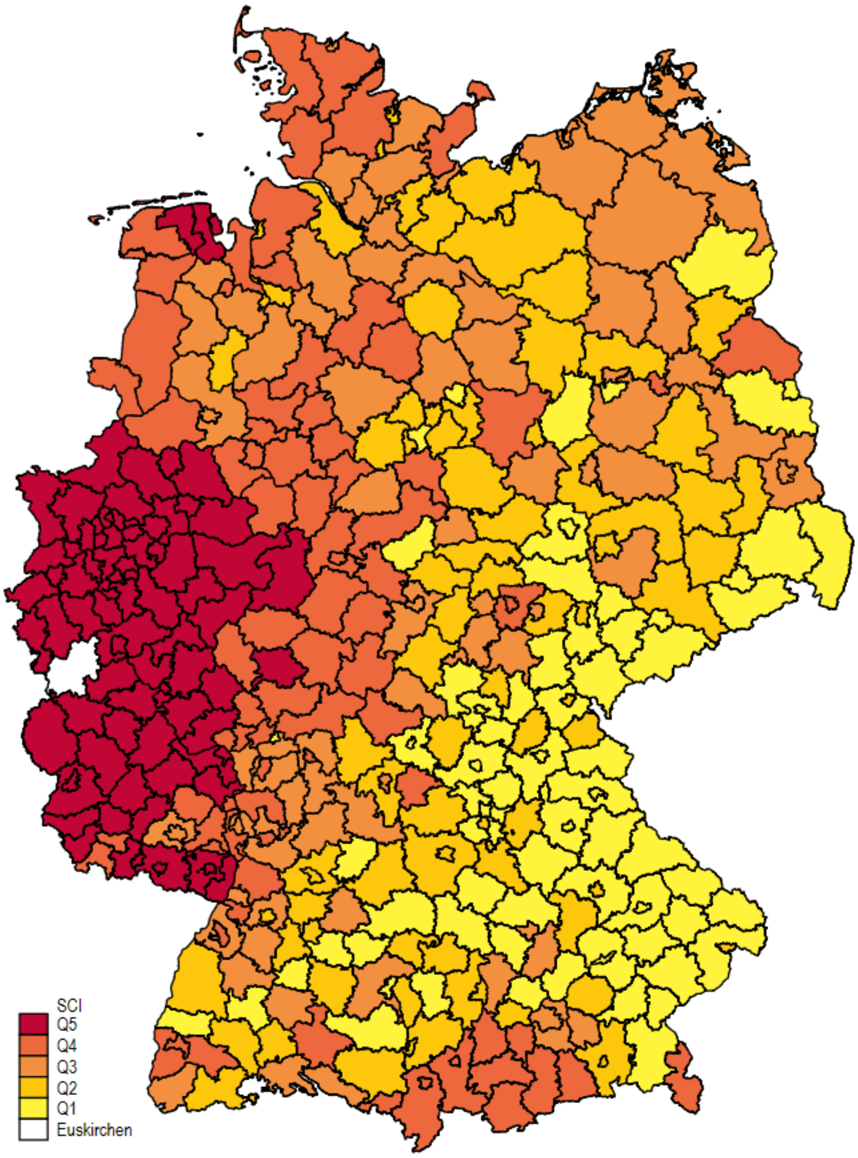


Figure 3.4: Social Connectedness Index Ahrweiler

This maps illustrates the Social Connecedness Index for the county Ahrweiler (white). Darker red shading indicates a higher connectedness.

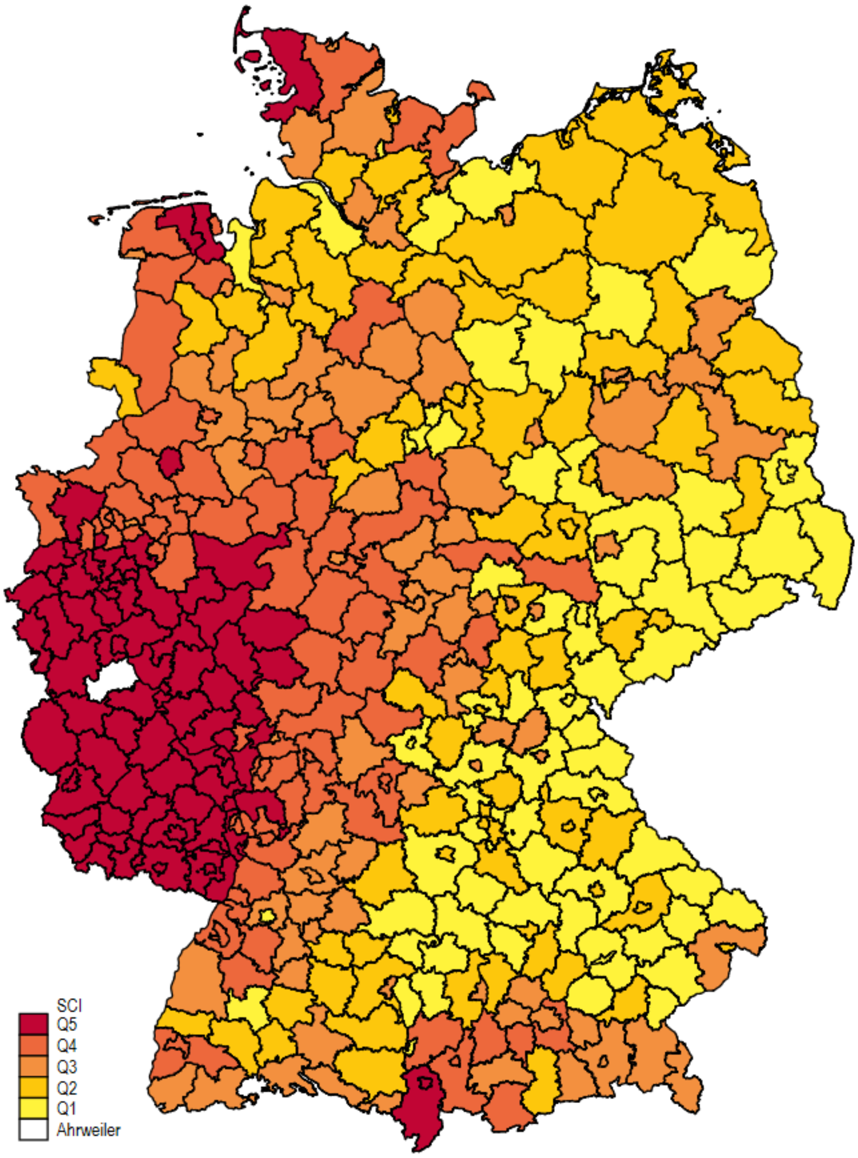


Figure 3.5: Sample Construction

This figure illustrates the intensity of the flood in 2021 on NUTS-3 level. The map shows three groups of counties based on their claim ratio for the universe of policies against elemental damages in the German insurance industry (GDV, 2022). 281 regions exhibit a claim ratio of 0%, colored in blue. The buffer category consist of 118 regions, displayed in white, that exhibit positive claim ratios below 12%. The two NUTS-3-regions with the highest claim ratios (23.93 (Landkreis Euskirchen) and 18.3 (Landkreis Ahrweiler) are colored in red.

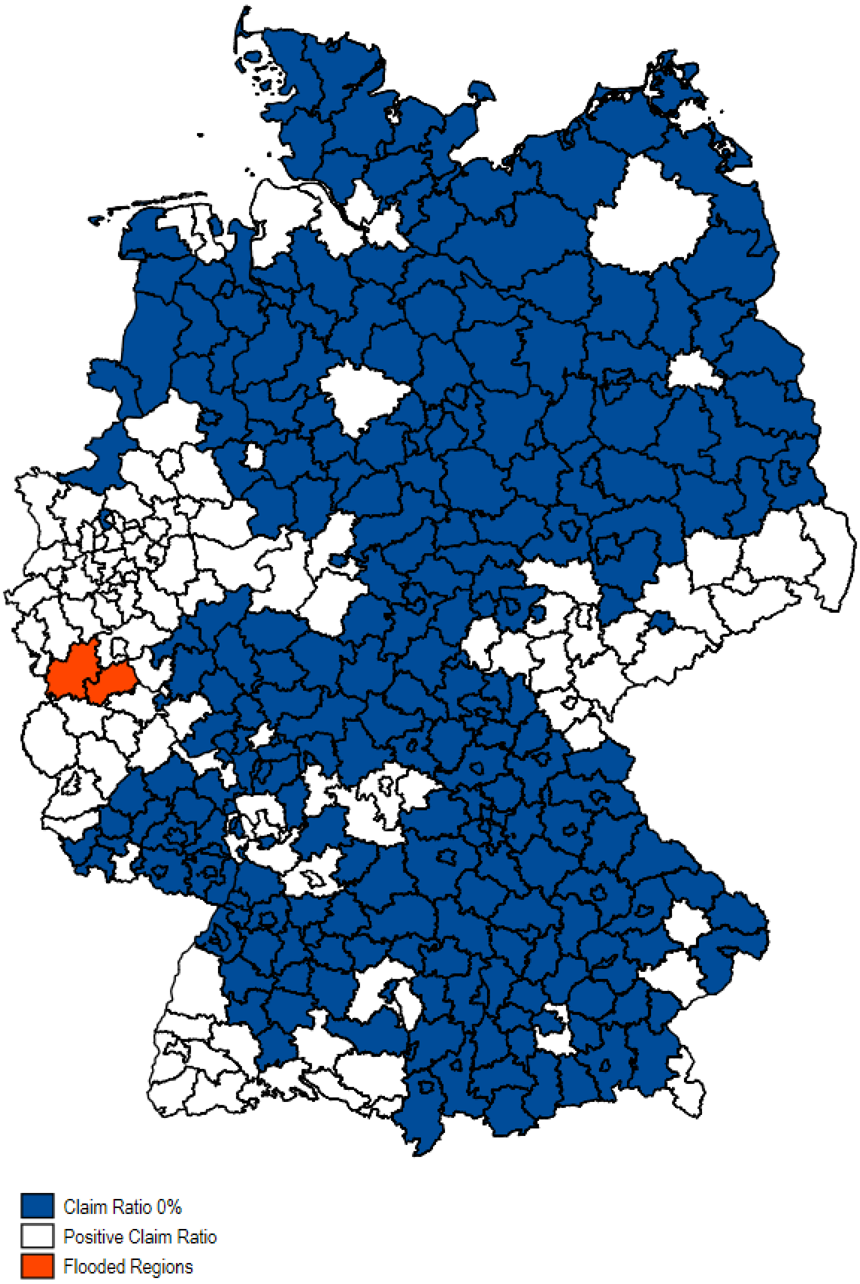


Figure 3.6: Flood-weighted Social Connectedness

This figure illustrates the flood weighted social connectedness of all regions that did not reported any insurance claims regarding the flood in July 2021 to the flooded regions (colored in red) by quintiles. Other regions that exhibit a positive claim ratio serve as a buffer category and are illustrated in white. Darker shading represents a higher connectedness. Data on social connectedness is taken from Bailey et al. (2018b) and has been introduced for Europe by Aref et al. (2020).

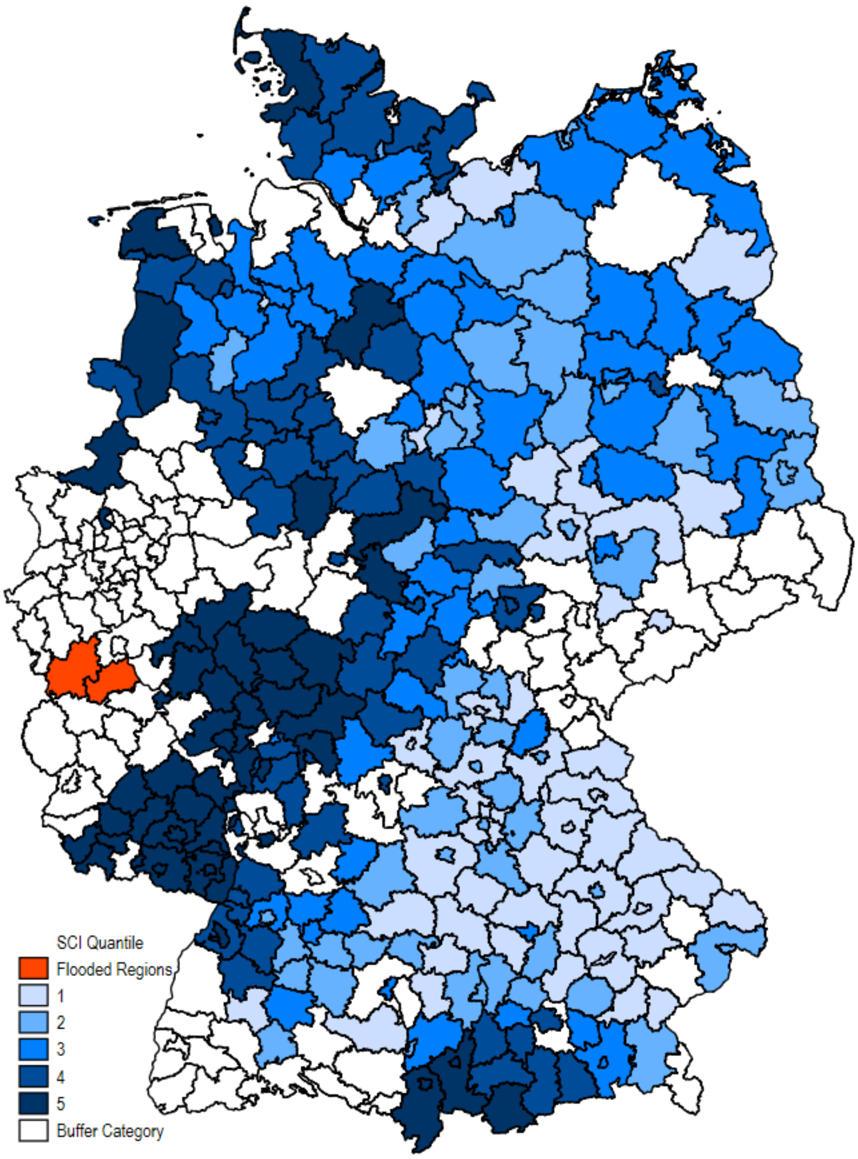


Figure 3.7: Share of Policies - Monthly Estimates

This figure illustrates the coefficient estimates of flood weighted social connectedness for every month before and after the flood in mid July 2021 (marked by the dashed red line). Dependent variable is the share of policies that include elemental damages. June 2021, one month before the flood is chosen as the baseline period. The orange bars represent 95% confidence intervals.

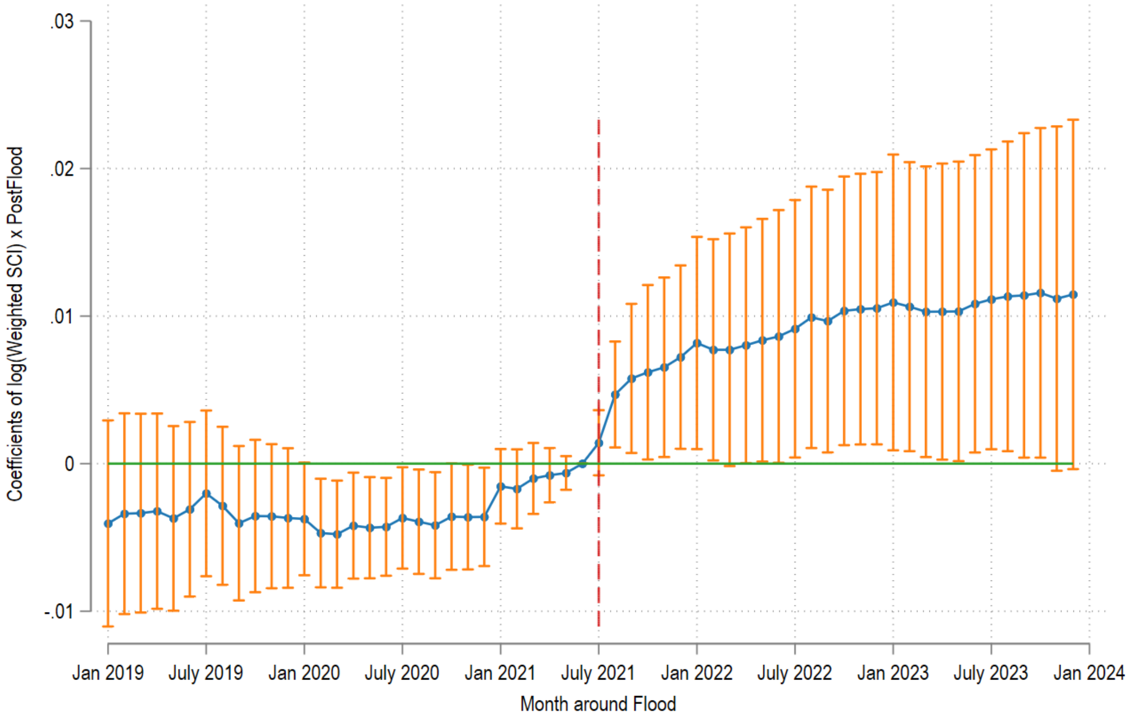
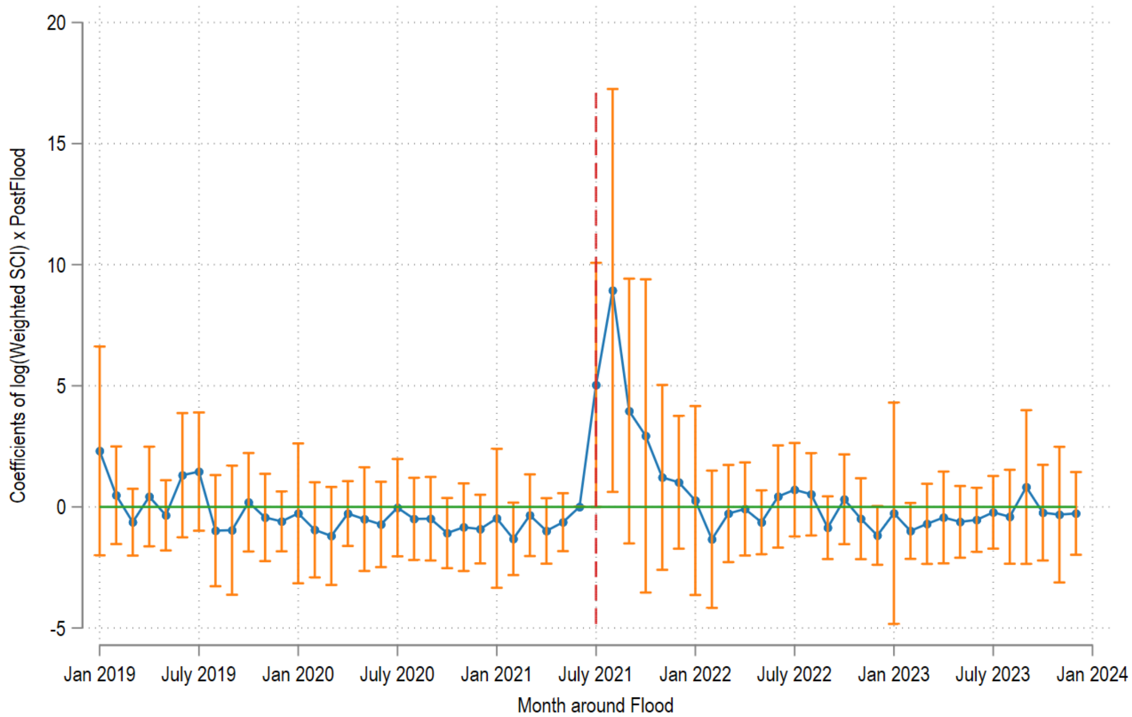


Figure 3.8: Timing of the Effect

This figure illustrates the coefficient estimates of flood weighted social connectedness for every month before and after the flood in mid July 2021 (marked by the dashed red line). Dependent variable is the change in the number of policies including elemental damages. June 2021, one month before the flood is chosen as the baseline period. The orange bars represent 95% confidence intervals.



Tables

Table 3.1: Descriptive Statistics

This table displays descriptive statistics for the main data set. All variables are at the NUTS-3 level. A detailed description of all variables can be found in Table A3.1.

	Mean	Median	SD	Min	Max	Obs
Main Variables						
Share of Policies Against Natural Hazard	0.53	0.53	0.15	0.12	0.94	16,860
Share of Premiums of Policies Against Natural Hazard	0.15	0.15	0.04	0.04	0.27	16,860
Share of Insurance Coverage of Policies Against Natural Hazard	0.56	0.56	0.15	0.12	0.96	16,860
log(Weighted SCI)	7.09	6.98	0.43	6.45	10.29	16,860
log(Weighted Distance)	4.38	4.46	0.42	2.76	5.04	16,860
Post Flood Dummy	0.50	0.50	0.50	0.00	1.00	16,860
Control Variables						
Fluvial Hazard	0.07	0.04	0.10	0.00	0.47	16,800
Population Share Exposed to Fluvial Flooding	0.04	0.02	0.07	0.00	0.50	16,800
Number of Natural Hazard Events (2002-2021)	10.32	10.00	5.88	1.00	29.00	16,800
log(Average Local Damage Cost (2002-2021))	8.62	8.62	0.41	7.34	10.35	16,800
Avg. Claim Ratio Heavy Rain (2002-2021)	0.04	0.04	0.02	0.01	0.11	16,800
log(Avg. Cost Heavy Rain (2002-2021))	8.55	8.54	0.30	7.59	9.36	16,800
Heavy Rain Events (2002-2021)	95.16	90.50	66.83	2.00	395.00	16,800
Projected Change in Heavy Precipitation Days	0.27	0.00	0.44	0.00	1.00	16,800
log(Length of Rivers)	5.46	5.78	1.12	2.78	7.67	16,860
Share of Area in Mountains	0.20	0.00	0.28	0.00	1.00	16,860
Climate Policy Attitudes						
Green Party Regional Votes (pre 2021)	0.14	0.13	0.06	0.01	0.33	16,860
Green Party State Votes (pre 2021)	0.13	0.11	0.08	0.02	0.41	16,860
Willingness to Pay Higher Cost for Fossil Fuels	0.37	0.37	0.04	0.29	0.53	16,860
Investment in Renewables Planned	0.33	0.33	0.04	0.18	0.45	16,860
Willingness to Protest against Local Wind Power Plants	0.42	0.42	0.05	0.28	0.57	16,860
Social Capital						
Voter Turnout (2017)	75.76	76.10	3.84	64.40	84.40	16,860
Covid Vaccinations per Capita June (2021)	0.49	0.45	0.25	0.11	2.07	16,860
SCI Dispersion	4.50	4.46	0.55	2.60	5.93	16,860
Educational Equality	-0.12	-0.09	0.10	-0.48	-0.00	16,800

Table 3.2: Baseline Result Shares

This table shows the results from estimating regression 3.3. The dependent variables is the share of policies including elemental damage. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3 level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Share of Policies Including Elemental Damages				
Spatial Level					
SCI	NUTS-3	NUTS-3	NUTS-3	NUTS-3	NUTS-3
Insurane Policies	NUTS-3	NUTS-3	Postcode	Postcode	Postcode
	(1)	(2)	(3)	(4)	(5)
log(Weighted SCI) × PostFlood	0.025***	0.012***	0.027***	0.016**	0.016**
	(0.004)	(0.005)	(0.004)	(0.007)	(0.007)
County FE	Yes	Yes	Yes	Yes	
Year-Month FE	Yes		Yes		
State × Year-Month FE		Yes		Yes	Yes
Postcode FE					Yes
R^2	0.985	0.991	0.627	0.631	0.953
Within-R^2	0.075	0.014	0.002	0.000	0.002
Observations	16,860	16,800	279,600	279,600	279,600

Table 3.3: Baseline Result Including Controls

This table shows the results from estimating regression 3.3 including a control for social connectedness derived from 10,000 placebo floods as well as a control term for distance to the flooded areas, weighted by flood intensity. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3 level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × PostFlood	0.025***	0.006***	0.027***	0.013**	0.003*	0.017**
	(0.008)	(0.002)	(0.008)	(0.007)	(0.002)	(0.007)
log(Random Weighted SCI) × PostFlood	0.000	-0.002	-0.004	0.002	-0.002	-0.002
	(0.009)	(0.002)	(0.009)	(0.012)	(0.003)	(0.013)
log(Weighted Distance) × PostFlood	0.001	0.003*	0.006	0.003	0.002	0.009
	(0.008)	(0.002)	(0.008)	(0.009)	(0.002)	(0.010)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.985	0.984	0.984	0.991	0.989	0.990
Within-R^2	0.075	0.041	0.063	0.014	0.009	0.018
Observations	16,860	16,860	16,860	16,800	16,800	16,800

Table 3.4: Baseline Result and Alternative Explanations

This table shows the results from estimating regression 3.3 including additional control terms for current risk of flood, past natural hazard events, climate projections into the future and geographic characteristics of regions. The dependent variable is the logarithmized share of policies including elemental damages. Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3 level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	Share of Policies Including Elemental Damages										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
log(Weighted SCI) × PostFlood	0.012***	0.012***	0.012***	0.012**	0.013***	0.012**	0.012***	0.012**	0.011**	0.013***	0.013***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Fluvial Hazard × PostFlood	0.000										0.007
	(0.017)										(0.023)
Population Share Exposed to Fluvial Flooding × PostFlood		-0.007									-0.015
		(0.021)									(0.026)
Number of Natural Hazard Events (2002-2021) × PostFlood			-0.000								-0.001*
			(0.000)								(0.000)
Avg. Local Damage Cost (2002-2021) × PostFlood				0.004							0.000
				(0.004)							(0.004)
Number of Heavy Rain Events (2002-2021) × PostFlood					-0.000						-0.000**
					(0.000)						(0.000)
Claim Ratio Heavy Rain Events (2002-2021) × PostFlood						0.006					0.023
						(0.085)					(0.106)
Avg. Damage Heavy Rain Events (2002-2021) × PostFlood							0.011*				0.012
							(0.006)				(0.008)
Projected Change in Heavy Precipitation Days × PostFlood								0.002			0.002
								(0.003)			(0.003)
log(Length of Rivers) × PostFlood									0.002		0.007**
									(0.001)		(0.003)
Share of Area in Mountains × PostFlood										-0.002	-0.004
										(0.005)	(0.006)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.991	0.991	0.991	0.991	0.991	0.991	0.991	0.991	0.991	0.991	0.991
Within-R^2	0.014	0.014	0.017	0.016	0.015	0.014	0.025	0.015	0.018	0.014	0.055
Observations	16,740	16,740	16,740	16,740	16,740	16,740	16,740	16,740	16,800	16,800	16,680

Table 3.5: Climate Policy Attitudes

This table shows the results from estimating regression 3.3. To explore the mechanism, several triple interaction terms are added. These include the share of votes for the green party in the last regional election prior to the flood, the share of votes for the green party in the last state election prior to the flood, the share of people that states willingness to pay higher cost for fossil fuels, the share of people that has planned an investment in renewable energy sources and the share of people that states willingness to protest against local wind power plants. The dependent variable is the share of policies including elemental damage. Data for climate policy attitudes taken from Levi et al. (2022). Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3 level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Share of Policies Including Elemental Damages				
	(1)	(2)	(3)	(4)	(5)
Log(Weighted SCI) × Postflood	0.037***	0.023***	0.065**	0.166***	-0.081**
	(0.012)	(0.006)	(0.027)	(0.056)	(0.039)
Green Party Regional Votes (pre 2021) × Postflood	0.851**				
	(0.400)				
Log(Weighted SCI) × Green Party Regional Votes (pre 2021) × Postflood	-0.131**				
	(0.056)				
Green Party State Votes (pre 2021) × Postflood		0.588			
		(0.385)			
Log(Weighted SCI) × Green Party State Votes (pre 2021) × Postflood		-0.092*			
		(0.054)			
Willingness to Pay Higher Cost for Fossil Fuels × Postflood			0.827*		
			(0.460)		
Log(Weighted SCI) × Willingness to Pay Higher Cost for Fossil Fuels × Postflood			-0.131**		
			(0.065)		
Investment in Renewables Planned × Postflood				3.180***	
				(1.108)	
Log(Weighted SCI) × Investment in Renewables Planned × Postflood				-0.457***	
				(0.161)	
Willingness to Protest against Local Wind Power Plants × Postflood					-1.515**
					(0.676)
Log(Weighted SCI) × Willingness to Protest against Local Wind Power Plants × Postflood					0.226**
					(0.098)
Income Controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.991	0.991	0.991	0.991	0.991
Observations	16,740	16,740	16,740	16,740	16,740

Table 3.6: Social Capital

This table shows the results from estimating regression 3.3. To explore the mechanism, several triple interaction terms are added. The dependent variable is the share of policies including elemental damage. Triple interactions include the voter turnout in the 2017 federal election, Covid vaccinations per capita in Junee 2021, SCI dispersion as computed by Equation 3.4 and educational equality. Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3 level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Share of Policies Including Elemental Damages			
	(1)	(2)	(3)	(4)
log(Weighted SCI) × Postflood	0.161**	0.020***	-0.058	0.022***
	(0.075)	(0.007)	(0.036)	(0.007)
Voter Turnout (2017) × Postflood	0.013*			
	(0.007)			
log(Weighted SCI) × Voter Turnout (2017) × Postflood	-0.002*			
	(0.001)			
Covid Vaccinations per Capita June (2021) × Postflood		0.089*		
		(0.048)		
log(Weighted SCI) × Covid Vaccinations per Capita June (2021) × Postflood		-0.013*		
		(0.007)		
SCI Dispersion × Postflood			-0.107*	
			(0.057)	
log(Weighted SCI) × SCI Dispersion × Postflood			0.016*	
			(0.008)	
Educational Equality × Postflood				-0.460*
				(0.241)
log(Weighted SCI) × Educational Equality × Postflood				0.067**
				(0.034)
Income Controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes
R^2	0.991	0.991	0.991	0.991
Observations	16,740	16,740	16,740	16,740

Concluding Remarks

This thesis examines key areas of financial intermediation, including cross-border investment flows, bank branch networks, and insurance markets. At the heart of all three chapters lies the transmission of information through social networks, measured by using Facebook's Social Connectedness Index (SCI). This data provides insights into the role of social ties across various geographic levels, enabling a nuanced understanding of how networks help to overcome information asymmetries, to reduce economic frictions, and to improve the efficiency of financial intermediation. Each chapter tackles issues that are likely to intensify in the future: Foreign direct investment (FDI) remains a critical source of external financing for emerging markets but is increasingly threatened by rising macroeconomic uncertainty, nationalism, and geopolitical tensions. The traditional banking model is under pressure as technological advancements accelerate, challenging established institutions. Lastly, as climate risks continue to escalate, the need for better protection against environmental hazards will become ever more pressing in the face of the ongoing climate crisis.

The first chapter investigates how cross-border social networks influence FDI. It shows that stronger social ties between countries significantly boost FDI. A 1% higher social connectedness is associated with an increase in FDI by 0.12%. The significant effect of social connectedness on FDI renders the effect of traditionally important determinants such as distance insignificant. Social ties are particularly important if legal, cultural, and institutional barriers exist. These connections help to bridge information frictions and to reduce the risks associated with investing in foreign markets. The analysis demonstrates that social networks not only encourage investment by reducing macroeconomic uncertainty but also mitigate challenges such as climate-related risks: In cases where both originating and destination countries are subject to higher physical risk from natural disasters, the importance of social connectedness for FDI is higher. Interestingly, this is also the case if countries transition towards a green economy. Investment decision-makers can leverage existing social networks to mitigate climate risks and macroeconomic uncertainties. Building and fostering new social ties can serve as an effective tool for policymakers to support economic integration through foreign direct investment.

The second chapter focuses on the impact of financial technology (fintech) on traditional banking services. It reveals that the rise of fintech lenders in the U.S. mortgage market is associated with a significant decline in local bank branches. The results suggest that on average 45% of the overall decline in the number of bank branches can be attributed to the increasing market share of fintech lenders. The relationship is particularly pronounced in metropolitan statistical areas and among smaller banks. Small banks rely on long-term relationships between borrower and lender, which makes the operation of branches more costly. Using the SCI, the study identifies how regions with strong social connectedness to Wayne County (Michigan), the headquarter location of the largest fintech lender, saw a larger fintech adoption, which is associated with a decrease in physical bank branches. This chapter highlights the disruptive effects of fintech on traditional banking structures and its broader implications for financial access and policy transmission. Higher market shares for fintech lenders are also linked to a decline in local deposits, indicating that branch closures following the rise of fintech have broader effects on the availability of banking services and bank balance sheets. Beyond impacting the bank lending channel of monetary policy, increased fintech lending also appears to have an effect via the deposit channel. Decreasing market power in retail deposits may also change the risk structure of banks balance sheets. Policymakers should constantly monitor the influence of new technologies in the financial sector with regard to changes in the risk structure of banks as well as the financial inclusion of households.

The third chapter examines the role of social networks in homeowners' decisions to purchase insurance for elemental damage, using the 2021 flood in Germany as a natural experiment. The analysis focuses on regions that have not been hit themselves. Using social connectedness to measure the exposure to the flooded regions, the results show positive and highly statistically significant peer effects on insurance uptake after the flood: Regions that are better connected to the flooded areas exhibit an increased share of policies that include elemental damages in their policies. Climate policy attitudes and social capital are exploited to identify heterogeneity in the responsiveness to peer effects. The results support the idea of information updating via social learning. People receive an information update about their own flood risk through first-hand experiences from directly affected peers. These updates are more pronounced in regions with lower bridging and higher bonding social capital. Limited bridging capital restricts the ability to receive and process information from outside the own community. Strong bonding capital reinforces a focus on the immediate environment and people similar to oneself. In regions, where warnings from authorities and politicians are less effective and climate risks are taken less seriously, social connections can play a critical role in enhancing disaster preparedness.

Taken together, this thesis highlights the critical role of social networks in financial intermediation across various actors. The first chapter demonstrates how social ties help to reduce information frictions in foreign direct investment. Chapter two explores how social networks influence product adoption, shedding light on the trend of bank branch closures amid the growing market share of fintech lenders. The last chapter uncovers peer effects in insurance markets. Using a natural disaster in distant but socially connected regions, it illustrates the impact of social networks on insurance uptake. These findings offer important insights for policymakers and financial decision-makers. Leveraging existing social connections can help to overcome information frictions. While other factors such as geographic distance, entrenched structures and risk preferences are impossible or at least very difficult to change, existing social networks can be easily identified. Carefully designed policies can exploit them to spur investment and boost economic integration. Evaluating social networks can be used to forecast the adoption of technologies and new financial products, the diffusion of financial risks and the propagation of economic shocks.

List of Appendix Figures

List of Appendix Tables

APPENDIX A1

Appendix for Chapter 1

A1.1 Implementation: Control Function Approach

The problems that occur when estimating a gravity equation with OLS, highlighted by Silva and Tenreyro (2006), also apply to standard IV strategies like a two stage least squares estimator. Next to that there is no efficient PPML estimator that allows to instrument for a variable and to control for a large set of fixed effects. To overcome these problems, we use a control function approach which is well explained by Wooldridge (2015) and Lin and Wooldridge (2019). This approach relies on two steps: In the first one we regress the SCI on the instrument, the gravity, economic and cultural controls and include the full set of fixed effects, using a simple OLS regression:

$$log(SCI_{i,j,t}) = \beta_1 \times IV_{i,j} + \beta_2 \times G_{i,j,t} + \beta_3 \times A_{i,j,t} + \delta_{i,t} + \delta_{j,t} + \eta_{i,j,t} \tag{A1.1}$$

Since this equation is estimated with OLS, the dependent variable SCI is included in a logarithmized term, IV represents the instrumental variable. All other notations are similar to the baseline model. To preserve consistency in the subsequent step, the error term is represented by $\eta_{i,j,t}$. Based on the results of Equation A1.1, the residuals are estimated and denoted by $Res_{i,j,t}$. These are subsequently included as another control variable in the PPML estimation, similar to the baseline model:

$$\begin{aligned} FDI_{i,j,t} = {} & \exp[\beta_1 \times log(SCI_{i,j}) + \beta_2 \times Res_{i,j,t} + \beta_3 \times G_{i,j,t} + \beta_4 \times A_{i,j,t} \\ & + \delta_{i,t} + \delta_{j,t}] \times \epsilon_{i,j,t} \end{aligned} \tag{A1.2}$$

This allows to decompose the effect of SCI into the part which can be explained by the instrument and a part that cannot. The coefficient of the SCI represents the elasticity of FDI using the respective instrument while the coefficient of the residual displays the remaining effect of the SCI in explaining FDI. The closer the latter coefficient is to zero, the smaller is the over- or underestimation of the role of social relationships as a determinant of FDI.

A1.2 Appendix Figures

Figure A1.1: Austria and the Former Countries (Partially) Included in its Empire

This figure highlights the countries that were part of the Austro-Hungarian Empire. Some countries were fully (or mostly) part of the Empire (Czechoslovakia, Hungary, Slovakia, Slovenia, Croatia, Bosnia and Herzegovina). In other countries smaller regions were formerly part of the Empire (Italy, Poland, Romania). Serbia was never part of the Empire, but was fully occupied in World War I.

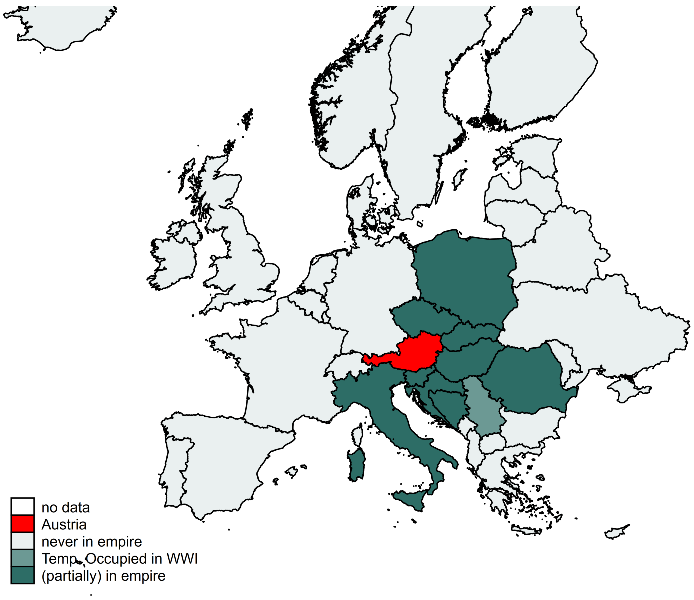
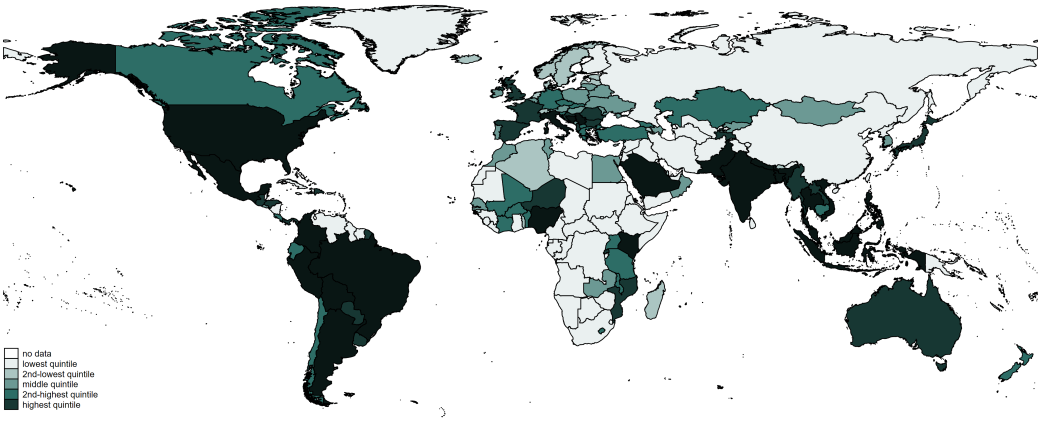


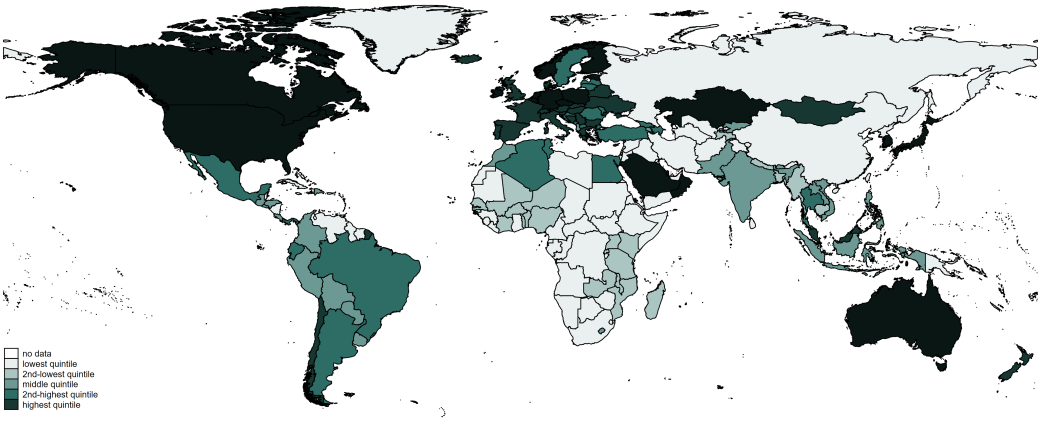
Figure A1.2: Physical Risk and Transition Risk Across the Globe

This figure shows that physical and transition risk is distributed differently across the world using two measures from our interactions in Table 1.5 and 1.6. Panel A plots the average number of floods in each country per year in quintiles. Panel B plots the average CO2 emissions per capita in quintiles. In both cases values represent averages over the years 2010-2019.

Panel A: Average Number of Floods per Year



Panel B: Average CO2 Emissions per Capita per Year



A1.3 Appendix Tables

Table A1.1: Variable Definitions and Data Sources

Variable Name	Description
Main Variables	
Inward FDI	Inward foreign direct investment flow in million USD from origin country i to destination country j in year t in nominal terms; source: UNCTAD WIR 2020
Social Connectedness Index	Relative probability of a user in country i being friends with a user in country j on Facebook taking values between 1 and 1 billion, winsorized at the 99$^{\text{th}}$ percentile; source: Humanitarian Data Exchange
Gravity Control Variables	
Distance	Population weighted distance measured in km between country i and country j; source: CEPII
Common Border	Binary variable equal to one if the origin country i and destination country j share the same border; source: CEPII
Common Official Language	Binary variable equal to one if the origin country i and destination country j share the same official language; source: CEPII
Common Colonizer	Binary variably equal to one if the origin country i and destination country j had a common colonizer after 1945; source: CEPII
Colonial Relationship after 1945	Binary variable equal to one if the origin country i and destination country j had a colonial relationship after 1945; source: CEPII
Economic Control Variables	
Avg. Trade Volume	Total average trade volume per country pair in million USD, i.e., country i's imports from country j added up with country j's imports from country i measured as cost, freight, insurance (CIF); averaged over 3 years prior to FDI in t; source: IMF

Table A1.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Avg. GDP Difference	Absolute value of the difference between the average GDP in current (2020) USD of origin country i and destination country j; averaged over three years prior to the FDI in t; source: The World Bank
Avg. GDP Growth Difference	Absolute value of the difference between the average GDP growth rate in percentage points of origin country i and destination country j; averaged over three years prior to the FDI in t; source: The World Bank
Regional Trade Agreement	Binary variable equal to one if origin country i and destination country j have a regional trade agreement in place in year t. Values are three years lagged, as all other economic covariates; source: CEPII
Common Currency	Binary variable equal to one if origin country i and destination country j have the same currency; source: DataHub
Cultural Control Variables	
Shared Religion Index	Index reaching from 0 to 1 indicating to what extend origin country i and destination country j share the same religion; data for 2010; source: PEW Research Center
Political Distance	Proximity of the political systems (democracy vs. autocracy) in origin country i and destination country j in year t; constructed as the squared distance between the originating and destination countries'; indicators originally vary from 0 to 9 and are divided by 9 to vary between 0 and 1; source: Varieties of Democracy Institute
Common Legal Origin	Binary variable equal to one if origin country i and destination country j have the same legal origin; source: CEPII
Common Ethnological Language	Binary variable equal to one if origin country i and destination country j share a common language spoken by at least 9% of the population; source: CEPII
Instrumental Variable	

Table A1.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Genetic Distance	Measure building on the time elapsed since the last common ancestors of people living in origin country i and destination country j; source Spolaore and Wacziarg (2018)
Unilateral Variables	
Worldwide Governance Indicators	
Dest./Orig. Regulatory Quality	Measure of the perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source: Worldwide Governance Indicators (WGI) project
Dest./Orig. Government Effectiveness	Measure of the perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source: Worldwide Governance Indicators (WGI) project
Cost of Doing Business	
Dest./Orig. Days to Start a Business	Unilateral measure of the days that are required to start a business in the destination country of the investment, winsorized at the 99^{th} percentile; source: CEPII
Dest./Orig. Procedures to Start a Business	Unilateral measure of the procedures that are required to start a business in the destination country of the investment, winsorized at the 99^{th} percentile; source: CEPII
Dest./Orig. Costs to Start a Business	Unilateral measure of the costs as percent of the gross national income (GNI) that occur when starting a business in the destination country of the investment, winsorized at the 99^{th} percentile; source: CEPII

Table A1.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Dest./Orig. Financial Market Development	Unilateral measure for the development of financial markets in the destination country of the investment, taking values between 0 (low) and 1 (high); source: IMF Financial Development Index
Climate Risk and Natural Disasters	
Dest./Orig. Global Climate Risk Index	Measure of the extent to what countries have been affected by the impacts of weather-related loss events (storms, floods, heat waves etc.).Higher values correspond to lower risk; source:https://www.germanwatch.org/en/cri
Dest./Orig. Number of Floods	number of floods per country and year. Data is collected from numerous sources such as UN agencies, non-governmental organizations, reinsurance companies, research institutes, and press agencies by the Centre for Research on the Epidemiology of Disasters (CRED); source: International Disaster Database (EM-DAT)
Dummy for Flood in Dest./Orig. Country	Dummy equal to one if there is at least one flood in the respective country and year. The dummy is generated using data of the number of floods in origin and destination countries; source: International Disaster Database (EM-DAT)
Dest./Orig. Percentage of Population Exposed to Wildfires	Share of population that is exposed to wildfires; source: Green Growth Indicators
Dest./Orig. Climate Change Physical Risk Sautner et al. (2023)	Physical Risk arising from climate change as measured by financial analysts attention paid to firms' climate change exposure in earnings conference calls. Data is aggregated on country-year level; source: Sautner et al. (2023)
Transition Risk	
Dest./Orig. CO2-Emissions	CO2-emissions measured as tons per capita from the Climate Watch Historical GHG Emissions (1990-2020) by the World Resources Institute; source:Worldbank
Dest./Orig. Renewable Energy Supply	Renewable energy supply as percentage of the total energy supply; source: Green Growth Indicators

Table A1.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Dest./Orig. Development of Environmental-related Technologies	Development of environmental-related technologies as percentage of all technologies; source: Green Growth Indicators
Dest./Orig. Changes in Emission Related Policies	Changes in emission related policies generated as the year-on-year changes in the count of emission related policies in place on country-year level; source: Gu and Hale (2023)
Dest./Orig. Energy Related Tax Revenue	Energy related tax revenue as percentage of total environmental tax revenue; source: Green Growth Indicators
Dest./Orig. Climate Change Regulatory Risk Sautner et al. (2023)	Physical Risk arising from regulatory changes with respect to climate change as measured by financial analysts attention paid to firms' climate change exposure in earnings conference calls. Data is aggregated on country-year level; source: Sautner et al. (2023)
Uncertainty Variables	
Dest./Orig. World Uncertainty Index	Index measuring uncertainty in the destination country of the investment computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is then rescaled by multiplying by 1,000,000. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index
Global World Uncertainty Index	Index measuring global uncertainty computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is then rescaled by multiplying by 1,000,000. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index
Global GDP-Weighted World Uncertainty Index	GDP-Weighted index measuring global uncertainty computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is rescaled to take values between 0 and 1. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index

Table A1.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Geopolitical Risk Index	Index that measure adverse geopolitical events based on a tally of newspaper articles covering geopolitical tensions, developed by Caldara and Iacoviello (2022). Higher values correspond to higher uncertainty; source: GPR Index
Macro Uncertainty Index Jurado et al. (2015)	Index that measures uncertainty for the broader macro economy. It quantifies the magnitude of unpredictability of the future by measuring a common component in the time-varying volatilities of 1-year-step-ahead forecast errors across a large number of macroeconomic variables. These cover a large number of macroeconomic time series including real economic variables as well as financial variables. The index and method is developed by Jurado et al. (2015). Higher values correspond to higher uncertainty; source: Macro Uncertainty Index
Real Uncertainty Index Ludvigson et al. (2021)	Index that measures uncertainty in real activity variables. It is introduced by Ludvigson et al. (2021) and constructed following the method developed by Jurado et al. (2015). It quantifies the magnitude of unpredictability of the future by measuring a common component in the time-varying volatilities of 1-year-step-ahead forecast errors across a large number of real activity variables. Higher values correspond to higher uncertainty; source: Real Uncertainty Index

Table A1.2: Correlation Matrix

This table displays the correlation between all bilateral control variables, instrumental variables and the log of SCI as well as the log of distance.

	log(SCI)	log(Distance)
log(SCI)	1.00	-0.65
log(Distance)	-0.65	1.00
log(Trade Volume)	0.19	-0.20
log(GDP Difference)	-0.16	0.28
log(GDP Growth Difference)	-0.12	0.14
Regional Trade Agreement	0.45	-0.60
Common Currency	0.29	-0.34
Shared Religion Index	0.36	-0.20
Political Distance	-0.19	0.18
Common Legal Origin	0.21	-0.12
Common Ethno. Language	0.32	-0.04
Common Border	0.23	-0.37
Common Official Language	0.36	-0.07
Common Colonizer	0.22	-0.13
Colonial Relationship after 1945	0.10	0.03
Genetic Distance	-0.46	0.60

Table A1.3: Social Connectedness and FDI: OLS Regressions with log(1+FDI)

This table presents the results for estimating Equation 1.3 using OLS but with log(1+FDI) as the dependent variable. In Column (3) genetic distance from Spolaore and Wacziarg (2018) is used as an instrument for the SCI. For this column, the coefficient of the first stage as well as the p-values and F-values for are included. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable		log(1+FDI)	
Instrumental Variable			Genetic Distance
	(1)	(2)	(3)
log(SCI)	0.476***	0.106***	0.423***
	(0.013)	(0.020)	(0.108)
log(Distance)		-0.473***	-0.245***
		(0.040)	(0.079)
log(Trade Volume)		0.122***	0.054**
		(0.014)	(0.027)
log(GDP Difference)		-0.148***	-0.145***
		(0.015)	(0.016)
log(GDP Growth Difference)		0.016	0.017
		(0.012)	(0.013)
Regional Trade Agreement		0.020	-0.047
		(0.053)	(0.047)
Common Currency		-0.187**	-0.154**
		(0.078)	(0.068)
Shared Religion Index		0.539***	0.181
		(0.091)	(0.132)
Political Distance		-0.094	-0.062
		(0.093)	(0.075)
Common Legal Origin		0.099***	0.056
		(0.038)	(0.034)
Common Ethno. Language		0.198*	0.108
		(0.105)	(0.081)
Common Border		0.041	-0.003
		(0.095)	(0.080)
Common Official Language		0.146	-0.040
		(0.111)	(0.082)
Common Colonizer		0.083	-0.080
		(0.088)	(0.088)
Colonial Relationship after 1945		0.500***	0.216
		(0.147)	(0.134)
Origin-Country FE	Yes	Yes	Yes
Destination-Country FE	Yes	Yes	Yes
R^2	0.429	0.452	0.121
Instrument (1st stage)			-17.371
p-value (1st stage)			0.000
F-value (1st stage)			200.578
Observations	45,313	45,249	41,384

Table A1.4: Estimates of Individual Control Variables: PPML

This table shows the results of estimating the effect of the SCI and all gravity, economic and cultural variables that are included as controls in Equation 1.2 on the dependent variable Inward FDI from country i to country j separately. Thereby the same PPML estimator as in Equation 1.2 is used. Column (1) represents a specification with no fixed effects and the log of social connectedness as the only independent variable. Column (2) represents a specification with fixed effects only, similar to Column (1) of Table 1.2. Columns (3)-(17) include all other controls one-by-one in decreasing order of explanatory power measured by the within-R^2. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI																
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
log(Trade Volume)			0.548***														
log(SCI)	0.274***			0.513***													
log(Distance)					-0.584***												
Regional Trade Agreement						0.880***											
Shared Religion Index							2.878***										
Common Border								0.855***									
Political Distance									-1.857***								
Common Official Language										0.664***							
Common Ethno. Language											0.710***						
Common Colonizer												2.247***					
Common Legal Origin													0.438***				
Common Currency														0.763***			
log(GDP Growth Difference)															-0.148***		
log(GDP Difference)																-0.107***	
Colonial Relationship after 1945																	0.005
Origin-Country × Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.037	0.727	0.769	0.762	0.758	0.744	0.740	0.737	0.736	0.736	0.736	0.735	0.734	0.733	0.730	0.729	0.727
Observations	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216	45,216

Table A1.5: Estimates of Individual Control Variables: OLS

This table shows the results of estimating the effect of the SCI and all gravity, economic and cultural variables that are included as controls in Equation 1.2 on the dependent variable log(Inward FDI) from country i to country j separately using OLS. Column (1) represents a specification with no fixed effects and the log of social connectedness as the only independent variable. Column (2) represents a specification with fixed effects only, similar to Column (2) of Table 1.2. Columns (3)-(17) include all other controls one-by-one in decreasing order of explanatory power measured by the R^2. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI																
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
log(SCI)	0.595***		1.098***														
log(Distance)				-1.980***													
log(Trade Volume)					0.988***												
Regional Trade Agreement						2.414***											
Common Official Language							2.955***										
Common Ethno. Language								2.852***									
Common Border									3.159***								
Shared Religion Index										3.264***							
Political Distance											-2.804***						
log(GDP Difference)												-0.401***					
Common Colonizer													2.404***				
Common Currency														2.091***			
Common Legal Origin															0.914***		
log(GDP Growth Difference)																-0.331***	
Colonial Relationship after 1945																	1.680***
Origin-Country × Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.038	0.534	0.614	0.613	0.608	0.566	0.559	0.559	0.551	0.550	0.545	0.541	0.541	0.541	0.540	0.538	0.536
Observations	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332	31,332

Table A1.6: First Stage Regressions for IV Regressions: Genetic Distance

This table presents the results for Equation A1.1, the first stage OLS regressions for Column (5) of Table 1.2. It shows the result from regressing the logarithmized SCI on the full set of controls and the respective instrument which is the genetic distance. Descriptive statistics for all variables are presented in Table 1.1, detailed descriptions are presented in Table A1.1. The regression includes a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(SCI)
	(1)
Genetic Distance	-17.371***
	(0.000)
log(Distance)	-0.572***
	(0.000)
log(Trade Volume)	0.225***
	(0.000)
log(GDP Difference)	-0.003
	(0.722)
log(GDP Growth Difference)	0.007
	(0.261)
Regional Trade Agreement	0.190***
	(0.000)
Common Currency	-0.100*
	(0.063)
Shared Religion Index	1.039***
	(0.000)
Political Distance	0.012
	(0.847)
Common Legal Origin	0.190***
	(0.000)
Common Ethno. Language	0.416***
	(0.000)
Gravity Controls	Yes
Origin-Country $\times$ Year FE	Yes
Destination-Country $\times$ Year FE	Yes
Adjusted R^2	0.796
Observations	41,384

Table A1.7: Social Connectedness and FDI: Baseline Results with Alternative FDI Data

This table shows robustness checks for Table 1.2, estimating Equation 1.2 for PPML and Equation 1.3 for OLS, using alternative FDI data. The dependent variable is FDI for PPML and log(FDI) for OLS, representing Inward FDI from country i to country j in year t. Data is taken from Damgaard et al. (2024). They use the IMF's Coordinated Direct Investment Survey to construct a global network of total FDI leveraging mirror reports from counterpart economies. The main explanatory variable is the log(SCI). The model controls for all gravity, economic and cultural covariates. Columns (1) and (2) include fixed effects only, Columns (3) and (4) include the full set of controls. In Columns (5)-(6) genetic distance from Spolaore and Wacziarg (2018) is used as an instrument for the SCI. IV specifications in PPML are estimated using a control function approach as presented in Section A1.1. First stage coefficients, p-value and F-value for the instruments are displayed at the bottom. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI	log(FDI)	FDI	log(FDI)	FDI	log(FDI)
Model	PPML	OLS	PPML	OLS	PPML	OLS
Instrumental Variable					Genetic Distance	
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.571***	1.050***	0.200***	0.385***	0.432***	0.797***
	(0.046)	(0.019)	(0.059)	(0.031)	(0.158)	(0.106)
Residuals					-0.249	
					(0.157)	
log(Distance)			0.048	-0.771***	0.267	-0.435***
			(0.085)	(0.059)	(0.182)	(0.096)
log(Trade Volume)			0.476***	0.271***	0.436***	0.182***
			(0.057)	(0.026)	(0.061)	(0.027)
log(GDP Difference)			-0.168***	-0.233***	-0.164***	-0.217***
			(0.034)	(0.021)	(0.036)	(0.017)
log(GDP Growth Difference)			-0.074***	0.010	-0.071***	0.004
			(0.018)	(0.016)	(0.018)	(0.015)
Regional Trade Agreement			0.060	0.210***	0.057	0.136***
			(0.107)	(0.074)	(0.107)	(0.051)
Common Currency			-0.292*	-0.210*	-0.258	-0.184**
			(0.177)	(0.117)	(0.175)	(0.080)
Shared Religion Index			1.614***	1.200***	1.286***	0.838***
			(0.317)	(0.148)	(0.378)	(0.138)
Political Distance			-0.074	0.115	0.070	0.178*
			(0.259)	(0.139)	(0.259)	(0.100)
Common Legal Origin			0.063	0.067	-0.003	-0.011
			(0.081)	(0.060)	(0.093)	(0.038)
Common Ethno. Language			0.073	0.280**	0.030	0.200**
			(0.176)	(0.138)	(0.176)	(0.101)
Common Border			-0.196	-0.180	-0.184	-0.317***
			(0.141)	(0.122)	(0.140)	(0.070)
Common Official Language			-0.048	0.232	-0.157	0.057
			(0.202)	(0.143)	(0.213)	(0.102)
Common Colonizer			0.207	0.259**	0.065	0.119
			(0.270)	(0.126)	(0.295)	(0.118)
Colonial Relationship after 1945			0.099	0.796***	-0.160	0.407***
			(0.150)	(0.177)	(0.208)	(0.117)
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.911		0.925		0.924	
R^2		0.678		0.709		0.338
Instrument (1st stage)					-28.048	-28.048
p-value (1st stage)					0.000	0.000
F-value (1st stage)					1030.1	853.8
Observations	82,946	41,497	82,781	41,470	75,504	38,184

Table A1.8: Social Connectedness and FDI: Real, Phantom and Ultimate Owner FDI

This table shows robustness checks for Columns (3) and (5) of Table 1.2 using alternative data for FDI. The dependent variable is Inward FDI from country i to country j in year t. Data is taken from Damgaard et al. (2024). They use the IMF's Coordinated Direct Investment Survey, construct a global network of total FDI leveraging mirror reports from counterpart economies. This data is used in Column (1) and (2). Damgaard et al. (2024) identifies phantom FDI that enters Special Purpose Entities (SPE). This data is used for Columns (3) and (4) and provides a placebo test. Finally, they match FDI flows to the Ultimate Investor Economies (UIE). This data is available from 2013 onwards and is used in Columns (5) and (6). The main explanatory variable is the log(SCI). The model controls for all gravity, economic and cultural covariates. Columns (2), (4) and (6) use genetic distance from Spolaore and Wacziarg (2018) is used as an instrument for the SCI. IV specifications in PPML are estimated using a control function approach as presented in Section A1.1. First stage coefficients, p-value and F-value for the instruments are displayed at the bottom. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI excluding SPEs		"Phantom FDI" (in SPEs)		"Real FDI" by UIE	
Instrumental Variable		Genetic Distance		Genetic Distance		Genetic Distance
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.263***	0.296**	-0.048	0.260	0.247***	0.369***
	(0.045)	(0.131)	(0.118)	(0.260)	(0.049)	(0.129)
Residuals		-0.033		-0.337		-0.127
		(0.128)		(0.281)		(0.121)
log(Distance)	0.012	0.044	-0.216**	0.105	-0.020	0.098
	(0.073)	(0.154)	(0.104)	(0.279)	(0.082)	(0.155)
log(Trade Volume)	0.529***	0.513***	0.324***	0.271***	0.524***	0.499***
	(0.046)	(0.048)	(0.071)	(0.087)	(0.051)	(0.053)
log(GDP Difference)	-0.154***	-0.152***	-0.155***	-0.138**	-0.164***	-0.159***
	(0.024)	(0.025)	(0.057)	(0.061)	(0.027)	(0.027)
log(GDP Growth Difference)	-0.045***	-0.045***	-0.075***	-0.072***	-0.037*	-0.036*
	(0.013)	(0.013)	(0.028)	(0.028)	(0.021)	(0.021)
Regional Trade Agreement	0.003	0.010	0.291**	0.269**	-0.026	-0.021
	(0.081)	(0.081)	(0.141)	(0.136)	(0.076)	(0.077)
Common Currency	-0.114	-0.103	-0.425**	-0.383*	-0.260**	-0.241*
	(0.123)	(0.123)	(0.204)	(0.205)	(0.125)	(0.124)
Shared Religion Index	1.532***	1.501***	2.597***	2.219***	1.398***	1.239***
	(0.271)	(0.332)	(0.564)	(0.584)	(0.253)	(0.304)
Political Distance	-0.305	-0.285	0.270	0.492	-0.286	-0.239
	(0.228)	(0.223)	(0.332)	(0.379)	(0.224)	(0.223)
Common Legal Origin	0.207***	0.204***	-0.347***	-0.415***	0.197***	0.166**
	(0.058)	(0.068)	(0.102)	(0.116)	(0.067)	(0.075)
Common Ethno. Language	-0.162	-0.162	0.421	0.347	-0.310**	-0.329***
	(0.125)	(0.127)	(0.264)	(0.277)	(0.127)	(0.127)
Common Border	-0.294***	-0.275***	-0.053	-0.060	-0.305***	-0.289**
	(0.101)	(0.101)	(0.254)	(0.256)	(0.112)	(0.114)
Common Official Language	0.304**	0.274**	-0.234	-0.325	0.602***	0.533***
	(0.121)	(0.138)	(0.268)	(0.254)	(0.132)	(0.154)
Common Colonizer	-0.057	-0.059	0.269	0.078	-0.172	-0.221
	(0.275)	(0.296)	(0.336)	(0.370)	(0.286)	(0.307)
Colonial Relationship after 1945	-0.041	-0.078	0.251	-0.148	-0.123	-0.265
	(0.147)	(0.197)	(0.319)	(0.435)	(0.168)	(0.208)
Origin-Country $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.935	0.934	0.952	0.952	0.935	0.934
R^2						
Instrument (1st stage)		-28.048		-28.048		-28.048
p-value (1st stage)		0.000		0.000		0.000
F-value (1st stage)		1030.1		1030.1		1030.1
Observations	82,781	75,504	32,963	29,590	53,052	48,123

Table A1.9: Social Connectedness and FDI: Cross-Section 2020

This table shows the baseline results from estimating Equation 1.2 for PPML and Equation 1.3 for OLS. The dependent variable is FDI for PPML and log(FDI) for OLS, representing Inward FDI from country i to country j in 2020. The main explanatory variable is the log(SCI). The model controls for gravity covariates (log of distance, a common border dummy, a common official language dummy, a dummy for a common colonizer post 1945, and a dummy for a colonial relationship post 1945), economic covariates (log of the total trade volume, log of the GDP difference, log of the GDP growth rate difference, a dummy for a regional trade agreement in place and a dummy indicating a shared currency) and cultural covariates (a shared religion index, a measure for political distance, a dummy for a common legal origin and a dummy indicating whether both countries share a common ethnological language). Columns (1) and (2) include fixed effects only, Columns (3) and (4) include the full set of controls. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI	log(FDI)	FDI	log(FDI)
Model	PPML	OLS	PPML	OLS
	(1)	(2)	(3)	(4)
log(SCI)	0.375***	1.277***	0.227*	0.321***
	(0.081)	(0.059)	(0.121)	(0.105)
log(Distance)			0.630***	-0.927***
			(0.210)	(0.214)
log(Trade Volume)			0.639***	0.489***
			(0.135)	(0.104)
log(GDP Difference)			-0.045	-0.203**
			(0.072)	(0.079)
log(GDP Growth Difference)			-0.047	-0.196**
			(0.106)	(0.093)
Regional Trade Agreement			0.622**	0.112
			(0.301)	(0.276)
Common Currency			-0.482	-0.443
			(0.380)	(0.434)
Shared Religion Index			0.470	1.509***
			(0.577)	(0.447)
Political Distance			-1.922***	-0.878
			(0.434)	(0.568)
Common Legal Origin			-0.114	-0.570***
			(0.198)	(0.201)
Common Ethno. Language			-0.137	0.638
			(0.372)	(0.491)
Common Border			-0.005	-0.317
			(0.328)	(0.362)
Common Official Language			-0.150	0.761
			(0.443)	(0.535)
Common Colonizer			-0.518	0.100
			(0.511)	(0.612)
Colonial Relationship after 1945			-0.235	2.435***
			(0.523)	(0.771)
Origin-Country × Year FE	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.735		0.759	
R^2		0.607		0.632
Observations	3,792	2,445	3,786	2,443

Table A1.10: Social Connectedness and FDI: Outflows

This table shows the baseline results from estimating Equation 1.2 for PPML and Equation 1.3 for OLS. The dependent variable is FDI for PPML and log(FDI) for OLS, representing outward FDI from country i to country j in year t. FDI data is compiled using outflows instead of inflows as in the rest of the paper. The main explanatory variable is the log(SCI). The model controls for gravity covariates (log of distance, a common border dummy, a common official language dummy, a dummy for a common colonizer post 1945, and a dummy for a colonial relationship post 1945), economic covariates (log of the total trade volume, log of the GDP difference, log of the GDP growth rate difference, a dummy for a regional trade agreement in place and a dummy indicating a shared currency) and cultural covariates (a shared religion index, a measure for political distance, a dummy for a common legal origin and a dummy indicating whether both countries share a common ethnological language). Columns (1) and (2) include fixed effects only, Columns (3) and (4) include the full set of controls. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI	log(FDI)	FDI	log(FDI)
Model	PPML	OLS	PPML	OLS
	(1)	(2)	(3)	(4)
log(SCI)	0.505***	1.118***	0.150**	0.176***
	(0.042)	(0.045)	(0.063)	(0.064)
log(Distance)			0.335***	-0.798***
			(0.102)	(0.128)
log(Trade Volume)			0.528***	0.684***
			(0.067)	(0.057)
log(GDP Difference)			-0.135***	-0.194***
			(0.038)	(0.041)
log(GDP Growth Difference)			-0.030	0.025
			(0.024)	(0.035)
Regional Trade Agreement			0.280**	-0.159
			(0.123)	(0.168)
Common Currency			-0.149	0.196
			(0.193)	(0.242)
Shared Religion Index			0.935***	0.147
			(0.315)	(0.269)
Political Distance			-0.572**	-1.039***
			(0.286)	(0.274)
Common Legal Origin			0.126	0.126
			(0.083)	(0.121)
Common Ethno. Language			0.064	0.729***
			(0.163)	(0.268)
Common Border			0.085	0.157
			(0.166)	(0.232)
Common Official Language			-0.016	-0.158
			(0.190)	(0.298)
Common Colonizer			0.179	0.802***
			(0.378)	(0.298)
Colonial Relationship after 1945			0.301	1.392***
			(0.259)	(0.477)
Origin-Country × Year FE	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.758		0.774	
R^2		0.628		0.653
Observations	32,611	21,753	32,575	21,736

Table A1.11: Social Connectedness and FDI: Previous FDI Flows

This table shows the baseline results from estimating Equation 1.2 for PPML and Equation 1.3 for OLS including previous FDI flows as control variables. Columns (1) and (2) use PPML with FDI as dependent variable, Columns (3) and (4) use OLS with log(FDI) as dependent variable and Columns (5) and (6) use OLS with log(FDI+1) as dependent variable. Descriptive statistics for all variables are presented in Table 1.1. Next to the gravity, economic and cultural controls, FDI flows from the previous year or the previous two years are included. Regressions include a full set of origin country and destination country fixed effects. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI		log(FDI)		log(FDI+1)	
Model	PPML		OLS		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.147***	0.155***	0.118***	0.088***	0.081***	0.070***
	(0.051)	(0.054)	(0.023)	(0.024)	(0.015)	(0.014)
FDI_{t-1}	0.000***	0.000**				
	(0.000)	(0.000)				
FDI_{t-2}		0.000				
		(0.000)				
$\log(FDI)_{t-1}$			0.643***	0.476***		
			(0.013)	(0.019)		
$\log(FDI)_{t-2}$				0.265***		
				(0.019)		
$\log(1+FDI)_{t-1}$					0.294***	0.229***
					(0.010)	(0.009)
$\log(1+FDI)_{t-2}$						0.200***
						(0.010)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.780	0.775				
R^2			0.798	0.824	0.498	0.513
Observations	38,685	32,820	20,649	14,222	38,710	32,862

Table A1.12: Social Connectedness and FDI: Investment Stocks

This table shows the results of estimating the influence of the SCI on Foreign Direct Investment Stocks and Portfolio Investment Stocks in 2019 as dependent variable using PPML. The data is taken from Coordinated Direct Investment Survey of the IMF and the Coordinated Portfolio Investment Survey of the IMF. Missing data for the year 2019 is filled with data for 2018, if available. Negative values are set to 0. Descriptive statistics for the independent variables are presented in Table 1.1. Regressions include all gravity, economic and cultural covariates and a full set of origin country and destination country fixed effects. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable:	Foreign Direct Investment Stock (2019)			Portfolio Investment Stock (2019)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)		0.525***	0.288***		0.564***	0.174*
		(0.091)	(0.071)		(0.083)	(0.093)
log(Trade volume)			0.271**			0.177***
			(0.110)			(0.067)
log(GDP differential)			-0.124***			-0.027
			(0.040)			(0.030)
log(GDP growth differential)			0.016			-0.013
			(0.047)			(0.042)
Regional trade agreement			0.193			0.229**
			(0.156)			(0.089)
Common Currency			-0.061			0.769***
			(0.173)			(0.228)
Shared religion index			1.935***			0.506
			(0.553)			(0.600)
Political distance			-0.002			-0.002
			(0.005)			(0.007)
Common legal origin			-0.020			0.176
			(0.149)			(0.120)
Common Ethno. Language			-0.053			0.121
			(0.183)			(0.209)
log(Distance)			0.068			0.040
			(0.144)			(0.072)
Common Border			-0.122			0.242
			(0.161)			(0.187)
Common Official Language			0.077			-0.119
			(0.167)			(0.165)
Common Colonizer			0.725			1.035
			(0.694)			(0.696)
Colonial Relationship after 1945			0.246			-0.127
			(0.185)			(0.208)
Origin-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.897	0.920	0.928	0.931	0.957	0.968
Observations	7,695	7,695	7,665	7,337	7,337	7,305

Table A1.13: World Governance Indicators

This table shows the presents a robustness check for the results of Table 1.3, estimating the influence of the SCI interacted with different measures for the quality of institutions in the destination and origin country on the FDI flows. Data is taken from the Worldwide Governance Indicators (WGI) project. Estimates for the institutional quality vary approximately between -2.5 (weak) and 2.5 (strong). Precise information on the collection and composition of all dimensions can be found in Kaufmann et al. (2006). Descriptive statistics can be found in Table A1.18. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Standard errors are clustered on the same level as fixed effects are defined and are depicted in parentheses. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.273***	0.267***	0.177***	0.190***	0.290***	0.288***
	(0.061)	(0.065)	(0.061)	(0.058)	(0.067)	(0.067)
log(SCI) × Dest. Control of Corruption	-0.165***					
	(0.029)					
log(SCI) × Orig. Control of Corruption	-0.065**					
	(0.030)					
log(SCI) × Dest. Rule of Law		-0.185***				
		(0.033)				
log(SCI) × Orig. Rule of Law		-0.050				
		(0.038)				
log(SCI) × Dest. Voice & Accountability			-0.133***			
			(0.040)			
log(SCI) × Orig. Voice & Accountability			0.009			
			(0.041)			
log(SCI) × Dest. Political Stability				-0.142***		
				(0.038)		
log(SCI) × Orig. Political Stability				-0.067		
				(0.042)		
log(SCI) × Dest. Government Effectiveness					-0.179***	
					(0.036)	
log(SCI) × Orig. Government Effectiveness					-0.054	
					(0.038)	
log(SCI) × Dest. Regulatory Quality						-0.197***
						(0.036)
log(SCI) × Orig. Regulatory Quality						-0.040
						(0.040)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.783	0.783	0.781	0.782	0.782	0.783
Observations	45,216	45,216	45,216	45,216	45,216	45,216

Table A1.14: Further Individual Rights and Participation Indicators and Alternatives for WGI

This table presents a robustness check for the results of Table A1.13. We interact the SCI with additional measures for the quality of institutions. The first three focus on individual rights and participation and are complemented by two corruption and two rule of law measures. Thereby we follow the Equation 1.2. We include the interaction term for the destination and origin country on the FDI. Data is taken from the Varieties of Democracy Institute, the Corruption Perception Index and the Freedom House Organization. Columns (1)-(6) include indices that vary between zero and one, the rule of law measure from Freedom House in Column (7) ranges from zero to 16. High values for the Exclusion of Socio-Economic Groups and for the Political Corruption Index indicate bad institutional quality, for all other variables higher numbers indicate higher levels of institutional quality. Descriptive statistics can be found in Table A1.18, detailed descriptions of all variables are presented in Table A1.1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(SCI)	0.570***	0.541**	-0.028	0.825***	-0.096	0.506***	0.378***
	(0.191)	(0.261)	(0.069)	(0.142)	(0.071)	(0.151)	(0.125)
log(SCI) × Dest. Access to Justice	-0.450***						
	(0.136)						
log(SCI) × Orig. Access to Justice	-0.121						
	(0.200)						
log(SCI) × Dest. Property Rights		-0.603**					
		(0.240)					
log(SCI) × Orig. Property Rights		0.099					
		(0.271)					
log(SCI) × Dest. Excl. of Socio-Econ. Groups			0.463***				
			(0.141)				
log(SCI) × Orig. Excl. of Socio-Econ. Groups			0.164				
			(0.222)				
log(SCI) × Dest. Corruption Perception Index				-0.873***			
				(0.162)			
log(SCI) × Orig. Corruption Perception Index				-0.338*			
				(0.180)			
log(SCI) × Dest. Political Corruption Index					0.486***		
					(0.105)		
log(SCI) × Orig. Political Corruption Index					0.246*		
					(0.137)		
log(SCI) × Dest. Rule of Law (V-DEM)						-0.390***	
						(0.112)	
log(SCI) × Orig. Rule of Law (V-DEM)						-0.118	
						(0.144)	
log(SCI) × Dest. Rule of Law (FH)							-0.026***
							(0.009)
log(SCI) × Orig. Rule of Law (FH)							-0.001
							(0.008)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.781	0.780	0.781	0.780	0.782	0.781	0.789
Observations	45,216	45,216	45,216	36,814	45,216	45,216	40,513

Table A1.15: Social Connectedness and Climate Risk: Physical Risk Controlling for Institutions

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with indicators for climate risk and natural disasters while controlling for interaction terms of all dimensions of the World Governance Indicators. These are: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Column (1) refers to the Global Climate Risk Index taken from Germanwatch. Higher values correspond to lower risk. Column (2) uses the number of floods per country and year, taken from the International Disaster Database (EM-DAT). Column (3) translates these data into a dummy, that is equal to one if there has been at least one flooding event in the respective country and year. Column (4) refers to the percentage of population exposed to wildfires per country and year. Data is taken from the Green Growth Indicators. In Column (5) data from Sautner et al. (2023) on physical risk arising from climate change is used. Higher values correspond to higher risk. In Columns (1) and (2), elasticity for the median, the 25^{th} and the 75^{th} percentile are presented for the Global Climate Risk Index and the Number of flooding events. Thereby the value of the interacted variable for the origin country is held at the median in all three cases. In Column (3), the elasticity of the SCI is shown for the case that there has been at least one flooding event in the destination county as well as in the origin country. Columns (4) and (5) display the elasticity of the SCI for the case that the percentage of population exposed to wildfire is at the 90^{th} percentile in the destination country while is is held at the median for the origin country because of a large fraction (>50%) of zeros in the interacted variables. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	log(FDI)				
	(1)	(2)	(3)	(4)	(5)
log(SCI)	0.446***	0.128	0.100	0.056	0.246**
	(0.100)	(0.085)	(0.087)	(0.354)	(0.112)
log(SCI) × Dest. Global Climate Risk Index	-0.364***				
	(0.082)				
log(SCI) × Orig. Global Climate Risk Index	-0.083				
	(0.081)				
log(SCI) × Dest. Number of Floods		0.041***			
		(0.013)			
log(SCI) × Orig. Number of Floods		0.036**			
		(0.015)			
log(SCI) × Dummy for Flood in Destination Country			0.120***		
			(0.044)		
log(SCI) × Dummy for Flood in Origin Country			0.080*		
			(0.043)		
log(SCI) × Dest. % Population Exposed to Wildfires				0.012**	
				(0.006)	
log(SCI) × Orig. % Population Exposed to Wildfires				0.013**	
				(0.006)	
log(SCI) × Dest. Climate Change Physical Risk Sautner et al. (2023)					0.261***
					(0.089)
log(SCI) × Orig. Climate Change Physical Risk Sautner et al. (2023)					0.276***
					(0.067)
WGI × SCI	Yes	Yes	Yes	Yes	Yes
Full Set of Controls	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms					
Origin Country at Median and Destination at:					
Median	0.140	0.205			
25^{th} Percentile	0.224	0.164			
75^{th} Percentile	0.046	0.246			
Dummy=1			0.300		
90^{th} Percentile				0.166	0.262
Pseudo R^2	0.787	0.812	0.812	0.759	0.745
Observations	39,491	24,134	24,134	10,237	17,817

Table A1.16: The Importance of Social Connectedness under Climate Risk: Transition Risk Controlling for Institutions

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with indicators for transition risk while controlling for interaction terms of all dimensions of the World Governance Indicators. These are: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Column (1) includes an interaction term of the SCI and the log of CO2-emissions in tons per capita taken from the Worldbank. Columns (2), (3) and (5) use measures for the renewable energy supply, the development of environmental related policies and energy related tax revenue from the Green Growth Indicators. Column (4) interacts the SCI with the changes in emission related policies as used by Gu and Hale (2023). Column (6) refers to climate change regulatory risk as measured by Sautner et al. (2023). At the bottom, the table reports the elasticity of SCI for different values of the interacted variable for the destination country of the investment. These are are the median, the 25^{th} and 75^{th} percentile, while the value of the interacted variable for the origin country is held at the median for all three cases. Column (6) and displays the elasticity of the SCI for the case that the climate change regulatory risk measure is at the 90^{th} percentile in the destination country while is is held at the median for the origin country because of a large fraction (>50%) of zeros in the interacted variable. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.371***	-0.177*	0.191*	0.072	0.130	0.243**
	(0.100)	(0.098)	(0.104)	(0.096)	(0.143)	(0.112)
log(SCI) × Dest. log(CO2 Emissions)	-0.181***					
	(0.045)					
log(SCI) × Orig. log(CO2 Emissions)	-0.027					
	(0.041)					
log(SCI) × Dest. Renewable Energy Supply		0.973***				
		(0.187)				
log(SCI) × Orig. Renewable Energy Supply		0.453**				
		(0.194)				
log(SCI) × Dest. Development of Env.-related Technologies			-0.041			
			(0.298)			
log(SCI) × Orig. Development of Env.-related Technologies			-0.230			
			(0.351)			
log(SCI) × Dest. Changes in Emission Related Policies				0.894**		
				(0.437)		
log(SCI) × Orig. Changes in Emission Related Policies				-0.544		
				(0.417)		
log(SCI) × Dest. Energy Related Tax Revenue					0.077	
					(0.097)	
log(SCI) × Orig. Energy Related Tax Revenue					0.037	
					(0.144)	
log(SCI) × Dest. Climate Change Regulatory Risk Sautner et al. (2023)						1.554***
						(0.490)
log(SCI) × Orig. Climate Change Regulatory Risk Sautner et al. (2023)						-0.298
						(0.276)
WGI × SCI	Yes	Yes	Yes	Yes	Yes	Yes
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms						
Origin Country at Median and Destination at:						
Median	0.040	0.008	0.158	0.086	0.213	
25^{th} Percentile	0.150	-0.040	0.160	0.059	0.200	
75^{th} Percentile	-0.047	0.135	0.157	0.122	0.219	
90^{th} Percentile						0.305
Pseudo R^2	0.789	0.790	0.786	0.785	0.780	0.743
Observations	44,522	42,461	41,250	34,191	29,975	17,817

Table A1.17: The Importance of Social Connectedness under Climate Risk: Transition Risk Controlling for Institutions and Economic Factors

This table shows the results for Equation 1.2, estimating the influence of the SCI interacted with indicators for transition risk while controlling for interaction terms of all dimensions of the World Governance Indicators. These are: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. In addition, the regression includes interactions of the SCI and bilateral economic controls (log bilateral trade, log GDP differences, log GDP growth differences, a dummy for a regional trade agreement lagged by three years and a dummy for a common currency) and unilateral economic controls (a country's development status, the log of GDP and GDP growth). Descriptive statistics can be found in Table 1.1, detailed descriptions of all variables are presented in Table A1.1. Column (1) includes an interaction term of the SCI and the log of CO2-emissions in tons per capita taken from the Worldbank. Columns (2), (3) and (5) use measures for the renewable energy supply, the development of environmental related policies and energy related tax revenue from the Green Growth Indicators. Column (4) interacts the SCI with the changes in emission related policies as used by Gu and Hale (2023). Column (6) refers to climate change regulatory risk as measured by Sautner et al. (2023). At the bottom, the table reports the elasticity of SCI for different values of the interacted variable for the destination country of the investment. These are are the median, the 25^{th} and 75^{th} percentile, while the value of the interacted variable for the origin country is held at the median for all three cases. Column (6) and displays the elasticity of the SCI for the case that the climate change regulatory risk measure is at the 90^{th} percentile in the destination country while is is held at the median for the origin country because of a large fraction (>50%) of zeros in the interacted variable. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by origin and destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	-1.043	-1.424	-0.835	-0.704	-0.529	-2.234*
	(0.901)	(0.896)	(0.915)	(1.055)	(1.112)	(1.246)
log(SCI) × Dest. log(CO2 Emissions)	-0.190***					
	(0.043)					
log(SCI) × Orig. log(CO2 Emissions)	-0.015					
	(0.045)					
log(SCI) × Dest. Renewable Energy Supply		0.962***				
		(0.184)				
log(SCI) × Orig. Renewable Energy Supply		0.446**				
		(0.218)				
log(SCI) × Dest. Development of Env.-related Technologies			0.128			
			(0.246)			
log(SCI) × Orig. Development of Env.-related Technologies			-0.234			
			(0.343)			
log(SCI) × Dest. Changes in Emission Related Policies				0.730*		
				(0.425)		
log(SCI) × Orig. Changes in Emission Related Policies				-0.443		
				(0.412)		
log(SCI) × Dest. Energy Related Tax Revenue					0.131	
					(0.095)	
log(SCI) × Orig. Energy Related Tax Revenue					0.102	
					(0.145)	
log(SCI) × Dest. Climate Change Regulatory Risk Sautner et al. (2023)						1.107**
						(0.443)
log(SCI) × Orig. Climate Change Regulatory Risk Sautner et al. (2023)						-0.390
						(0.306)
WGI × SCI	Yes	Yes	Yes	Yes	Yes	Yes
Interactions with Bilateral Economic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Interactions with Unilateral Economic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Elasticity of SCI for Interaction Terms						
Origin Country at Median and Destination at:						
Median	-1.368	-1.241	-0.848	-0.693	-0.362	
25^{th} Percentile	-1.252	-1.289	-0.852	-0.714	-0.383	
75^{th} Percentile	-1.459	-1.116	-0.844	-0.663	-0.350	
90^{th} Percentile						-2.190
Pseudo R^2	0.792	0.793	0.789	0.788	0.783	0.748
Observations	44,522	42,461	41,250	34,191	29,975	17,817

Table A1.18: Descriptive Statistics for Additional Controls

This table shows summary statistics for the additional variables used in the appendix to identify the channels through which social connectedness influences FDI.

	Mean	Median	SD	Min	Max	Obs
Dest. Control of Corruption	0.41	0.19	1.00	-1.67	2.41	45,345
Orig. Control of Corruption	0.55	0.37	1.09	-1.82	2.41	45,281
Dest. Rule of Law	0.50	0.43	0.93	-1.59	2.13	45,345
Orig. Rule of Law	0.59	0.54	1.00	-1.85	2.13	45,281
Dest. Voice & Accountability	0.48	0.66	0.83	-2.07	1.74	45,345
Orig. Voice & Accountability	0.44	0.60	0.93	-2.12	1.74	45,281
Dest. Political Stability	0.19	0.39	0.82	-2.81	1.64	45,345
Orig. Political Stability	0.23	0.41	0.87	-2.81	1.64	45,281
Dest. Government Effectiveness	0.58	0.53	0.84	-1.64	2.24	45,345
Orig. Government Effectiveness	0.65	0.65	0.95	-2.08	2.24	45,281
Dest. Regulatory Quality	0.64	0.65	0.81	-2.24	2.26	45,345
Orig. Regulatory Quality	0.65	0.71	0.93	-2.35	2.26	45,281
Dest. Access to Justice	0.80	0.89	0.21	0.18	1.00	45,345
Orig. Access to Justice	0.80	0.89	0.21	0.09	1.00	45,345
Dest. Property Rights	0.85	0.91	0.12	0.23	0.97	45,345
Orig. Property Rights	0.83	0.90	0.15	0.12	0.97	45,345
Dest. Excl. of Socio-Econ. Groups	0.24	0.14	0.23	0.01	0.96	45,345
Orig. Excl. of Socio-Econ. Groups	0.24	0.16	0.23	0.01	0.96	45,345
Dest. Corruption Perception Index	0.53	0.50	0.19	0.15	0.92	37,128
Orig. Corruption Perception Index	0.55	0.53	0.21	0.14	0.92	36,889
Dest. Political Corruption Index	0.36	0.28	0.30	0.00	0.95	45,345
Orig. Political Corruption Index	0.33	0.22	0.30	0.00	0.97	45,345
Dest. Rule of Law (V-DEM)	0.73	0.83	0.27	0.03	1.00	45,345
Orig. Rule of Law (V-DEM)	0.73	0.84	0.28	0.03	1.00	45,345
Dest. Rule of Law (FH)	10.42	11.00	4.15	0.00	16.00	41,213
Orig. Rule of Law (FH)	10.34	11.00	4.58	0.00	16.00	40,620

Bibliography

Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. American Economic Review, 112(4): 1194–1225.

Damgaard, J., Elkjaer, T., and Johannesen, N. (2024). What is real and what is not in the global FDI network? Journal of International Money and Finance, 140: 102971.

Gu, G. W. and Hale, G. (2023). Climate risks and FDI. Journal of International Economics, 146: 103731.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring Uncertainty. American Economic Review, 105(3): 1177–1216.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2006). Governance matters V: Aggregate and individual governance indicators for 1996-2005. World Bank Policy Research Working Paper, (4012).

Lin, W. and Wooldridge, J. M. (2019). Testing and correcting for endogeneity in nonlinear unobserved effects models. In Panel Data Econometrics, pages 21–43. Elsevier.

Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and business cycles: Exogenous impulse or endogenous response? American Economic Journal: Macroeconomics, 13(4): 369–410.

Sautner, Z., Van Lent, L., Vilkov, G., and Zhang, R. (2023). Firm-level climate change exposure. The Journal of Finance, 78(3): 1449–1498.

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

Spolaore, E. and Wacziarg, R. (2018). Ancestry and development: New evidence. Journal of Applied Econometrics, 33(5): 748–762.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2): 420–445.

APPENDIX A2

Appendix for Chapter 2

A2.1 Implementation: Spatial Durbin Model

A census tract in the US is defined as an area with around $4,000$ inhabitants and a relatively homogeneous distribution regarding socio-economic characteristics. It therefore more likely captures the operating area of a bank branch. Still, on the demand side, many tracts that exhibit fintech loans do not host a bank branch. Borrowers might be located in neighboring tracts to those hosting a physical bank branch. People could also locally move but do not change their bank account but keep on using a branch that locates in an adjacent tract. It is reasonable to assume, that the fintech shares of neighboring tracts (as well as all other covariates) affect the number of bank branches in a tract. On the supply side, operating a branch in one tract might be sufficient to cover demand of neighboring tracts as well. The number of branches in a tract is also dependent of the number of branches in neighboring branches. Taking these dynamics into account, I estimate a Spatial Durbin Model (SDM) model allowing for endogenous spatial spillovers. A detailed discussion on this type of spatial model is provided by LeSage and Pace (2009), Elhorst (2010) and Elhorst (2014). It draws on the early contributions of Durbin (1960), Ord (1975) and Manski (1993).

Spatial Durbin Models (SDMs) are a class of spatial econometric models that explicitly account for spatial dependence in both the dependent and independent variables. These models are particularly suitable for examining phenomena where the effects of explanatory variables are not confined to the units under observation but also spill over to neighboring units. In a typical regression analysis, we assume that observations are independent of one another. However, this assumption is often violated in spatial data where geographic units, such as census tracts, can exhibit spatial autocorrelation. That is, values observed in one location are influenced by values in nearby locations. SDMs address this issue by incorporating spatial lags of both the dependent variable and the independent variables into the regression model.

In the context of analyzing the impact of rising fintech market shares on bank branches at the census tract level, SDMs are particularly appropriate for several reasons: First, the presence and performance of bank branches in one tract are likely influenced by the characteristics of neighboring tracts. For instance, if a fintech lender captures a significant market share in one tract, nearby tracts might also experience a shift in banking behavior due to customer mobility and regional economic factors. Second, fintech market penetration often follows patterns that transcend administrative boundaries. The use of spatial lags in SDMs allows to capture the diffusion of fintech's influence across neighboring tracts. Third, the closure of bank branches can have ripple effects, influencing local economies and communities beyond the immediate tract. By including spatial lags of the independent variables, SDMs account for these externalities, offering insights into how fintech-induced branch closures might propagate through space. Fourth, standard regression models that ignore spatial dependencies can suffer from biased estimates due to omitted variable bias stemming from spatial autocorrelation. SDMs mitigate this problem by explicitly modeling the spatial processes, leading to more reliable and valid inferences.

$$Branches_{i,t+2} = \rho \times \sum_{i \neq j} w_{ij} \times Branches_{j,t+2} + \beta \times X_{it} + \theta \times \sum_{i \neq j} w_{ij} \times X_{jt} + \delta_i + \epsilon_{i,t}, \tag{A2.1}$$

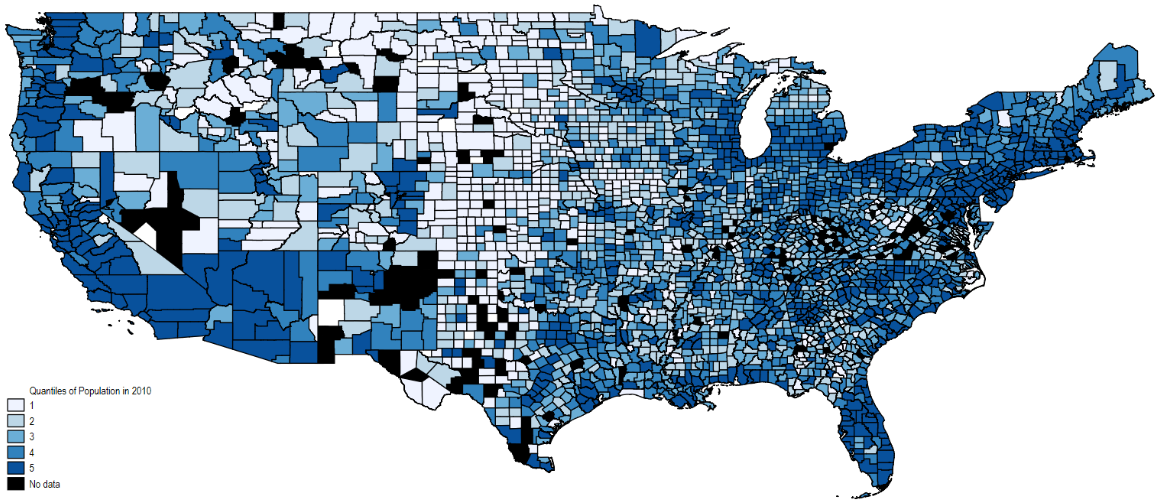
The model estimates the number of bank branches, lagged by two years, in county i with respect to the fintech share and the full set of controls, both subsumed in x_{it}. δ_i include census tract fixed effects, the term $\epsilon_{i,t}$ represents the error term. In addition, the model allows for endogenous effects between census tracts: The term $\sum_{i \neq j} w_{ij} \times Branches_{jt}$ refers to the spatially lagged number of bank branches of neighboring census tracts of tract i in year t. The preceding parameter ρ captures this endogenous interaction effect. The term $\sum_{i \neq j} w_{ij} \times x_{jt}$ represents spatial lags of the independent variables, e.g. the fintech share of adjacent tracts to tract i in year t. This exogenous interaction effect is captured by vector θ. Covariates enter on tract level. Only GDP which is not available and enters on county level. For the sake of simplicity, however, I stick to the vector notation with i for tract. δ_i captures county fixed effects. I use the Stata command `spxtregress` applying a quasi-maximum likelihood estimator including individual fixed effects following Lee and Yu (2010). The spatial weights matrix governing the neighbor-relationships is a queen-type row normalized matrix. To limit the impact of boundary changes in the census tracts I restrict the panel to the period from 2012 to 2017. I further limit the analysis to the state of Illinois which is closest to the whole US in terms of racial markup, education, age and income of the population regarding to the census data. In addition, there are no missing values for tracts in Illinois which is a crucial condition.

A2.2 Appendix Figures

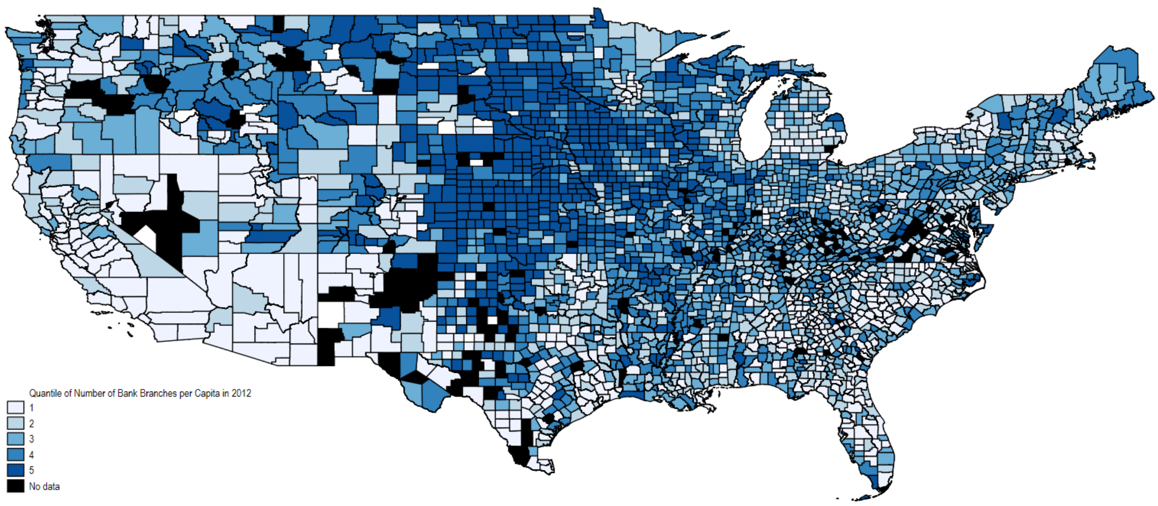
Figure A2.1: Distribution of Population and Branches per Capita

Panel A shows the population in 2010 on county level by quantiles. Darker shading represents a higher population. The population is highest in the Northwest as well as counties in California, Arizona, Florida and around centers such as Chicago, Denver, and Seattle. By far the most populated county is Los Angeles County with a population of just over 10 million people, followed by Cook County around Chicago (5.2 million). The lowest population of 455 people can be found in Arthur County in Nebraska. Panel B presents the number of bank branches per 10,000 people on county level in 2012 by quintiles. Darker areas show a higher number of bank branches per population.

Panel A: Population (2010)



Panel B: Number of Bank Branches per Capita (2012)



A2.3 Appendix Tables

Table A2.1: Variable Definitions and Data Sources

Variable Name	Description
Main Variables	
Bank branches per County	Number of bank branches in county i in year t; source: Federal Deposit Insurance Corporation (FDIC)
Bank branches from Small Banks per County	Number of bank branches in county i in year t that are operated by banks with less than 10bn USD in assets; source: Federal Deposit Insurance Corporation (FDIC)
Fintech Market Share	Share of mortgages by value in county i in year t, that is originated by a lender classified as a fintech leder; source: Consumer Financial Protection Bureau (CFPB)
Fintech market Share in purchases	Share of mortgages for home purchases by value in county i in year t, that is originated by a lender classified as a fintech leder; source: Consumer Financial Protection Bureau (CFPB)
Fintech market Share in Refinancings	Share of mortgages for refinancings by value in county i in year t, that is originated by a lender classified as a fintech leder; source: Consumer Financial Protection Bureau (CFPB)
Fintech market Share Fuster et al. (2019)	Share of mortgages by value in county i in year t, that is originated by a lender classified as a fintech leder as computed by Fuster et al. (2019); source: Fuster et al. (2019)
Institutes per County	Number of banks operating one or more branches in county i in year t; source: Federal Deposit Insurance Corporation (FDIC)
Deposits per County	Total county-specific deposits held by all branches in county i in year t in $1,000$ USD; source: Federal Deposit Insurance Corporation (FDIC)
Assets per County	Total national assets held by all banks that are active in county i in year t in $1,000$ USD; source: Federal Deposit Insurance Corporation (FDIC)
Instrumental Variables	

Table A2.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Social Connectedness Index	Relative probability of a user in county i being friends with a user in Wayne County on Facebook taking values between 1 and 1 billion, winsorized at the 99^{th} percentile; source: Humanitarian Data Exchange
Control Variables	
Population	Population in county i in year t; source: United states Census Bureau
County GDP	GDP in county i in year t; source: U.S. Bureau of Economic Analysis
Median Home Value	Median home value of owner-occupied housing units in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Building Permits	Number of building permits in terms of units in county i in year t; source: Building Permits Survey (BPS)
Median Household Income	Median household income in the past 12 months of year t in county i(in dollars, inflation-adjusted to year t) ; source: American Community Survey accessed via the Social Determinants of Health Database
Percentage of Population over 64	Percentage of population age 65 and over in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Percentage of Unemployed Population	Percentage of population that was unemployed (ages 16 years and over) in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Percentage of Population with Bachelor Degree	Percentage of population with a bachelor's degree (ages 25 and over) in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Percentage of Population reporting White Race	Percentage of population reporting White race in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database

Table A2.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Percentage of Population Reporting Black or African American Race	Percentage of population reporting Reporting Black or African American race in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Percentage of Female Population	Percentage of population that is female in county i in year t; source: American Community Survey accessed via the Social Determinants of Health Database
Additional Variables	
Metro Dummy	Dummy equal to one if the county is located within a Major Statistical Area (MSA) of the US; source: Economic Research Service U.S. Department of Agriculture

Table A2.2: Fintech Mortgage Lenders

This table lists the fintech mortgage lenders in my sample and the year in which they first enter the sample as a fintech based on Buchak et al. (2018) and Fuster et al. (2019).

	Included since	Buchak et al. (2018)	Fuster et al. (2019)
AMERISAVE MORTGAGE CORPORATION	2010	2008	-
CASHCALL INC	2010	2008	-
GUARANTEED RATE INC	2010	2008	2010
QUICKEN LOANS, INC	2010	2000	2010
MOVEMENT MORTGAGE, LLC	2014	2013	2014
AVEX FUNDING CORPORATION	2016	-	2016
EVERETT FINANCIAL INC	2016	-	2016
LOANDEPOT INC	2016	-	2016
21ST MORTGAGE CORP.	2017	-	2017
AMERICAN INTERNET MORTGAGE	2017	-	2017
AMERICAN NEIGHBORHOOD MORTGAGE	2017	-	2017
ARK-LA-TEX FINANCIAL SERVICES	2017	-	2017
ENVOY MORTGAGE, LTD	2017	-	2017
EVERGREEN MONEYSOURCE MORTGAGE	2017	-	2017
FBC MORTGAGE, LLC	2017	-	2017
HOMEWARD RESIDENTIAL, INC	2017	-	2017
MORTGAGE INVESTORS GROUP	2017	-	2017
RPM MORTGAGE INC	2017	-	2017
SKYLINE FINANCIAL CORP	2017	-	2017

Table A2.3: Correlation Matrix

This table displays the correlation between the most important variables throughout the paper.

	Bank Branches (2012-2019)	Δ Bank Branches (2012-2019)	Fintech Market Share	Δ Fintech Market Share
Bank Branches	1.00	-0.79	-0.08	-0.07
Δ Number of Bank Branches (2012-2019)	-0.78	1.00	0.03	0.03
Fintech Market Share	-0.08	0.03	1.00	0.92
Δ Fintech Market Share	-0.07	0.03	0.92	1.00
Fintech Market Share in Purchases	-0.04	0.01	0.75	0.70
Fintech Market Share in Refinancings	-0.09	0.04	0.72	0.63
Fintech Market Share Fuster et al. (2019)	-0.08	0.03	0.93	0.85
log(Total Population)	0.61	-0.56	-0.15	-0.13
log(County GDP)	0.63	-0.57	-0.18	-0.14
log(Median Home Value)	0.39	-0.32	-0.22	-0.20
log(Building Permits)	0.49	-0.45	-0.22	-0.17
log(Median Household Income)	0.29	-0.24	-0.28	-0.25
% of Population Over 64	-0.22	0.16	0.11	0.07
% of Unemployed Population	0.11	-0.15	0.18	0.16
% of Population with Bachelor Degree	0.37	-0.29	-0.25	-0.21
% of Population Reporting White Race	-0.19	0.17	-0.16	-0.15
% of Population Reporting Black or African American Race	0.07	-0.09	0.18	0.16
% of Female Population	0.14	-0.13	-0.05	-0.04
Purged SCI to Wayne County (Detroit)	0.14	-0.13	0.20	0.19

Table A2.4: Spatial Model

This table shows the results from estimating regression A2.1. Following LeSage and Pace (2009), Columns (1)-(3) show the results for a Spatial Durbin Model, reporting direct, indirect and total effect for the fintech share as the covariate effect. The dependent variable is the Number of Bank Branches in census tract i in year t. The main explanatory variable is the Fintech Market Share of census tract i in year t in Column (1), the Fintech Market Share in Purchases in Columns (2) and the Fintech Market Share in Refinancings in Column (3). The model controls on county level for the log of total population, the log of the median home value, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. Log GDP and the log of building permits vary on county level only, all other control variables enter on the tract level. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. The spatial weights matrix governing the neighbor-relationships is a simple queen-type row normalized matrix. Regressions include individual fixed effects as specified by Lee and Yu (2010). Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	Bank Branches (2014-2019)		
	(1)	(2)	(3)
Bank Branches (2014-2019)	-0.299***	-0.298***	-0.298***
	(0.018)	(0.018)	(0.018)
Direct Effects			
Fintech Market Share	-0.098		
	(0.083)		
Fintech Market Share in Purchases		0.005	
		(0.064)	
Fintech Market Share in Refinancings			-0.078
			(0.054)
Indirect Effects			
Fintech Market Share	-0.248*		
	(0.128)		
Fintech Market Share in Purchases		-0.285***	
		(0.105)	
Fintech Market Share in Refinancings			-0.130**
			(0.092)
Total Effects			
Fintech Market Share	-0.346***		
	(0.116)		
Fintech Market Share in Purchases		-0.280***	
		(0.098)	
Fintech Market Share in Refinancings			-0.208**
			(0.087)
County FE	Yes	Yes	Yes
AIC	13,137.92	13,138.49	13,140.54
BIC	13,341.7	13,342.27	13,344.32
Observations	18,726	18,726	18,726

Table A2.5: Excluding Outliers

This table shows the results from estimating Equation 2.1 while excluding the four outstanding counties with more than 700 bank branches as well as counties with fintech shares above 30%. The dependent variable is change the number of bank branches in county i between 2012 and 2019. The main explanatory variables are the changes from 2010 to 2017 in fintech market share in Column (1), in the fintech market share in purchases in Column (2) and in refinancings in Column (3) and the change in the fintech market share from Fuster et al. (2019) in Column (4). The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017. Standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Bank Branches (2012-2019))			
	(1)	(2)	(3)	(4)
Δ Fintech Market Share	-0.206***			
	(0.078)			
Δ Fintech Market Share in Purchases		-0.084		
		(0.083)		
Δ Fintech Market Share in Refinancings			-0.167***	
			(0.046)	
Δ Fintech Market Share Fuster et al. (2019)				-0.297***
				(0.086)
Δ log(Total Population)	0.051	0.054	0.056	0.128**
	(0.062)	(0.061)	(0.062)	(0.062)
Δ log(County GDP)	0.023	0.023	0.020	-0.001
	(0.015)	(0.015)	(0.015)	(0.016)
Δ log(Median Home Value)	0.207***	0.208***	0.205***	0.245***
	(0.029)	(0.029)	(0.029)	(0.029)
Δ log(Building Permits)	0.002	0.002	0.002	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Δ log(Median Household Income)	0.029	0.030	0.030	0.057
	(0.047)	(0.047)	(0.047)	(0.047)
Δ % of Population Over 64	-0.524**	-0.541**	-0.478*	-0.599**
	(0.266)	(0.266)	(0.268)	(0.273)
Δ % of Unemployed Population	0.526***	0.539***	0.529***	0.496***
	(0.174)	(0.175)	(0.174)	(0.182)
Δ % of Population with Bachelor Degree	-0.111	-0.077	-0.122	-0.308
	(0.220)	(0.223)	(0.219)	(0.221)
Δ % of Population Reporting White Race	-0.050	-0.061	-0.056	-0.000
	(0.109)	(0.111)	(0.107)	(0.110)
Δ % of Population Reporting Black or African American Race	0.174	0.165	0.211	0.436
	(0.317)	(0.317)	(0.315)	(0.346)
Δ % of Female Population	-0.101	-0.085	-0.071	0.166
	(0.341)	(0.340)	(0.342)	(0.405)
R^2	0.064	0.062	0.066	0.079
Observations	2,855	2,855	2,855	2,677

Table A2.6: Excluding Outliers in Panel

This table shows the results from estimating Equation 2.2 using an OLS panel data estimator in Column (1) and (2) and a PPML panel data estimator in Column (3) and (4) while excluding the four outstanding counties with more than 700 bank branches as well as counties with fintech shares above 30%. The dependent variable is the Number of Bank Branches in County i, logarithmized for the OLS case. The main explanatory variable is the fintech market share. The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. Further δ_i and δ_t capture county and year fixed effects, absorbing all county-specific as well as year-specific common shocks. Standard errors are clustered on metro level. Columns (2) and (4) serve as a robustness check by using the fintech market share data from Fuster et al. (2019). Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	log(Bank Branches (2012-2019))		Bank Branches (2012-2019)	
Model	OLS		PPML	
	(1)	(2)	(3)	(4)
Fintech Market Share	-0.091**		-0.110**	
	(0.034)		(0.055)	
Fintech Market Share Fuster et al. (2019)		-0.109***		-0.129***
		(0.038)		(0.047)
log(Total Population)	0.082	0.084	0.197***	0.202***
	(0.060)	(0.061)	(0.055)	(0.055)
log(County GDP)	0.013	0.013	0.032***	0.032***
	(0.010)	(0.010)	(0.012)	(0.012)
log(Median Home Value)	0.159***	0.158***	0.122***	0.122***
	(0.023)	(0.023)	(0.032)	(0.031)
log(Building Permits)	0.001	0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
log(Median Household Income)	0.043	0.043	0.102**	0.100**
	(0.038)	(0.038)	(0.048)	(0.048)
% of Population Over 64	-0.476	-0.476	-0.668**	-0.679***
	(0.289)	(0.289)	(0.260)	(0.262)
% of Unemployed Population	0.483***	0.481***	0.750***	0.747***
	(0.172)	(0.172)	(0.213)	(0.213)
% of Population with Bachelor Degree	-0.290**	-0.289**	-0.175	-0.178
	(0.122)	(0.122)	(0.120)	(0.119)
% of Population Reporting White Race	-0.092	-0.090	-0.026	-0.026
	(0.130)	(0.130)	(0.070)	(0.070)
% of Population Reporting Black or African American Race	0.307	0.310	0.286	0.285
	(0.223)	(0.223)	(0.318)	(0.319)
% of Female Population	0.124	0.128	0.042	0.058
	(0.222)	(0.221)	(0.235)	(0.234)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.996	0.996		
Pseudo R^2			0.923	0.923
Observations	22,486	22,486	22,486	22,486

Table A2.7: Branches per Population

This table shows the results from estimating Equation 2.1. The dependent variable is the Number of Bank Branches in County i per $10,000$ people. The main explanatory variables are the Fintech Market Share in Column (1), the Fintech market Share in Purchases in Column (2) and Refinancings in Column (3) and the Fintech Market Share from Fuster et al. (2019) in Column (4).The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017. Standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(Bank Branches per 10,000 People (2012-2019))			
	(1)	(2)	(3)	(4)
Δ Fintech Market Share	-0.133*			
	(0.070)			
Δ Fintech Market Share in Purchases		-0.076		
		(0.071)		
Δ Fintech Market Share in Refinancings			-0.087**	
			(0.043)	
Δ Fintech Market Share Fuster et al. (2019)				-0.151**
				(0.075)
Δ log(Total Population)	-0.974***	-0.973***	-0.969***	-0.904***
	(0.057)	(0.056)	(0.057)	(0.059)
Δ log(County GDP)	0.029**	0.030**	0.027**	0.006
	(0.013)	(0.013)	(0.013)	(0.016)
Δ log(Median Home Value)	0.186***	0.187***	0.185***	0.207***
	(0.028)	(0.028)	(0.028)	(0.028)
Δ log(Building Permits)	-0.004	-0.004	-0.004	-0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Δ log(Median Household Income)	-0.041	-0.040	-0.040	-0.013
	(0.044)	(0.044)	(0.044)	(0.044)
Δ % of Population Over 64	-0.461*	-0.471*	-0.438*	-0.652**
	(0.253)	(0.252)	(0.254)	(0.272)
Δ % of Unemployed Population	0.537***	0.543***	0.542***	0.557***
	(0.165)	(0.166)	(0.165)	(0.174)
Δ % of Population with Bachelor Degree	-0.212	-0.192	-0.210	-0.425**
	(0.199)	(0.200)	(0.198)	(0.198)
Δ % of Population Reporting White Race	-0.047	-0.051	-0.054	0.034
	(0.095)	(0.096)	(0.095)	(0.108)
Δ % of Population Reporting Black or African American Race	0.142	0.132	0.157	0.254
	(0.314)	(0.315)	(0.313)	(0.369)
Δ % of Female Population	-0.088	-0.083	-0.072	-0.006
	(0.342)	(0.341)	(0.342)	(0.435)
R^2	0.155	0.155	0.155	0.161
Observations	2,907	2,907	2,907	2,693

Table A2.8: Preference for Convenience

This table shows the results from estimating Equation 2.1. The dependent variable is the change in the Number of McDonald's Restaurants between 2010 and 2017 in County i in Column (1) and the Number of Fast Food Restaurants between 2011 and 2016 in County i in Column (2). Data is taken from the Food Environment Atlas. The main explanatory variable is the change are the Fintech Market Share. The model controls on county level for the log of total population, the log of GDP, the log of the median home value, the log of building permits, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years. All other variables enter as the difference between 2010 and 2017. Standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Δ log(McDonald's Restaurants (2010-2017))	Δ log(Fast Food Restaurants (2011-2016))
	(1)	(2)
Δ Fintech Market Share	0.180*	0.389***
	(0.098)	(0.135)
Δ log(Total Population)	0.262***	0.824***
	(0.083)	(0.144)
Δ log(County GDP)	-0.029	0.013
	(0.021)	(0.038)
Δ log(Median Home Value)	0.017	0.019
	(0.037)	(0.066)
Δ log(Building Permits)	-0.001	-0.009
	(0.005)	(0.007)
Δ log(Median Household Income)	-0.098	0.042
	(0.066)	(0.095)
Δ % of Population Over 64	0.049	1.185**
	(0.367)	(0.487)
Δ % of Unemployed Population	-0.140	-0.379
	(0.222)	(0.463)
Δ % of Population with Bachelor Degree	0.656*	-0.460
	(0.336)	(0.434)
Δ % of Population Reporting White Race	0.063	0.230
	(0.219)	(0.323)
Δ % of Population Reporting Black or African American Race	0.334	0.761
	(0.428)	(0.632)
Δ % of Female Population	0.313	-0.641
	(0.485)	(0.868)
R^2	0.010	0.051
Observations	2,215	2,770

Table A2.9: Purged Social Connectedness Index

This table shows the construction of the Purged Social Connectedness Index to Wayne County (Detroit) which is used as an instrumental variable for the change in fintech market share. The dependent variable is the change in the Fintech Market Share in County i between 2010 and 2017. The model controls on county level for the log of Distance to Wayne County, a dummy indicating that the county is located in the same state as Wayne County (Michigang), a dummy indicating that the county shares a border with Wayne County, a dummy indicating that the county lies within the commuting zone of Wayne County and a dummy indicating that the county is inside a metropolitan area. The Residuals from this regressions are the used as instrumental variable, labeled Purged Social Conceitedness Index. Standard errors are clustered on metro level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	SCI to Wayne County (Detroit)
	(1)
log(Distance to Wayne County (Detroit))	-0.353***
	(0.022)
Same State	2.331***
	(0.077)
Common Border	1.215***
	(0.076)
Same Commuting-Zone	0.261***
	(0.099)
Indicator for Metro Counties	0.391***
	(0.043)
R^2	0.427
Observations	2,906

Table A2.10: Spatial Durbin Model

This table shows the coefficients from estimating Equation A2.1. The dependent variable is the Number of Bank Branches in County i and year t. The main explanatory variable is the Fintech Market Share of census tract i in year t in Column (1), the Fintech Market Share in Purchases in Column (2) and the Fintech Market Share in Refinancings in Column (3), all between 2012 and 2017. The model controls on county level for the log of total population, the log of the median home value, the log of median household income, the percentage of population over 64, the percentage of unemployed population, the percentage of population with bachelor degree, the percentage of population reporting white race and the percentage of female population. Log GDP and the log of building permits vary on county level only, all other control variables enter on the tract level. To account for the fact that banks need a certain time to react to changing market structures and adjust their branch network, all dependent variables enter the model with a lag of 2 years (2014-2019). The spatial weights matrix governing the neighbor relationships is a simple queen-type row normalized matrix. Regressions include individual fixed effects as specified by Lee and Yu (2010). Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	Bank Branches (2014-2019)		
	(1)	(2)	(3)
Fintech Market Share	-0.109		
	(0.081)		
Fintech Market Share in Purchases		-0.008	
		(0.062)	
Fintech Market Share in Refinancings			-0.084
			(0.053)
log(County GDP)	0.007	0.006	0.006
log(Building Permits)	0.014	0.014	0.015
log(Total Population in Tract)	0.063	0.070	0.064
log(Median Home Value in Tract)	-0.030	-0.027	-0.029
log(Median Household Income in Tract)	-0.050**	-0.053**	-0.050**
Percentage of Population Over 65 in Tract	-0.040	-0.061	-0.051
Percentage of Unemployed Population in Tract	0.004	0.003	0.014
Percentage of Population with Bachelor Degree in Tract	0.024	0.017	0.027
Percentage of Population Reporting White Race in Tract	-0.138	-0.133	-0.142
Percentage of Population Reporting Black or African American Race in Tract	-0.129	-0.131	-0.136
Percentage of Female Population in Tract	0.076	0.076	0.072
W			
Fintech Market Share	-0.387**		
	(0.175)		
Fintech Market Share in Purchases		-0.409***	
		(0.144)	
Fintech Market Share in Refinancings			-0.211*
			(0.127)
log(County GDP)	-0.147	-0.148	-0.153
log(Building Permits)	-0.061***	-0.061***	-0.061***
log(Total Population in Tract)	0.128	0.143	0.121
log(Median Home Value in Tract)	0.218***	0.229***	0.223***
log(Median Household Income in Tract)	-0.175***	-0.187***	-0.177***
Percentage of Population Over 65 in Tract	-3.059***	-3.165***	-3.159***
Percentage of Unemployed Population in Tract	0.583**	0.592**	0.679**
Percentage of Population with Bachelor Degree in Tract	0.136	0.099	0.149
Percentage of Population Reporting White Race in Tract	0.533**	0.553***	0.533**
Percentage of Population Reporting Black or African American Race in Tract	-0.959***	-0.955***	-0.981***
Percentage of Female Population in Tract	0.039	0.029	0.032
Bank Branches (2014-2019)	-0.299***	-0.299***	-0.299***
County FE	Yes	Yes	Yes
Pseudo R^2	0.024	0.025	0.023
Observations	18,726	18,726	18,726

Bibliography

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. Journal of Financial Economics, 130(3): 453–483.

Durbin, J. (1960). Estimation of parameters in time-series regression models. Journal of the Royal Statistical Society: Series B (Methodological), 22(1): 139–153.

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. Spatial Economic Analysis, 5(1): 9–28.

Elhorst, J. P. (2014). Spatial econometrics from cross-sectional data to spatial panels. Springer.

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The role of technology in mortgage lending. The Review of Financial Studies, 32(5): 1854–1899.

Lee, L.-f. and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. Journal of Econometrics, 154(2): 165–185.

LeSage, J. and Pace, R. K. (2009). Introduction to spatial econometrics. Chapman and Hall/CRC.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The Review of Economic Studies, 60(3): 531–542.

Ord, K. (1975). Estimation methods for models of spatial interaction. Journal of the American Statistical Association, 70(349): 120–126.

APPENDIX A3

Appendix for Chapter 3

A3.1 Appendix Figures

Figure A3.1: Flood-weighted, Distance-purged Social Connectedness

This figure illustrates quintiles of the flood weighted social connectedness of all regions that did not reported any insurance claims regarding the flood in July 2021 to the flooded regions (colored in red) purged by distance. Log SCI is regressed on log distance, both variables weighted by the flood intensity (see Table A3.3). Residuals of this regression are displayed by quintiles. Other regions that exhibit a positive claim ratio serve as a buffer category and are illustrated in white. Darker shading represents a higher connectedness. Data on social connectedness is taken from Bailey et al. (2018) and has been introduced for Europe by Aref et al. (2020).

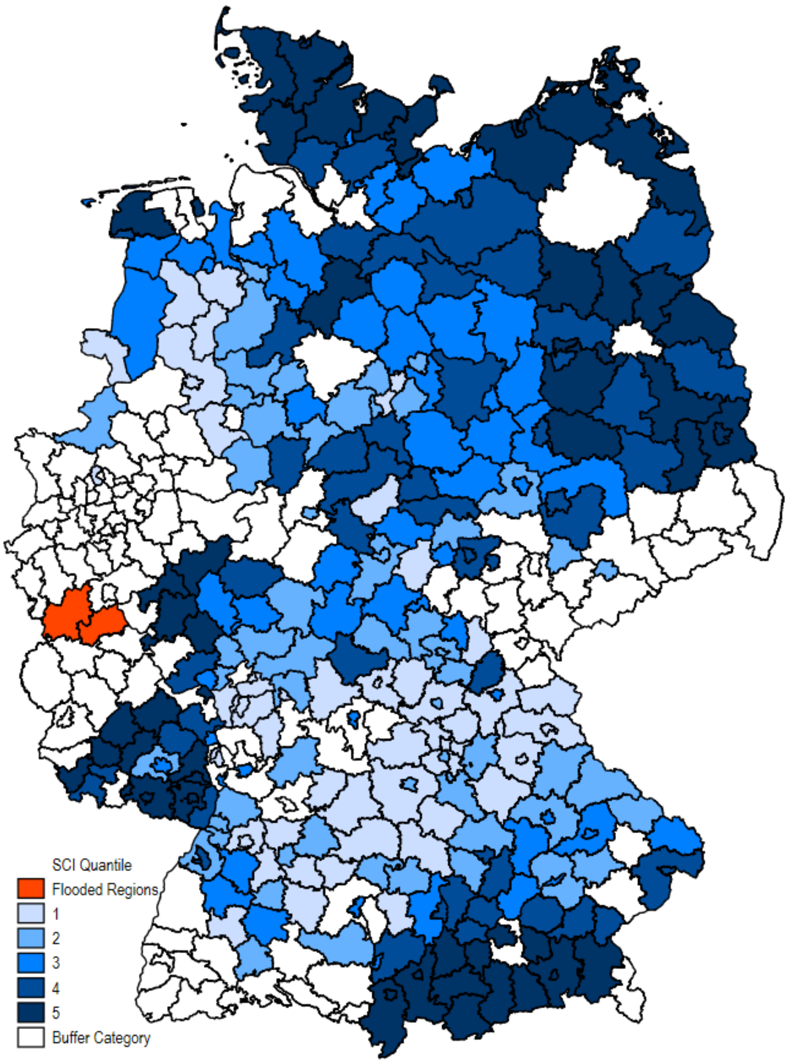


Figure A3.2: Alternative Sample Construction - 5%-Threshold

This figure illustrates the intensity of the flood in 2021 on NUTS-3 level. The map shows three groups of counties based on their claim ratio for the universe of policies against elemental damages in the German insurance industry (GDV, 2022). 281 regions exhibit a claim ratio of 0%, colored in blue. The buffer category consist of all regions that exhibit positive claim ratios below 5%, displayed in white. All NUTS-3-regions with claim ratios above 5% are defined as flooded and are illustrated in red.

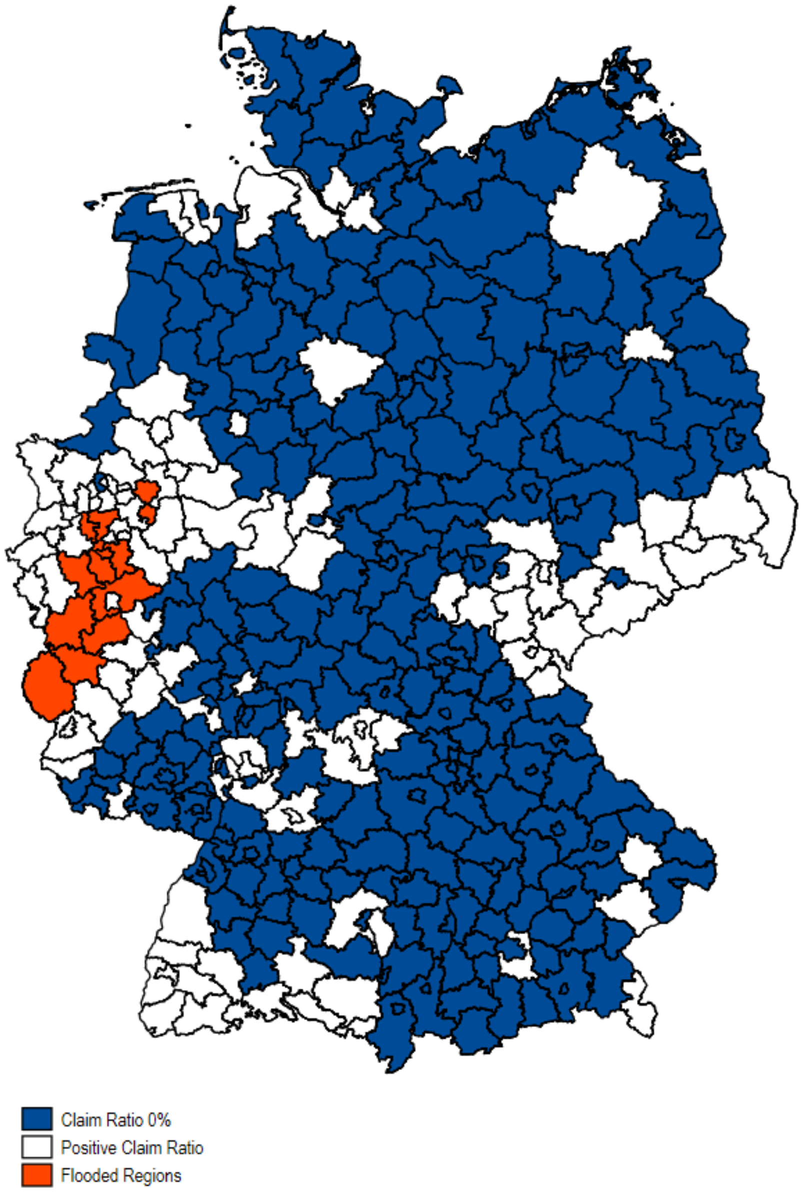


Figure A3.3: Flood-weighted Social Connectedness - 5%-Threshold

This figure illustrates the flood weighted social connectedness of all regions that did not reported any insurance claims regarding the flood in July 2021 to the flooded regions (colored in red) by quintiles. Other regions that exhibit a claim ratio below 5% serve as a buffer category and are illustrated in white. Darker shading represents a higher connectedness. Data on social connectedness is taken from Bailey et al. (2018) and has been introduced for Europe by Aref et al. (2020).

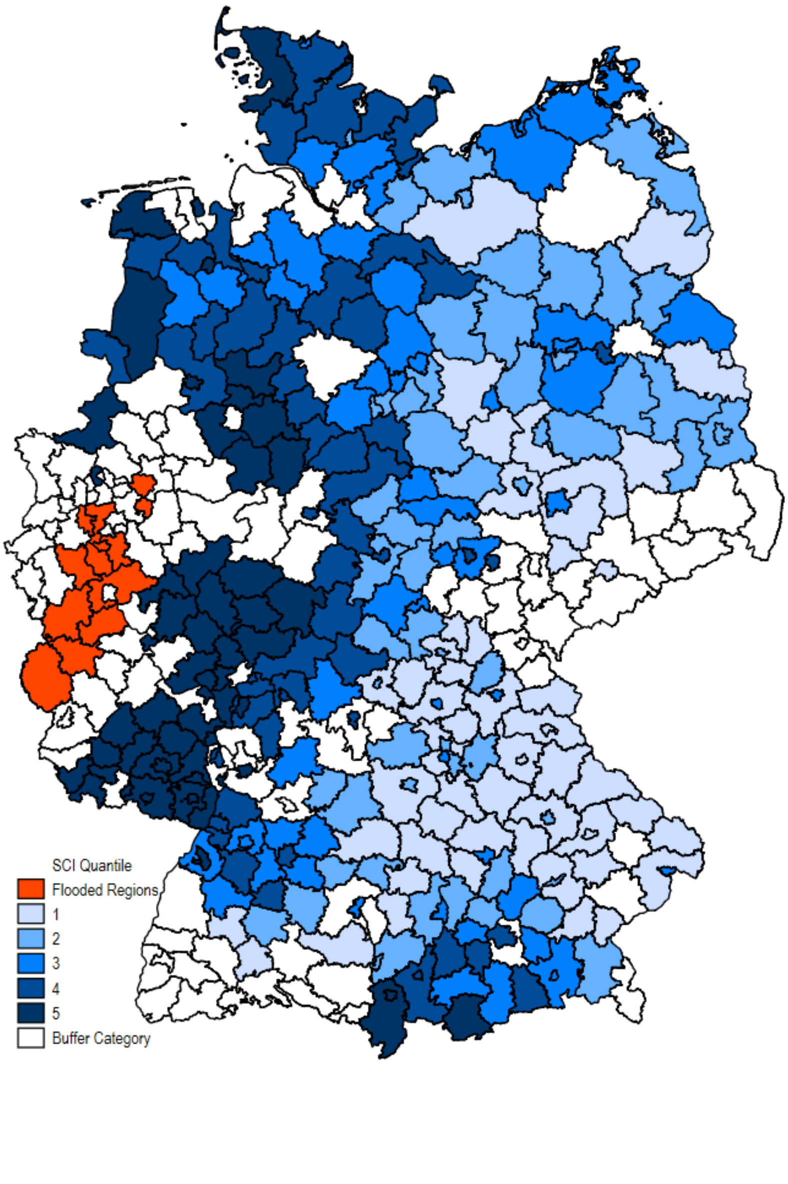


Figure A3.4: Alternative Sample Construction - 1%-Threshold

This figure illustrates the intensity of the flood in 2021 on NUTS-3 level. The map shows three groups of counties based on their claim ratio for the universe of policies against elemental damages in the German insurance industry (GDV, 2022). 281 regions exhibit a claim ratio of 0%, colored in blue. The buffer category consist of all regions that exhibit positive claim ratios below 1%, displayed in white. All NUTS-3-regions with claim ratios above 1% are defined as flooded and are illustrated in red.

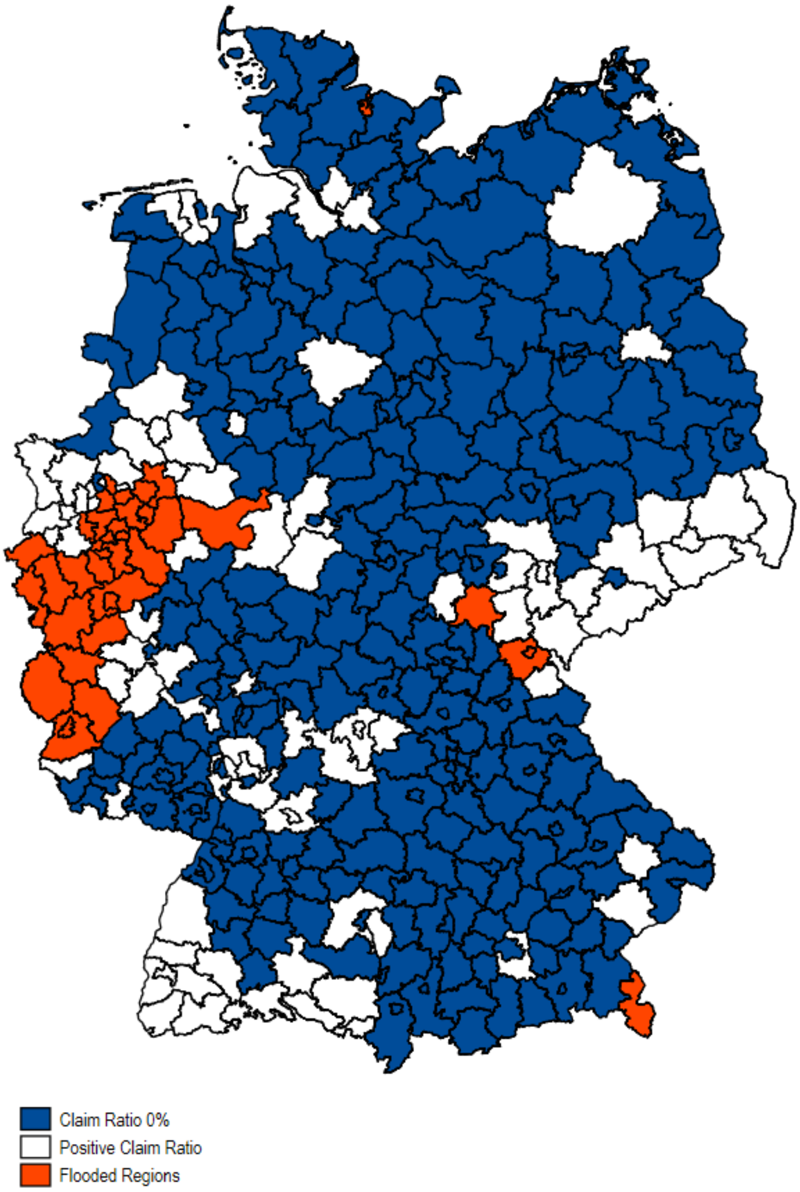
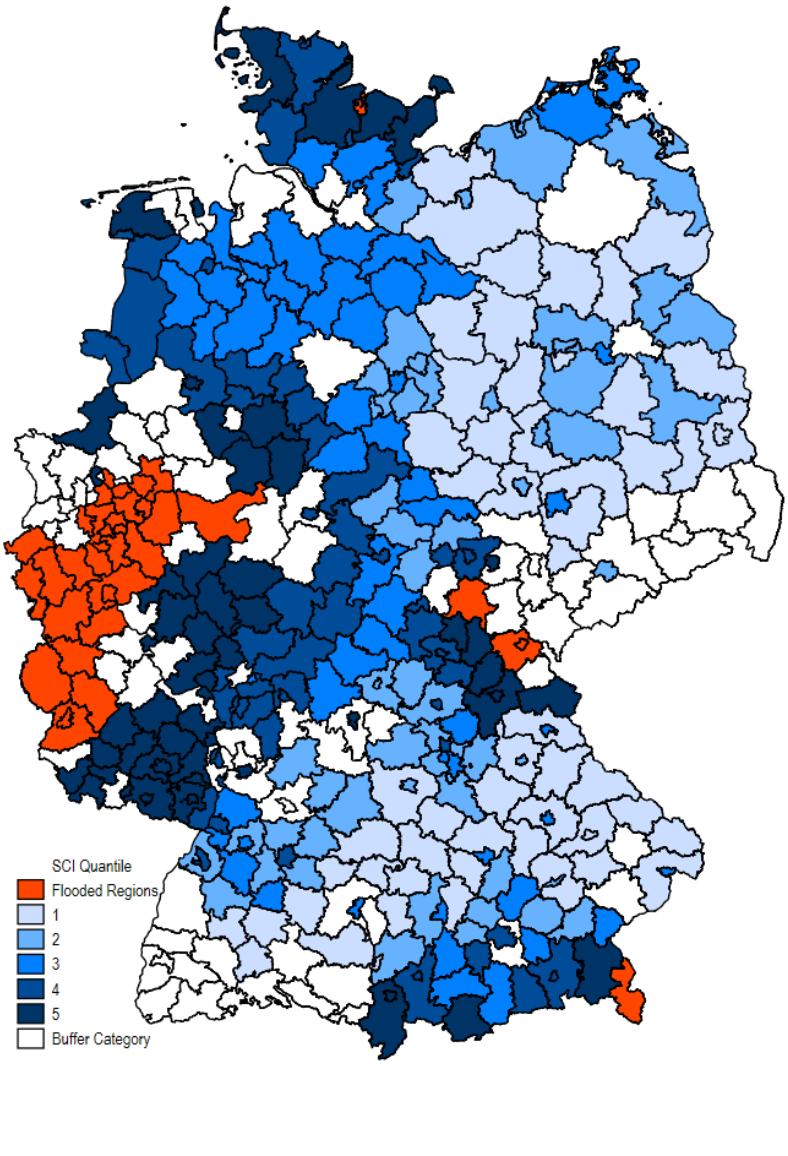


Figure A3.5: Flood-weighted Social Connectedness - 1%-Threshold

This figure illustrates the flood weighted social connectedness of all regions that did not reported any insurance claims regarding the flood in July 2021 to the flooded regions (colored in red) by quintiles. Other regions that exhibit a claim ratio below 1% serve as a buffer category and are illustrated in white. Darker shading represents a higher connectedness. Data on social connectedness is taken from Bailey et al. (2018) and has been introduced for Europe by Aref et al. (2020).



A3.2 Appendix Tables

Table A3.1: Variable Definitions and Data Sources

Variable Name	Description
Main Variables	
Log(Policies)	Log number of insurance policies that include elemental damage coverage in NUTS-3-region i in year-month t; source: DEVK Versicherungen
Log(Premiums)	Log value of premiums of all insurance policies that include elemental damage coverage in NUTS-3-region i in year-month t; source: DEVK Versicherungen
Log(Coverage)	Log value of the total coverate for all insurance policies that include elemental damage coverage in NUTS-3-region i in year-month t; source: DEVK Versicherungen
Social Connectedness Index	Relative probability of a user in NUTS-3-region i being friends on Facebook with a user in NUTS-3-region j, taking values between 1 and 1 billion, winsorized at the 99th percentile; source: Humanitarian Data Exchange
Weighted SCI	Pooled social connectedness between region i and flooded regions j, weighted by claim ratios as a proxy the local flood intensity; source: German Insurance Association (GDV, 2022)
Control Variables	
Weighted Distance	Population-weighted distance between NUTS-3-regions, weighted by claim ratios as a proxy the local flood intensity; source: TERCET Flat Files by the European Commission.
Claim Ratios	Share of contracts for which a loss has been reported for the flood in 2021. Date is taken from the German Insurance Association (GDV) and covers the entire German insurance industry.
PostFlood	Dummy that switches to 1 in July 2021, the month in which the flood occured
Fluvial Hazard	Percentage of the total area of the NUTS3 area that is prone to flooding in the event of a 1 in 100 year fluvial flood; source: (Hincks et al., 2023).
Population Share Exposed to Fluvial Flooding	Percentage of the total population of the NUTS3 area that lives in settlements that would be exposed to flooding in the event of a 1 in 100 year fluvial flood; source: (Hincks et al., 2023).

Table A3.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Number of Natural Hazard Events (2002-2021)	Number of events per NUTS-3-region that are categorized as major events by the German Insurance Association; source: German Insurance Association (GDV, 2022)
Average Local Damage Cost (2002-2021)	Average local cost of damages caused by major natural hazard events between 2002 and 2021 in EUR, according to the German Insurance Association; source: German Insurance Association (GDV, 2022)
Claim Ratio Heavy Rain Events (2002-2021)	Share of contracts for which a loss has been reported for heavy rain events per NUTS-3-region that are categorized as major events by the German Insurance Association; source: German Insurance Association (GDV, 2022)
Avg. Damage Heavy Rain Events (2002-2021)	Average local cost of damages caused by major heavy events between 2002 and 2021 in EUR, according to the German Insurance Association; source: German Insurance Association (GDV, 2022)
Number of Heavy Rain Events (2002-2021)	Number of heavy rain events per NUTS-3-region that are categorized as major events by the German Insurance Association; source: German Insurance Association (GDV, 2022)
Projected Change in Heavy Precipitation Days	Difference between the 1981-2010 period (observed baseline) and the 2036-2065 period (future projection) in the number of days with precipitation greater than or equal to 10mm; source: (Hincks et al., 2023).
Lenght of Rivers	Aggregated lenght of rivers in each NUTS-3-region; source: German Federal Institute of Hydrology.
Share of Area in Mountains	Share of area of a NUTS-3-regions that lies in the mountains; source: Bundesamt für Kartographie und Geodäsie.
Local GDP per Capita	GDP per Capita of NUTS-3-region i in 2021; source: Federal Office for Building and Regional Planning.
Purchasing Power	Purchasing Power in NUTS-3-region i in 2021; source: Federal Office for Building and Regional Planning.
Climate Policy Attitudes	

Table A3.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Green Party Regional Votes (pre 2021)	Vote share for the green party in the last regional election before the flood in NUTS-3-region i; source: multiple sources from official homepages of the Germany states, see for example Landesdatenbank NRW or Bayrisches Landesamt für Statstik for Bavaria.
Green Party State Votes (pre 2021)	Vote share for the green party in the last state election before the flood in NUTS-3-region i; source: multiple sources from official homepages of the Germany states, see for example Landesdatenbank NRW.
Willingness to Pay Higher Cost for Fossil Fuels	Share of people in NUTS-3-region i that is willing to pay higher cost for fussile fuels, based on the "Social Sustainability Barometer" (SNB) and the Ariadne Warmth and Housing Panel (WWP); source: Hertie School Centre for Sustainability.
Investment in Renewables Planned	Share of people in NUTS-3-region i that plans to invest in a wind power plant or a solar PV panel within the next two years, either alone, with friends or with a cooperative; source: Hertie School Centre for Sustainability.
Willingness to Protest Against Local Wind Power Plants	Share of people in NUTS-3-region i that can imagine participating in a collection of signatures, protest activity or civil society initiative against planned wind power plants in the surrounding area; source: Hertie School Centre for Sustainability.
Social Capital	
Voter Turnout (2017)	Voter turnout in the 2017 national election in NUTS-3-region i; source: Federal Office for Building and Regional Planning.
Covid Vaccinations per Capita June (2021)	Number of vaccinations per capita for the population between 18 and 59 years in June 2021 in county i; source: Bade et al. (2024).
SCI Dispersion	Average social connectedness of NUTS-3-region i to the own and all adjacent NUTS-3-regions j divided by the average social connectedness to all other counties k; source: own calculation using the Social Connectedness Index from Humanitarian Data Exchange.

Table A3.1: Variable Definitions and Data Sources (continued)

Variable Name	Description
Educational Equality	Absolute difference between the share of people with at least a high school degree in NUTS-3-region i and the share of people with less than a high school degree in NUTS-3-region i, multiplied by minus one, such that higher values can be interpreted as larger equality.; source: own calculation using the Social Connectedness Index from the Census 2022.

Table A3.2: Triple Interaction with Distance

This table shows the results from estimating regression 3.3 including an additional interaction term that controls for heterogeneity in the distance to the flooded areas. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × PostFlood	0.025***	0.004***	0.024***	0.012***	0.002	0.012***
	(0.005)	(0.001)	(0.005)	(0.004)	(0.001)	(0.005)
log(Weighted SCI) × log(Weighted Distance) × PostFlood	0.000	0.001**	0.001	0.001	0.000	0.001
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.985	0.984	0.984	0.991	0.989	0.990
Within-R^2	0.076	0.040	0.064	0.014	0.009	0.018
Observations	16,860	16,860	16,860	16,800	16,800	16,800

Table A3.3: SCI Purged by Distance

This table shows the results from regressing log social connectedness to the flood on log distance to the flood. Both variables are weighted by the flood intensity using claim rations. The Residuals from this regression are interacted with the PostFlood dummy in Tabel 3.5. Standard errors are clustered on NUTS-3-region level. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(Weighted SCI)
	(1)
log(Weighted Distance)	-0.760***
	(0.087)
R^2	0.554
Observations	16,860

Table A3.4: SCI and Distance (Purged)

This table shows the results from estimating Regression 3.3 with a purged version of the social connectedness to the flooded ares. The logarithmized SCI is regressed on the logarithmized distance to the flooded areas, results are presented in Table A3.3. Since both variables are time-invariant, this regression is estimated without any fixed effects, standard errors are clustered on NUTS-3-region level. The residuals from this regression enter Equation 3.2 interacted with *PostFlood*. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
Residuals × PostFlood	0.025***	0.006***	0.027***	0.013*	0.003*	0.017**
	(0.007)	(0.002)	(0.007)	(0.007)	(0.002)	(0.007)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.984	0.984	0.983	0.991	0.989	0.990
Within-R^2	0.035	0.035	0.039	0.011	0.009	0.017
Observations	16,860	16,860	16,860	16,800	16,800	16,800

Table A3.5: Unweighted SCI

This table shows the results from estimating regression 3.3. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Unweighted SCI) × PostFlood	0.024***	0.003***	0.021***	0.012**	0.002*	0.013***
	(0.004)	(0.001)	(0.004)	(0.005)	(0.001)	(0.005)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.985	0.983	0.984	0.991	0.989	0.990
Within-R^2	0.073	0.026	0.058	0.013	0.005	0.014
Observations	16,860	16,860	16,860	16,800	16,800	16,800

Table A3.6: Difference-in-difference with Control Group

This table shows the results from estimating regression 3.3 with a binary definition of treatment and control group. A treatment and control group is constructed by discretization of social connectedness. NUTS-3-regions with a social connectedness to the flooded areas above the median are defined as treated, those with values below the median are defined as being in the control group. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. All dependent variables enter logarithmized. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
PostFlood × Treated	0.020***	0.003***	0.020***	0.007**	0.001	0.008**
	(0.004)	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.984	0.984	0.984	0.714	0.617	0.685
Within-R^2	0.069	0.033	0.063	0.007	0.014	0.006
Observations	16,860	16,860	16,860	16,800	16,800	16,800

Table A3.7: Placebo Flood in 2019

The table shows the effect of SCI on the share of the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages based on a placebo flood. We assume that the flood took place in July 2019 and restrict the data set to the period from January 2019 to June 2021 in order to exclude the period after the actual flood. The placebo postflood dummy is set to 1 for all months after July 2019 and 0 for the period before. Social Connectedness is constructed by using the real flood. This approach tests, whether the SCI to the flooded regions has an influence on insurances even before the flood occurred. Regressions include county and year-month or state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × Placebo PostFlood 2019	0.015***	0.002**	0.015***	0.003	0.001	0.007
	(0.005)	(0.001)	(0.005)	(0.006)	(0.002)	(0.007)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes			
State × Year-Month FE				Yes	Yes	Yes
R^2	0.986	0.984	0.986	0.992	0.987	0.991
Within-R^2	0.030	0.011	0.029	0.001	0.002	0.004
Observations	15,174	15,174	15,174	15,120	15,120	15,120

Table A3.8: Extended Buffer Zone

This table shows the results from estimating regression 3.3 with an extended buffer zone. In Columns (1) to (3), the counties that have positive claim ratios below 15%, all counties in the states of North Rhine-Westphalia (NRW) and Rhineland-Palatinate (RLP) are excluded. In Column (4) to (6) the counties that have positive claim ratios below 15% and all counties within a 100km band of average distance to Euskirchen and Ahrweiler are excluded. This controls for effects caused by spatial proximity, but also effects triggered by local media, since I exclude all counties that are located in the same local media markets as the flooded counties Euskirchen and Ahrweiler. The dependent variables are the shares of policies including elemental damages, the premiums of insurance contracts including elemental damages and the coverage of insurance contracts including elemental damages. All dependent variables enter logarithmized. Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × PostFlood	0.022*	0.004	0.027**	0.017*	0.003	0.023**
	(0.012)	(0.003)	(0.012)	(0.010)	(0.003)	(0.010)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.986	0.981	0.985	0.986	0.980	0.984
Within-R^2	0.010	0.004	0.015	0.008	0.003	0.012
Observations	20,832	20,832	20,832	23,268	23,268	23,268

Table A3.9: Alternative Definition of Flooded Areas: Different Cut-offs

This table shows the results from estimating regression 3.3 for a different definition of the flooded areas. For Columns (1)-(3), all areas that exhibit a claim ratio above 5% are defined as flooded. For Columns (4)-(6), all areas that exhibit a claim ratio above 1% are defined as flooded. The dependent variables are the share of policies including elemental damages, the share of premiums of insurance contracts including elemental damages and the share of coverage of insurance contracts including elemental damages. Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × PostFlood	0.013**	0.001	0.014**	0.007	0.001	0.007
	(0.005)	(0.001)	(0.006)	(0.005)	(0.001)	(0.006)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.991	0.989	0.990	0.991	0.989	0.990
Observations	16,800	16,800	16,800	16,800	16,800	16,800

Table A3.10: Alternative Definition of Flooded Areas: Euskirchen or Ahrweiler

This table shows the results from estimating regression 3.3. For Columns (1)-(3) NUTS3-region Ahrweiler is considered to be the only flooded region, Euskirchen enters the buffer zone and is not taken into account when the social connectedness to the flooded region is calculated. For Columns (4)-(6) NUTS3-region Euskirchen is considered to be the only flooded region, Ahrweiler enters the buffer zone and is not taken into account when the social connectedness to the flooded region is calculated. The dependent variables are the share of policies including elemental damages, the share of premiums of insurance contracts including elemental damages and the share of coverage of insurance contracts including elemental damages. Regressions include county and state-year-month fixed effects, standard errors are clustered on NUTS-3-region level and depicted in parentheses. Descriptive statistics are presented in Table 3.1, detailed descriptions of all variables are presented in Table A3.1. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent Variable	Shares Including Elemental Damages					
	Policies	Premiums	Coverage	Policies	Premiums	Coverage
	(1)	(2)	(3)	(4)	(5)	(6)
log(Weighted SCI) × PostFlood	0.009**	0.001	0.010**	0.017***	0.003*	0.018***
	(0.004)	(0.001)	(0.004)	(0.006)	(0.001)	(0.006)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.991	0.989	0.990	0.991	0.989	0.990
Observations	16,800	16,800	16,800	16,800	16,800	16,800

Bibliography

Aref, S., Bontcheva, K., Braghieri, M., Dignum, F., Giannotti, F., Grisolia, F., and Pedreschi, D., editors (2020). Social Informatics: 12th International Conference, SocInfo 2020, Pisa, Italy, October 6–9, 2020, Proceedings, volume 12467 of Lecture Notes in Computer Science. Springer International Publishing, Cham.

Bade, V., Schmitz, H., and Tawiah, B. B. (2024). Regional variations in vaccination against COVID-19 in Germany. PLOS ONE, 19(4): e0296976.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. Journal of Economic Perspectives, 32(3): 259–280.

GDV (2022). Naturgefahrenreport 2022. Technical report, Gesamtverband der Deutschen Versicherungswirtschaft e. V.

Hincks, S., Carter, J., and Connelly, A. (2023). A new typology of climate change risk for European cities and regions: Principles and applications. Global Environmental Change, 83: 102767.