The London School of Economics and Political Science

Essays on Economics of Natural Disasters in China

Yuxiao Hu

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy. London, April 2025

I confirm that Chapter 1 was jointly-coauthored with Yifan Wang. I contributed 50% of the work in Chapter 1. I confirm that Chapter 2 was jointly-coauthored with Yifan Wang and Runhong Ma. I contributed 33% of the work in Chapter 2.

Abstract

This thesis investigates how flood management interventions and flood events shape the spatial distribution of economic development, firm adaptation, and innovation in China— one of the most flood-prone countries in the world. It examines both deliberate, policy-driven reallocations of flood risk and economic adaptations triggered by actual flooding. Chapters 1 and 2 analyze the impacts of China's national Flood Detention Basin (FDB) policy, which redirects floodwaters into designated rural areas to protect downstream urban centers. Chapter 1 uses reduced-form empirical methods to quantify the economic costs taken by FDB-designated counties, while Chapter 2 develops a spatial general equilibrium model to assess the broader economic benefits of the policy. Chapter 3 investigates the role of floods in shaping the geographical pattern of patenting activities. Together, the three chapters provide a comprehensive analysis of how both flood risks and flood management strategies influence the geography of economic activity and adaptive responses in China.

Chapter 1 examines the economic costs of China's Flood Detention Basin (FDB) policy, implemented in 2000. Under this national policy, the government designated 96 counties to host FDBs — low-lying areas intended to absorb excess floodwater during extreme weather events. While protecting downstream urban centers, this policy imposes concentrated flood risks on rural counties. Using difference-indifferences methods, the chapter documents significant economic costs for FDB counties: a 10.7% reduction in nighttime light intensity, a 15.9% decline in new firm entries, and a 19.7% drop in fixed asset investment.

These losses are persistent and primarily driven by firms ' aversion to locating in high-risk areas, rather than migration responses by individuals. Overall, using causal identification tools, this chapter shows that FDB policy has led to a substantial economic cost in counties selected to take more flood risks.

Chapter 2 builds a structural spatial general equilibrium model to quantify both the aggregate benefits of the FDB policy. The model captures trade linkages across FDB counties, protected cities, and the rest of the country. Counterfactual simulations reveal that the policy indeed enhances national output by protecting high-productivity urban centers. Overall, the benefit to cost ratio of the Flood Detention Basin policy exceeds one. However, these gains come at the expense of lower-productivity rural areas bearing theflood risk. The model shows that removing high-productivity counties from the FDB list would not substantially reduce output gains, while greatly improving equity. These findings suggest that the current FDB configuration may overprioritize output over equality. The results indicate the necessity for a more balanced compensation scheme to support vulnerable regions.

Chapter 3 explores how floods shape the spatial distribution of innovation in China. Using satellite-derived flood maps and detailed patent data, I created a dataset to measure collaborative patents among different regions. The chapter finds that floods decrease local patenting activity but simultaneously encourage cross-county collaboration in innovation. A one-day increase in average flood duration is related to a 12% increase in collaborative patents between counties. These partnerships are especially strong between counties that share similar flood histories and are more likely to yield disaster mitigation technologies. Mechanism analysis shows that historical flood experience—rather than unexpected shocks—drives the shift toward collaboration, suggesting a strategic adaptation to long-term climate risks.

Acknowledgements

I am deeply grateful for the guidance and support of my supervisors—Prof. Tim Besley, Prof. Robin Burgess, and Prof. Michael Callen—throughout my PhD journey. Prof. Besley has been an inspiring role model who showed me the importance of adhering to high standards. Prof. Burgess sparked my interest in environmental economics, and I am especially thankful that he encouraged me to study floods in China from the early stages of my work. Prof. Callen has consistently offered generous encouragement that not only improved my academic work but also helped me stay motivated. I thank them for preparing me for the exciting academic journey ahead. I also thank Prof. Oriana Bandiera and Prof. Clare Balboni for their detailed feedback and insightful discussions.

I would like to thank my coauthors, Yifan Wang and Runhong Ma. This journey would not have been nearly as rewarding without their invaluable contributions, intellectual insights, and unwavering support. I truly cherish the incredible synergy within our team. More importantly, we have been each other's strongest support during many challenging times. The deep mutual trust within the team gives me confidence that we will continue to work on research that we are proud of for many years to come.

I am also thankful to my friends at the Economics department—especially Sitong Ding, Yuhao Qian, Linchuan Xu, Yu Yi, Yuanhang Yu, and Wei Wang—for making my time there so memorable. Thank you for all the laughter, support, and unforgettable moments. A special thanks to my best friend outside the LSE, Yuchen Du, whose support and encouragement has meant so much to me since we met.

Above all, my family has been my greatest source of strength. I am endlessly thankful to my mother, who has always encouraged me to aim high and has never stopped believing in me. My deepest love goes to my partner Runhong. The love between us is the most precious gift the universe has given me, and I thank the universe everyday for how truly fortunate I am.

Lastly, I thank how wonderful and splendid the world can be.

Contents

List of Figures

List of Tables

Chapter 1

Government and Nature: Evidence from the Distribution of Flood Damages in China

With increasing disaster risks, it is increasingly important to understand the impact of government interventions that reallocate environmental damages. In 2000, the Chinese government designated 96 Flood Detention Basin (FDB) counties, allocating lower-elevation areas within these counties for temporary floodwater storage. During severe flood events, floodwater may be diverted to these FDB counties to protect downstream urban centers. We evaluate the aggregate and distributional impacts of the FDB policy. Difference-in-differences results show that if a county is selected to the FDB list, county-level firm entry and firm-level fixed asset investments would decrease by 15.9% and 19.7%, respectively. Overall, FDB designation results in a 10.7% reduction in county-level nighttime light intensity.

1.1 Introduction

A key challenge in natural disaster management is determining how to allocate environmental damages. Should government intentionally expose certain areas to higher risks to protect broader regions from severe damages? Similar to environmental policies that often create winners and losers (e.g., He et al. 2020, Taylor and Druckenmiller 2022), flood management interventions could have uneven distributional impacts. For instance, building dams and levees would lead to uneven effects across different regions (Duflo and Pande 2007, Bradt and Aldy 2023). Currently, floods impact more than 1.8 billion people globally (Tellman et al. 2021), and by 2050, severe flooding events are projected to double in frequency across 40% of the world (Arnell and Gosling 2016). As the threat of severe floods intensifies, it is increasingly important for policymakers in high-risk countries to understand the impact of flood management policies that lead to reallocation of flood damages.

This paper explores the aggregate and distributional impacts of Flood Detention Basins (FDBs), the last-resort solution in flood management. In extreme flood events, when reservoirs reach capacity, governments divert floodwaters into FDBs, which are regular lands during non-flooding periods, to protect a broader region from severe flood damages. A well-known example of an FDB is the Birds Point - New Madrid Floodway, located on the west bank of the Mississippi River in the United States.[1] In our paper, we focus on the world's largest FDB program: Flood Detention Basins in China, for two primary reasons. First, China ranks among the top three countries in terms of flood risk globally, with more than 395 million people exposed to floods. Floods also result in significant and persistent economic losses in China (Kocornik-Mina et al. 2020). Second, the FDB policy in China is a large-scale and explicitly designed national flood control policy, which has been in place for over two decades. Hence, we are able to clearly analyze its persistent impacts.

In 2000, the Chinese government officially implemented the Flood Detention

[1] The Birds Point-New Madrid Floodway in Missouri is engineered to divert up to 550,000 cubic feet per second from the Mississippi River during an extreme flood event. According to the US Army Corps of Engineers, "the purpose of the floodway is to lower flood stages upstream and adjacent to the floodway during major flood events."

Basin (FDB) policy, designating 98 low-lying wetlands in 96 counties as flood detention basins, covering over 30,000 km^2 and directly affecting more than 15 million residents. Over the past two decades, the government has used FDBs, which are mainly located in rural counties, to absorb floodwaters in nine different years. According to the Ministry of Water Resources in China, residents in FDB counties—counties designated for floodwater storage—make significant sacrifices to protect collective social welfare and improve economic resilience against floods.

In this paper, we ask the research question: What are the aggregate and distributional impacts of Flood Detention Basin policy in China? Regarding the distributional impact of the policy, we quantify the economic costs on counties where FDBs are located. It allows us to examine the extent to which rural FDB counties, which are initially more economically vulnerable, have made sacrifices to enhance overall economic resilience against floods. In terms of the aggregate impact of the policy, we evaluate whether the policy has resulted in a net gain in total output by extending our analysis to a general equilibrium context.

First, we find that the FDB policy has effectively redistributed flood exposures across different regions. Using the Global Flood Database (Tellman et al. 2021), a satellite-based flood dataset, we construct proxies to measure county-level flood exposures. Through a fixed-effect regression, we find that the size of flood inundation in FDB counties is over 50% larger compared to other counties, after controlling for key geographical attributes. Additionally, we use a hydro-dynamic engineering model to simulate a counterfactual scenario without FDBs absorbing excess floodwaters. In one case study, we find that an economically important city, Wuhan, would experience 45% more flooding during a severe flood event.

Second, we find that the FDB designation has had a negative and persistent impact on the economic development of FDB counties. Using a difference-in-differences approach, we compare the economic development of FDB counties with that of counties not affected by the FDB policy. However, FDBs are not randomly distributed. According to the Chinese government, FDBs should be low-lying areas that are hydrologically feasible for absorbing floodwater. To address this selection issue, we employ the synthetic difference-in-differences estimation method proposed by Arkhangelsky et al. (2021) for the main analysis, while also providing traditional and alternative DID estimations for robustness (Callaway and Sant'Anna 2021, de Chaisemartin and D'Haultfœuille 2020, Gardner 2022). Overall, we find a negative and statistically significant impact of FDB designation on economic development: the FDB designation reduces annual nighttime light intensity by approximately 10%. Henderson et al. (2012) and Martinez (2022) estimate the elasticity of GDP to nighttime light at around 0.3. Hence, we are able to translate the reduction in light intensity to an approximate 3% annual GDP loss. This cost estimation is also consistent with findings from hydrologists (e.g., Wang et al. 2021). Event studies using a 20-year window centered around the year of FDB designation further support the validity of our identification methods.

To investigate the mechanism behind the reduction in nighttime light intensity, we examine the impact of the 2010 policy change in which the Chinese government added 20 counties to the FDB list and removed 10 counties from it. This policy change allows us to compare the treatment effect of being selected into the FDB list and that of being removed from the list. We first examine whether people make location decisions in response to the FDB policy. However, unlike previous studies that provide evidence of migration following floods (e.g., Hornbeck and Naidu 2014), our findings do not find evidence of migration in response to this policy, possibly due to the mobility restriction in China. Instead, our findings suggest that the firm-response effect is the major mechanism, as firms are reluctant to enter and invest in FDB counties with higher flood risks. Our empirical analysis supports the firm-response mechanism as follows:

(i) On average, the number of new firm entries has declined by 15.9% in the newly designated FDB counties after the 2010 policy change. Focusing on larger manufacturing firms with a turnover above \$3 million, the number of such firms has declined by 21.7% in newly designated FDB counties following the policy change. This result is also consistent with Jia et al. (2022) and Balboni et al. (2023), which find that firms make location decisions in response to flood risk change;

(ii) Using detailed firm investment data, we apply a spatial regression discontinuity approach (Imbens and Wager 2019 and He et al. 2020) to compare firm investments in FDB counties versus neighboring counties. We find that investment in fixed assets is 19.7% lower in FDB counties compared to neighboring counties, with this gap in fixed asset investment only emerging after the 2010 policy change;

(iii) In contrast, we find significant evidence that firm entries and firm investments have increased in counties that were removed from the FDB list in 2010. Specifically, the number of new firm entries has increased by 16.8%, and investments in fixed assets have increased by 25.7%. Compared to the treatment of being selected into the list, we view the balanced and symmetrical effect of being removed from the list as compelling evidence that the FDB policy significantly influences firms' decision-making.

This paper makes three key contributions. First, we contribute to the discussion on flood costs by illustrating that flood management policies, while aimed at reducing damage, can also lead to significant economic costs. We find that governments have incentives to mitigate floods by shifting flood damages onto regions of lower economic values. Kocornik-Mina et al. (2020) finds that while urban areas experience frequent flooding, lower-elevation cities tend to recover as quickly as higher-elevation cities. Our paper helps explain this phenomenon by suggesting that governments may strategically channel flood damages to rural areas. Also, our findings are consistent with prior studies that report negative impacts of floods on economic development in both developing (e.g., Patel 2023) and developed countries (e.g., Strobl 2011). A cross-country study by Hsiang and Jina (2014) also demonstrates the causal effect of cyclones on long-term economic growth across various regions, while Desmet et al. (2021) predicts that permanent flooding due to climate change could reduce global real GDP by 0.19 percent. For studies focusing on China, Elliott et al. (2015) identifies that typhoons impose a significant but short-lived negative impact on local economic activity in China. Our study contributes to this relatively limited literature on the economic costs of floods in China—a country with severe flood risk, where approximately 395 million people are exposed to floods.

Second, our study contributes to the literature on how individuals and firms adapt to both natural disasters and government interventions. We find that firms adjust their entry and investment decisions in response to changes in flood risk. Balboni et al. (2023)

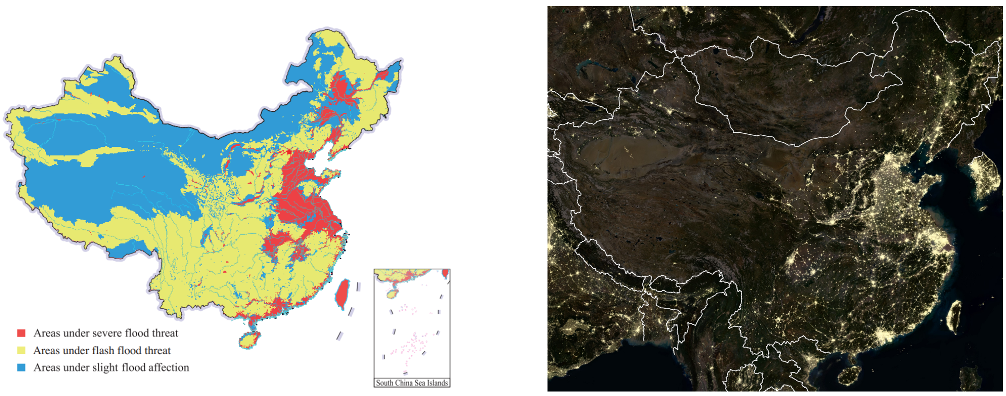
observes a similar trend, with firms in Pakistan relocating from flood-affected areas to less flood-prone regions. Similarly, Jia et al. (2022) also finds that flood risk will affect firm location decisions in the United States. Our study further expands the discussion by examining how firms adapt to government interventions. We find that environmental damages tend to be disproportionately concentrated in economically less valuable areas. Meanwhile, economic activity becomes more concentrated in urban centers. This aligns with recent findings by Hsiao (2023) that government interventions may create moral hazard, encouraging greater economic concentration in coastal regions. In terms of individual response, we find no evidence of migration in reaction to the policy, which is different from previous studies (e.g., Boustan et al. 2012; Hornbeck and Naidu 2014; Gröger and Zylberberg 2016; Boustan et al. 2020). Understanding the underlying reasons will be an important area for future research.

Third, our research contributes to the discussion on the aggregate and distributional impacts of environmental policies and government interventions. Environmental policies often have distributional impacts. He et al. (2020) shows that firms located upstream of pollutant monitoring stations in China experience larger reductions in productivity than downstream firms. Similarly, Taylor and Druckenmiller (2022) finds spatial heterogeneity in benefits from the Clean Water Act in the United States. With climate change intensifying, the allocation of environmental damages becomes an increasingly important topic. For instance, Duflo and Pande (2007) finds that residents upstream of dams in India face greater constraints in economic mobility than those downstream. For example, Balboni (2019) examines the spatial distribution of large infrastructure investments in Vietnam, a country highly threatened by sea-level rise. Hsiao (2023) also assesses the spatial distributional impacts of constructing sea walls. We contribute to this strand of literature by incorporating Flood Detention Basins (FDBs) into the discussion. Consistent with prior findings, we observe substantial distributional impacts of these policies. Meanwhile, while much of the literature on flood management policy has focused on flood insurance programs in the United States (e.g., Gallagher 2014, Mulder 2021, Georgic and Klaiber 2022), our study extends this discussion by investigating the impact of FDB policy that intentionally reallocates flood damages.

1.2 Research Background

1.2.1 Substantial Flood Risk in China

China ranks among the top countries globally for flood risk, due to its large population exposed to both coastal and river flooding. According to the Aqueduct Global Flood Risk Country Rankings, China ranks third in the world for the absolute number of people exposed to flood risks, with approximately 395 million people at risk annually. This places China among the most flood-exposed countries, alongside India and Bangladesh. About 27.5% of China's population is vulnerable to flooding, driven by river floods in the Yangtze, Huai, and Yellow River basins, as well as coastal areas prone to typhoons and rising sea levels. From 2000 to 2017, floods caused economic damage exceeding $150 billion, according to the EM-DAT International Disaster Database. Furthermore, Arnell and Gosling (2016) predicts that the likelihood of a 100-year flood occurring in China could increase by 33-67% by 2050.



(a) Flood Risk Distribution in China **(b)** Nighttime Light in China

Figure 1.1: Richer regions in China face higher river flood risk.

A key feature of China's floods is their disproportionate impact on economically important regions. Jiangsu Province, for instance, ranks second in GDP among China's provinces, yet faces severe flood risks due to its location along the Yangtze River and Huai River. As shown in Figure 1.1, regions with higher flood risks, identified by Zhang and Song (2014), are also more economically significant, as indicated by higher nighttime light intensity. For instance, the Yangtze River Basin, home to one-third of China's population, is a crucial economic hub. Frequent flooding, exacerbated by seasonal rainfall and extreme weather events, poses significant risks to infrastructure and livelihoods in these areas. Similarly, the Huai River Basin, another key region, faces recurring flood threats. Flooding in these economically vital regions could hinder China's overall economic growth, making flood management a critical concern for the government.

Due to rapid urbanization, urban populations in major cities (e.g., Beijing, Wuhan, and Nanjing) are increasingly exposed to severe flood risks. The urbanization rate surged to 64.72% in 2021, up from 36.00% in 2000, which has significantly heightened the vulnerability of urban areas to flooding. For instance, the 2012 Beijing flood, triggered by extreme rainfall, resulted in over 79 fatalities and caused approximately $2 billion in economic damage. The 2021 Zhengzhou flood led to over 350 deaths and caused around $6 billion in economic losses. This underscores the severe impact of urban flooding on densely populated areas.

1.2.2 Flood Detention Basins: the Last Resort of Managing Floods

Flood Detention Basins (FDBs) are areas designated for the temporary storage of floodwater to protect broader regions from flood damage. FDBs are an essential component of the flood management strategy, particularly when other measures are insufficient to mitigate severe flood impacts. A famous example of such an approach is the Birds Point-New Madrid Floodway in Missouri, USA, where controlled flooding mitigates the risk of severe damage to surrounding communities. The Birds Point-New Madrid Floodway, established in 1928 after the Great Mississippi Flood, spans approximately 130,000 acres and is part of the Mississippi River and Tributaries Project. During times of extreme flooding, levees are intentionally breached to divert water away from populated areas, thereby reducing flood risks to downstream communities such as Cairo, Illinois. This floodway has been activated multiple times, most recently in 2011, to protect both urban and rural areas from catastrophic flood damage.

The *Flood Control Law of the People's Republic of China*, implemented in 2000,

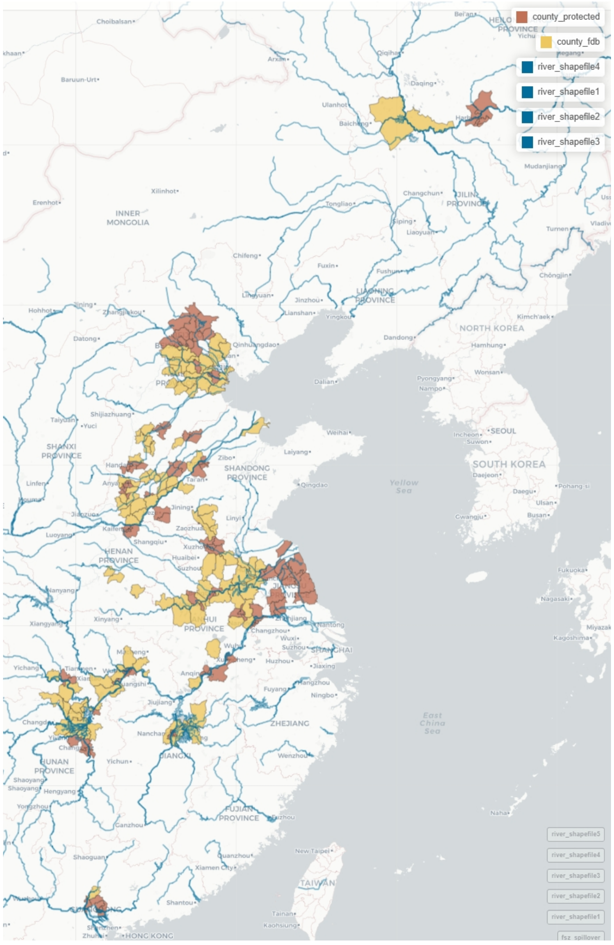


Figure 1.2: FDB Counties and FDB-Protected Counties
Note: (1) FDB counties are marked using color yellow, and FDB-protected counties are marked using color red; (2) FDB counties are located near the river, and FDB-protected counties are located to the downstream of FDB counties.

is the country's first piece of legislation specifically governing flood management. This law officially designates certain areas as Flood Detention Basins (FDBs). According to the law, FDBs are low-lying lands and lakes used for the temporary storage of floodwaters. To facilitate floodwater diversion, the Chinese government constructs dams and dikes in these FDB counties, enabling effective flood management during extreme events. The law specifies that the purpose of establishing FDBs is to "safeguard the interests of pivotal regions and the whole watershed." Additionally, the government acknowledges that residents in these FDBs make significant sacrifices for the greater collective welfare. As shown in Table A1, the FDB policy directly affects about 1.1%

of China's total population. The aggregate area of FDBs is 30,443 km^2 (0.3% of China's total land), which is comparable with the entire territory of Switzerland.

FDB counties protect downstream urban areas from severe flood impacts and play a key role in diverting floodwaters to protect downstream areas. As illustrated in Figure 1.2, FDB counties are located in the upstream so that urban cities to the downstream could be protected from being severely damaged during floods. For example, the Mengwa Flood Detention Basin, located in Funan County, Anhui Province, has been activated more than 16 times since its establishment. During flood detention, more than 200,000 residents in the Mengwa Flood Detention Basin are temporarily relocated to neighboring counties. More details about this case study can be found in Appendix A.1.2.

Policy Change

According to the 2000 Flood Control Law, 96 counties were designated as FDB counties, the first time that the specific locations of these basins for flood detention were officially confirmed. In 2010, the Ministry of Water Resources revised the earlier law in the *National Flood Detention Basin Construction and Management Plan*. As indicated in Table A2, under this new plan 13 FDBs were added and 12 were removed. Consequently, the specific counties classified as FDB counties changed, with 20 new additions to the list and 10 removed. Table A1 and Table A2 offer an overview of the FDBs in China's major river basins in 2000 and 2010.

1.2.3 Key Features of Flood Detention Basins in China

Selection

According to national law, detention basins are typically placed in topographically low areas conducive to floodwater containment, as these areas naturally accumulate water, making them ideal for mitigating flood impacts. The selection of FDB counties is determined by the Ministry of Water Resources, indicating a centralized decision-making process. Among all factors, hydrological feasibility for absorbing floodwater is the most critical determinant in this decision-making process. Key considerations include soil permeability, water retention capacity, and the ability to minimize adverse downstream effects. Research in hydrology has consistently emphasized the importance of these factors in optimizing FDB selection.[2] In Table A4, we also present a linear proabability regression model to identify factors influencing the selection of Flood Detention Basin locations. Our findings indicate that the choice of FDB sites is predominantly influenced by hydrological and geographical characteristics. This is consistent with the official stance of the Chinese government, which defines FDBs as "low-lying lands and lakes that are hydrologically suitable for temporary storage of floodwaters."

Migration

During the flood detention period, residents are temporarily relocated to neighboring counties or 'zhuangtai'—areas with higher elevation that remain unflooded during floodwater diversion. Unlike the case of building reservoirs, the government does not force residents in FDB counties to leave. Although the government may encourage local residents to relocate, the financial incentives provided are insufficient. According to a survey conducted in a Flood Detention Basin, 73% of local residents are dissatisfied with the current migration incentive scheme, and 94% are dissatisfied with the migration destinations offered by the government. Overall, 69% of participating residents are unwilling to leave the FDB county. We present more empirical findings about migration in Section 1.6.1.

Compensation

Subsidies to FDB counties during normal periods, when there is no floodwater diversion, are limited. However, according to the *Temporary Measures for the Use of*

---

[2] Mays and Bedient (1982) developed an optimal model based on dynamic programming, aiming to determine the ideal size and location of detention basins to maximize flood absorption while minimizing construction costs. This model was further refined by Bennett and Mays (1985) by incorporating the cost implications of detention basin structures and downstream channel designs. Using this refined model, Taur et al. (1987) optimized the detention basin system in Austin, Texas, highlighting the significance of hydrological suitability in site selection. Mays and Bedient (1982) advanced this research by optimizing the placement and sizing of retention basins in a watershed, specifically targeting reductions in aggregated costs related to construction, maintenance, and sediment removal, while considering hydrological efficiency.

Compensation in Flood Storage and Detention Areas initiated by the Chinese government in 2000, the government is supposed to compensate up to 70% of damages caused by direct floodwater diversion. The specific compensation standards are determined by the provincial-level government and are based on the actual damage caused by the flood within these parameters. However, the requirements for receiving compensation are not clearly specified. For example, the government will not compensate for livelihood losses if assets could have been relocated according to government orders but were not. But there is no clear specification regarding how to assess whether the asset could or could not have been transferred.

Due to increasing flood risk, the Chinese government has emphasized using financial tools to help alleviate flood risk. For instance, in 2024, the People's Bank of China allocated an additional $15 billion in relending funds for agricultural and small business support in 12 provinces (regions, municipalities). These funds are intended to support flood prevention, disaster relief, and post-disaster reconstruction efforts in severely affected areas. However, during the period of this research, this type of subsidy or compensation remains limited. A more detailed discussion on compensation can be found in Appendix A.1.4, along with an example of actual compensation from the 2023 floodwater diversion in Zhuozhou County, Hebei Province.

1.3 Data and Empirical Strategies

1.3.1 Data

FDB List - The Ministry of Water Resources officially announced the list of Flood Detention Basins (FDB) in 2000, and revised the list in 2010. We then define counties that hold flood detention basins as FDB counties. The original policy document can be found in Appendix A.1.1.

Data on Light - Given possible threats to GDP estimation in datasets provided by the National Bureau of Statistics (NBS), as suggested by Martinez (2022), we use nighttime light data as a proxy of economic activity. Specifically, we use the 1984-2020 'Prolonged Artificial Nighttime-light Dataset of China' data by Zhang et al. (2024).

Data on Firm-level Outcomes - Firm-level data is collected from National Enterprise Credit Information Publicity System (NECIPS) and Annual Survey of Industrial Enterprises (ASIE). NECIPS, administered by China's State Administration for Market Regulation (SAMR), provides annual registration records for all Chinese enterprises spanning from 1960 to 2023. This dataset is rich in detail, encompassing key information such as the date of establishment, ownership type, and geographical location of each firm. Using the geo-located data within this resource, we are able to accurately track the entry of firms in counties and towns designated as Flood Detention Basins (FDB). The firm-level data derived from ASIE spans from 1998 to 2014. ASIE encompasses private industrial enterprises with annual sales exceeding 5 million RMB (approximately 0.7 million USD) and all state-owned industrial enterprises (SOEs). Compiled and maintained by the National Bureau of Statistics (NBS), this dataset offers an extensive array of information sourced from the accounting records of these firms. It includes data on inputs, outputs, sales, taxes, and profits. This dataset contrasts with the National Enterprise Credit Information Publicity System (NECIPS) in two key aspects. Firstly, ASIE's temporal scope is confined to the period between 1998 and 2014, whereas NECIPS provides a wider temporal range for analysis (1960 to 2023). Secondly, ASIE primarily concentrates on collecting comprehensive details about firm activities, whereas NECIPS is oriented towards the registration of new firms.

Data on Other Socio-economic Outcomes - Other county level data is collected from the County-level Statistical Annual Yearbooks from 1999 to 2022. The National Bureau of Statistics (NBS) conducts county-level survey each year. It is a longitudinal survey that collects county-level socio-economic data for all counties in China. County-level variables include local output (disaggregated by sector), number of firms, fiscal income, fiscal expenditure, savings and etc.

Geographical Data - Elevation and gradient information is obtained from the NASA ASTER Global Digital Elevation Model (GDEM). The GDEM, with its extensive coverage from 83 degrees north to 83 degrees south latitude, encompasses 99 percent of the Earth's landmass. This comprehensive database enabled us to gather detailed elevation and gradient data for all counties and towns across China. For precipitation data, we turned to the Global Surface Summary of the Day (GSOD), sourced from the Integrated Surface Hourly (ISH) dataset. GSOD provides daily summaries typically within 1-2 days of the observation date. It encompasses data from over 9,000 stations worldwide, offering historical records from 1929 onwards, with the period from 1973 to the present being the most complete. Utilizing this resource, we calculated the mean monthly precipitation for each village and town in China.

1.3.2 Descriptive Statistics

In Table A3, we compare several descriptive statistics of FDB counties and non-FDB counties. FDB counties, compared to non-FDB counties, exhibit differences in geographical, flood, and socio-economic characteristics. Geographically, FDB counties have lower elevations and slopes but more permanent water pixels. This is consistent with the government claim that flood detention basins are typically low-lying lands and lakes used for temporary storage of floods. In descriptive results, we find that FDB counties experience higher flood exposure and larger areas of flood inundation. Contrary to the claim that FDB counties should hold less population and be poorer, the data demonstrates that FDB counties actually have larger populations and higher nighttime light intensity, which is often an indicator of greater economic activity. Additionally, FDB counties have a slightly greater number of firms compared to non-FDB counties.

These socio-economic indicators suggest that FDB counties are not poorer; rather, they have significant economic activities. This evidence contradicts the assumption that FDB counties are less populated and economically disadvantaged.

1.3.3 Empirical Strategies

Identification Challenge: FDB Location Choice

From a geographical perspective, detention basins are typically placed in topographically low areas conducive to floodwater containment. The field of hydrology has provided a wealth of research on optimizing the selection of flood detention basins. Mays and Bedient (1982) developed an optimal model based on dynamic programming, aiming to determine the ideal size and location of detention basins, with the goal of minimizing system construction expenditures. This model was further refined by Bennett and Mays (1985) by incorporating the cost implications of detention basin structures and downstream channel designs. Utilizing this evolved model, Taur et al. (1987) optimized the detention basin system in Austin, Texas. Travis and Mays and Bedient (1982) advanced this line of research by optimizing the placement and sizing of retention basins in a watershed, targeting the reduction of aggregated costs encompassing construction, maintenance, and sediment removal. Subsequent studies have integrated various optimization techniques, such as genetic algorithms and simulated annealing, and incorporated detailed engineering cost assessments into the design frameworks for detention basin-river-protected region systems (e.g., Perez-Pedini et al. 2005; Park et al. 2014).

However, potentially non-random FDB location choice remains the major challenge in identifying the effects of the Flood Detention Basin (FDB) policy. The selection or removal of counties from the FDB list is likely influenced by factors other than geographical factors. For instance, the government may designate less economically developed counties to host those basins, or conversely, remove a county from the FDB list due to its better economic performance.

In Table A4, we apply a logit regression model to identify the determinants influencing the selection of Flood Detention Basins (FDB) locations. Our findings suggest that the choice of FDB sites is predominantly influenced by geographical characteristics. This aligns with the official stance of the Chinese government, which defines FDBs as 'low-lying lands and lakes situated beyond the back scarps of dikes, inclusive of flood diversion outfalls, utilized for the temporary storage of floodwaters.' Our analysis corroborates this definition, revealing a significant tendency for counties with lower elevation levels to be selected as FDBs. We do not find empirical evidence to claim that the Chinese government intentionally selected relatively poorer counties as FDBs.

Two-Way-Fixed-Effects (TWFE) Difference-In-Differences

Our logit regression results, as shown in Table A4, reveal no significant correlation between a county's FDB status and its GDP, which suggests that FDB policy implementation may not be directly related to economic output. However, this does not entirely rule out the possibility that socioeconomic factors influence FDB selection decisions.To address the endogeneity concern, we use three identification strategies: traditional TWFE Difference-in-Differences, the Synthetic Difference-In-Differences (SDID) and spatial regression discontinuity (SRD).

We first use the most traditional Two-Way-Fixed-Effects (TWFE) Difference-In-Differences approach to investigate the imapct of FDB policy. The regression specification takes the form of:

$$ln(Y)_{it} = \alpha + \beta_1 FDB_{it} + \gamma_i + \lambda_t + \varepsilon_i$$

where Y_{it} measures the outcome of interest of county i in year t, FDB_{it} is a dummy variable that equals 1 if the county i is an FDB county in year t, and 0 if not. γ_i, and λ_t indicate county and year fixed effects, respectively. Standard errors are clustered at the county level. In this regression specification, β_1 is the difference-in-difference estimate that measures the impact of FDB policy on outcomes of interests.

Synthetic Difference-In-Differences (SDID)

Considering recent discussions on the properties of the staggered Difference-in-Differences (DID) approach, particularly regarding potential biases stemming from the weighting problem as highlighted by Borusyak et al. (2024), we argue that the Synthetic Difference-in-Differences (SDID) method, proposed by Arkhangelsky et al. (2021). Central to the SDID framework is its ability to derive a counterfactual for each treated entity by computing a weighted average from a comprehensive set of potential controls. We argue that SDID is well-suited for our empirical setting for several reasons.

First, constructing a counterfactual group using synthetic weights, as proposed by Abadie et al. (2010), effectively addresses concerns about the weighting problem inherent in traditional TWFE DID. SDID ensures that the synthetic control group closely mirrors the treatment group's pre-treatment characteristics, thereby enhancing the validity of causal inferences.

Second, Roth et al. (2023) suggest that clustering at the unit level is inappropriate when the number of treated groups is small. In our context, the 2010 policy change by the Chinese government, which added 20 new counties to the list and removed 10, involves a limited number of treated clusters. Given this small sample size, employing bootstrap standard errors, as facilitated by the SDID approach, provides a more reliable measure.

Third, the construction of synthetic weights mitigates potential threats to exogeneity by ensuring that the counterfactual group exhibits pre-treatment outcomes that are parallel to those of the treatment group. This parallel trend assumption is crucial for the validity of DID estimates, and the SDID method's ability to create a closely matched synthetic control group strengthens this assumption.

In summary, the SDID approach offers a robust solution to the potential biases associated with traditional DID methods, particularly in settings with small numbers of treated units and concerns about weighting and exogeneity. This makes it a particularly suitable choice for our analysis of the economic impacts of the 2010 policy change in China. Following Arkhangelsky et al. (2021), the average treatment effect on the treated, or ATT, is denoted as τ. Estimation of the ATT proceeds as follows:

$$\left(\widehat{\tau}^{sdid}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \widehat{\omega}_i^{sdid} \widehat{\lambda}_t^{sdid} \right\}$$

weights $\hat{\omega}_i^{\text{sdid}}$ and $\widehat{\lambda}_t^{\text{sdid}}$ are optimally chosen given the design by Arkhangelsky et al. (2021). Time fixed effects are denoted by β_t and unit fixed effects are denoted by α_i. Y_{it} is the outcome of a county i at year t. W_{it} is the treatment dummy that equals 1 if county i is treated in year t, and 0 if not. μ is the constant term.

Spatial Regression Discontinuity (SRD)

We also employ a spatial regression discontinuity design based on a firm-level dataset, the Annual Survey of Industrial Enterprises (ASIE). Both parametric and non-parametric methods can estimate the discontinuity. Imbens and Wager (2019) demonstrated that the parametric RD method, employing a polynomial function of the running variable as a regression control, often produces RD estimates sensitive to the polynomial's degree and exhibits several other unfavorable statistical characteristics. Consequently, we adopt the advised local linear method and proceed to estimate the equation below.:

$$\begin{aligned} Y_{ij} = \alpha_1 \text{ FDB }_{ij} + \alpha_2 \text{ Dist }_{ij} + \alpha_3 \text{ FDB }_{ij} \cdot \text{ Dist }_{ij} \\ + \varepsilon_{ij} \quad \text{ s.t. } -h \leqslant \text{ Dist }_{ij} \leqslant h, \end{aligned}$$

where Y_{ij} is the assets per worker of firm i in county j. FDB $_{ij}$ is an indicator variable that equals 1 if firm i is treated by policy shock (in the new FDB region or in the newly abolished FDB region), and 0 otherwise. Dist $_{ij}$ measures the distance between firm i and new FDB county border (or abolished FDB county border) j (negative if outside the county and positive within the county), and h is the estimated MSE-optimal bandwidth following Calonico, Cattaneo, and Farrell (2018). The standard error is clustered at the county level to deal with the potential spatial correlation of the error term, as suggested by Cameron and Miller (2015).

1.3.4 Counties as the Unit of Analysis

In this study, we concentrate on the county level rather than the town level within China's administrative hierarchy. Counties, situated between prefectures and townships, form the third tier of the administrative structure. Mainland China comprises 2,851 county-level divisions. According to 2000 and 2010 FDB policy, in total, 96 and 106 counties could be identified as a FDB county, respectively. We focus on counties for two reasons. First, county-level data is more comprehensive. The National Bureau of Statistics (NBS) provides the most extensive collection of socioeconomic variables at the county level. By focusing our analysis here, we can more effectively examine the impact of policies on crucial socioeconomic indicators, such as the output of various sectors. Second, flood detention typically will impact most towns in a county. Although dams are situated in towns, we observed that in the event of a flood, the impact typically extends to encompass the entire county.

1.4 FDB Policy and Flood Redistribution

Before analyzing the economic impacts of the Flood Detention Basin (FDB) policy, this section presents the first-stage results on whether the policy has successfully redistributed floodwaters. Using fixed effects regression and hydrological dynamic model, we aim to quantify the extent to which FDB counties have absorbed excess floodwaters due to the policy.

1.4.1 Measuring Floods

We gathered data on each flood event from the Global Flood Database (GFD), which provides comprehensive tracking of floods in China from 2000 to 2018. This database documents a total of 189 flood events within China. Given GFD offers satellite maps that record flood events for every county, we are able to collect data regarding the length of flooding experienced by each pixel (30m $\times$ 30m). Additionally, the database allows us to identify whether a pixel includes permanent water bodies, which "are consistently identified with the presence of surface water for the majority of observations in 2000-2018 at 30 meter resolution which was resampled to 250m resolution in Google Earth Engine using nearest neighbor resampling.", according to GFD. Using Global Flood Database (GFD), we are able to construct three county-level proxies of flood exposures.

Size of Flood Inundation (total size of inundation in a flood event in a county)

$$\text{Size of Inundation}_{ift} = \sum_{j \in A_i} I(\text{Flood Duration}_{jft}) > 0$$

where A_i represents pixels that have not contained permanent water in county i, Flood Duration$_{jft}$ indicates the flood duration in a non-permanent water pixel j in flood event f at time t.

Flood Duration (total flood duration experienced by all non-permanent water pixels in a flood event in a county)

$$\text{Total Flood Duration}_{ift} = \sum_{j \in A_i} \text{Flood Duration}_{jft}$$

<u>*Size-Adjusted Flood Exposure*</u> (average flood duration of each non-permanent water pixel in a flood event in a county)

First, we identify all the pixels within a county that are not occupied by permanent water bodies. Next, we look at every flood event individually, adding together the duration of flooding for each non-permanent water pixel to get the county's total flood duration for each flood event. Finally, to proxy flood risk of each county, we divide the county's flood duration by the count of non-permanent water pixels. We believe that this index provides a nuanced quantification of flood risk, adjusted for the spatial extent of the county's land area susceptible to flooding.

Following this thought, we define the size-adjusted flood duration as

$$AdjustedFloodExposure_{ift} = \frac{\sum_{j \in A_i} FloodDuration_{fjt}}{|A_i|}$$

where $AdjustedFloodExposure_{ift}$ indicates the size-adjusted flood exposure at flood event f that happened at time t. A_i represents pixels that have not contained permanent water in county i. $FloodDuration_{fjt}$ is the number of flooded days experienced by non-permanent water pixel j at the flood event f of time t. It will be 0 if the non-permanent water pixel has not been flooded at the flood event. And it will take a positive value if that non-permanent water pixel has been flooded at the flood event. Here, we define a pixel as a flood-pixel at a flood event f if that pixel: (i) has not contained permanent water previously, which means $j \in A_i$; (ii) but has been marked as flooded by Global Flood Database in the flood event f of time t. Hence, $\sum_{j \in A_i} FloodDuration_{fjt}$ measures the total sum of flood duration experienced by non-permanent water pixels in county i at flood event f of time t. By dividing this sum by total number of non-permanent water pixels $|A_i|$, we adjust the total sum of flood duration by the size of non-permanent water in county i. Figure 1.3 demonstrates that size-adjusted flood exposure is higher in FDB counties. From 2000 to 2018, FDB counties consistently experience higher levels of flood exposure. Notably, the peaks in the graph around 2003, 2006, 2010, and 2014 highlight periods where FDB counties face substantially increased flood risks, due to flood water detention.

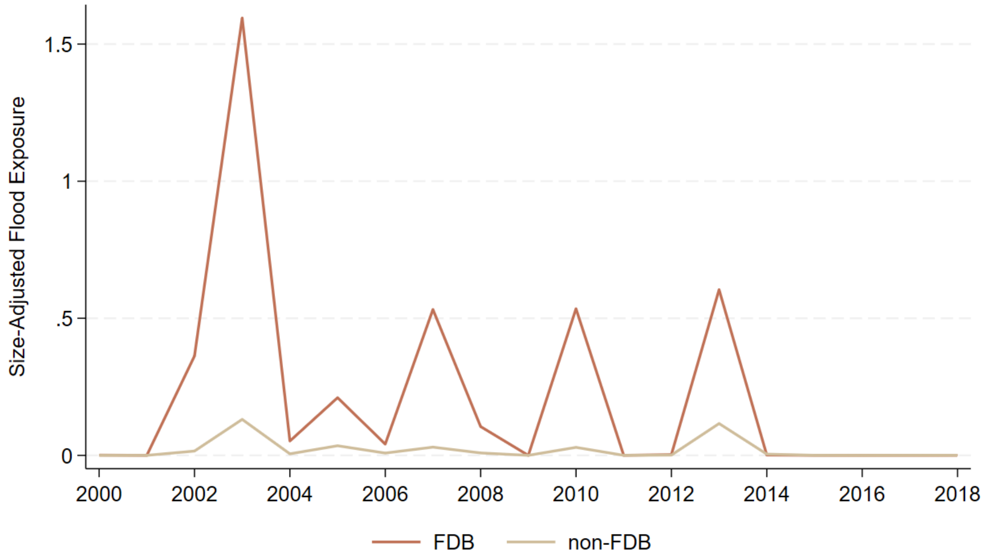


Figure 1.3: Size-Adjusted Flood Exposure in FDB and non-FDB Counties
Note: The size-adjusted flood exposure is calculated using Global Flood Database and measures the average days of inundation experienced by a non-permanent water pixel in a county.

1.4.2 Quantify the Flood Exposure Redistribution Rate

Figure 1.3 straightforwardly demonstrates that the size-adjusted flood exposure is much higher in FDB counties, compared to non-FDB counties. We then use the following specification to determine whether the flood exposure in FDB counties is significantly higher than non-FDB counties.

$$ln(Exposure_{ijt}) = \alpha + \beta_1 FDB_{ijt} + \beta_2 X_{ijt} + \gamma_j + \theta_t + \varepsilon_i$$

where $ln(Exposure_{ijt})$ is the proxy of flood risk in county i, city j, at year t. In our setting, we use two proxies to investigate the impact of FDB policy on flood exposure. The first proxy is the size of inundation area. And the second one is the size-adjusted flood exposure (detailed explanation can be found in Section 1.3.1), which measures the average days of flood inundation of a county in a flood event. FDB_{ijt} is a dummy that equals 1 if the county i is a FDB county, and 0 if not. γ_j represents the city fixed effect, and θ_t represents time fixed effect. ε_i is the standard error that is clustered at city level. X_{ijt} contains geographical controls (precipitation, elevation and slope), which are important determinants of floods. β_1 then measures whether FDB counties have a higher flood exposure than other counties in a given city, holding geographical factors constant.

As indicated in Column (1) and (2) of Table 1.1, we find that after controlling for important geographical controls, the size of flood inundation area in FDB counties is more than 50% higher in FDB counties than other counties in the same city. Column (5) and (6) also suggest that the size-adjusted flood exposure is 5% higher in FDB counties, compared to other counties in the same city. This empirical evidence supports the claim that FDB policy induces flood risk redistribution across different regions. In other words, FDB counties tend to absorb more flood water according to the policy design.

Table 1.1: Impacts of FDB Policy on Flood Exposure

Sample Period:	Flood Size		Flood Duration		Flood Exposure per Pixel	
2000-2020	(1)	(2)	(3)	(4)	(5)	(6)
FDB	0.602***	0.547***	0.662***	0.574***	0.050***	0.043***
	(0.090)	(0.087)	(0.096)	(0.090)	(0.010)	(0.010)
N(obs)	52,307	52,307	52,307	52,307	52,307	52,307
Controls						
Precipitation	N	Y	N	Y	N	Y
Slope	N	Y	N	Y	N	Y
Elevation	N	Y	N	Y	N	Y
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
City	Y	Y	Y	Y	Y	Y

Note: (1) This table presents results of fixed-effect regression: $ln(Flood_{ijt}) = \alpha + \beta_1 FDB_{ijt} + \beta_2 X_{ijt} + \gamma_j + \theta_t + \varepsilon_i$, $ln(Flood)_{ijt}$ indicates flood-related outcomes in county i, city j, at year t, FDB_{ijt} is a dummy variable that equals 1 if the county i is an FDB county in year t, and 0 if not, X_{ijt} are geographical controls, γ_j is city fixed effect, λ_t is time fixed effect, standard errors are clustered at the county level; (2) We have three types of flood-related outcomes. 'Size of Flood Inundation' measures the area of flood inundation in each county, 'Total Flood Duration' measures the total flooded day experienced by all non-permanent while 'Size-Adjusted Flood Exposure' measures the average days of flood inundation experienced by a non-permanent water pixel in a county. Detailed calculation is introduced in Section 1.3.1.

1.4.3 Hydrological Analysis based on Hydro-Dynamic Model

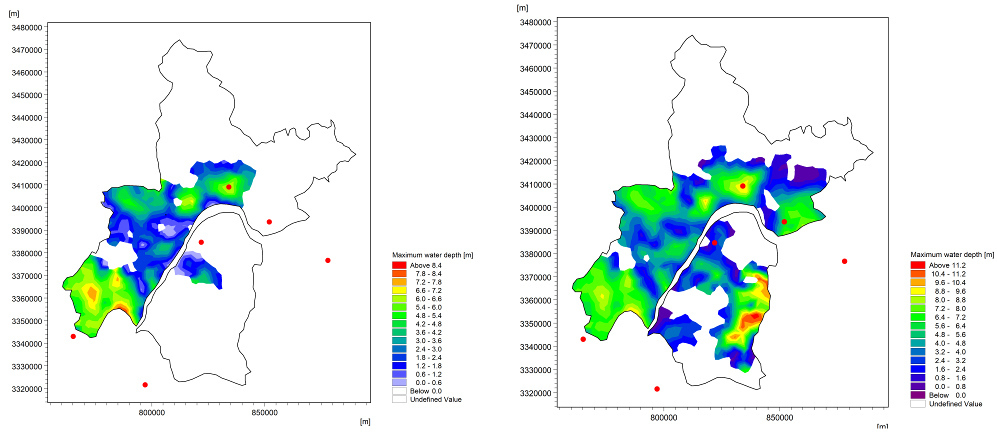
According to a hydro-logical research by Mingkai and Kai 2017, "inundated farmland in the downstream would be increased to 2530 hectares, with an increased area of 1340 hectares more than the use of the Mengwa Detention Basin." To rigorously quantify the level of floodwater redistribution, we incorporate an interdisciplinary approach and employ a hydro-dynamic engineering model developed under the supervision of the Danish Hydraulic Institute (DHI) to measure the flood exposure redistribution rate during a real flood event. The hydro-dynamic model is a sophisticated tool used for simulating water flow, particularly in river basins and floodplain areas. It accounts for variables such as topography, water velocity, flow rates, and human interventions, making it highly suitable for assessing the impacts of floodwater management policies like the Flood Detention Basin (FDB) policy. We specifically choose Wuhan for this analysis because of its economic importance, and its size is comparable to that of the FDB counties. This makes it easier to translate the flood protection benefits observed in Wuhan to the flood water absorbed by the FDB regions.

The process of implementing the model consists of several key steps. First, we collect high-resolution geographical shape data, river runoff data, and detailed policy information on floodwater diversion. These inputs are essential to build an accurate representation of the river system and floodplain in question, including the areas designated as FDB zones. The geographical data defines the physical characteristics of the region, while the runoff data provides insight into how much water the rivers and floodplains can handle during heavy rainfall or extreme flood events. The FDB policy details, on the other hand, establish the parameters of water diversion in our model.

Next, we calibrate the model using historical flood data to ensure its accuracy. This involves adjusting model parameters until the simulated outcomes closely match the observed data from past flood events. Calibration is a crucial step because it ensures that the model is reliable and that its predictions reflect real-world conditions. By checking the consistency of model predictions with actual flood patterns, we validate the model's capacity to predict the effects of floodwater redistribution accurately.

After calibration, we simulate a counterfactual scenario where the floodwaters are not diverted to the FDB areas. This simulation allows us to assess what would happen in the absence of the flood diversion policy. The model predicts how floodwaters would behave if allowed to flow freely without the designated intervention, providing us with a comparison between the actual and hypothetical scenarios.

Finally, we compare the size of the inundation area in Wuhan City between the actual scenario, where floodwaters are diverted into the FDB regions, and the counterfactual scenario without diversion. As shown in Figure 1.4, the inundation area in Wuhan, an important city intended to be protected by the FDB policy, increases by 45% in the absence of floodwater diversion. This significant increase in the flooded area highlights the crucial role that the FDB policy plays in mitigating flood risks for urban centers.



(a) Actual Case: with FDB **(b)** Counterfactual Case: without FDB

Figure 1.4: Inundation Map in Wuhan City (Actual v.s. Counterfactual)
Note: (1) The map is drawn using MIKE hydrological modelling software launched by Danish Hydraulic Institute (DHI); (2) Model: hydro-dynamic model; (3) We select Wuhan city because this city is a major protected city by FDBs in Yangtze Rivers; (4) The flood exposure redistribution rate based on this estimation is 45%.

1.5 Economic Costs on FDB Counties

After confirming that flood exposures in FDB counties are significantly higher than other counties, we extend our analysis into economics. In this section, we aim to quantify the economic impact of FDB policy on selected FDB counties. Here, we mainly focus on nighttime light intensity, a proxy of GDP. However, in future research, we plan to extend our analysis to more individual level outcomes, for example, education and health outcomes.

1.5.1 Main Result: Impacts of FDB Selection on Nighttime Light

To quantify the economic costs on FDB counties, we examine the impact of FDB policy on nighttime light intensity. We choose nighttime light as a proxy for economic activity over GDP for two reasons. First, county-level GDP data before 2000 is unavailable, making it impossible for us to compare pre-treatment and post-treatment outcomes of the 2000 policy change. Second, nighttime light is a more credible indicator of economic activity in China in that Chinese GDP figures announced by the government may not be accurate (Martinez 2022), and Zeng and Zhou 2024).

In our Difference-in-Differences approach, the treatment is the designation of a county as an FDB site. Since the government first announced the FDB list in 2000, and made revisions in 2010. In other words, if a county is selected into the FDB list in 2000 (2010), then this county would be considered as treated in and after 2000 (2010). For the control group, we exclude four types of counties: (1) counties located within protected urban areas, as these counties receive a different treatment by being protected through FDBs; (2) counties that were removed from the FDB list in 2010, as the treatment status has changed across time; (3) counties adjacent to FDB counties, since floodwaters may flow into these neighboring areas; and (4) counties adjacent to protected cities, as these counties may receive implicit protection. Thus, our control group includes counties that are not directly targeted by the FDB policy. An illustrative explanation can be found in Figure 1.5. We also label different counties in Figure 1.6.

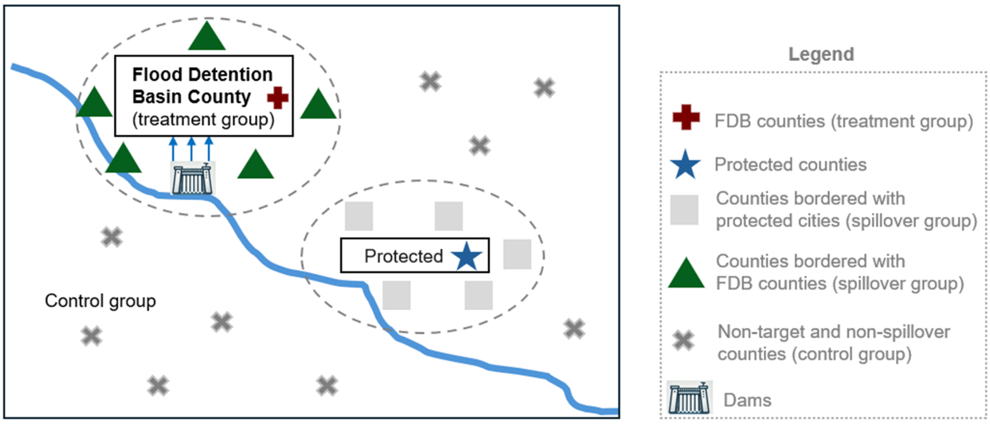
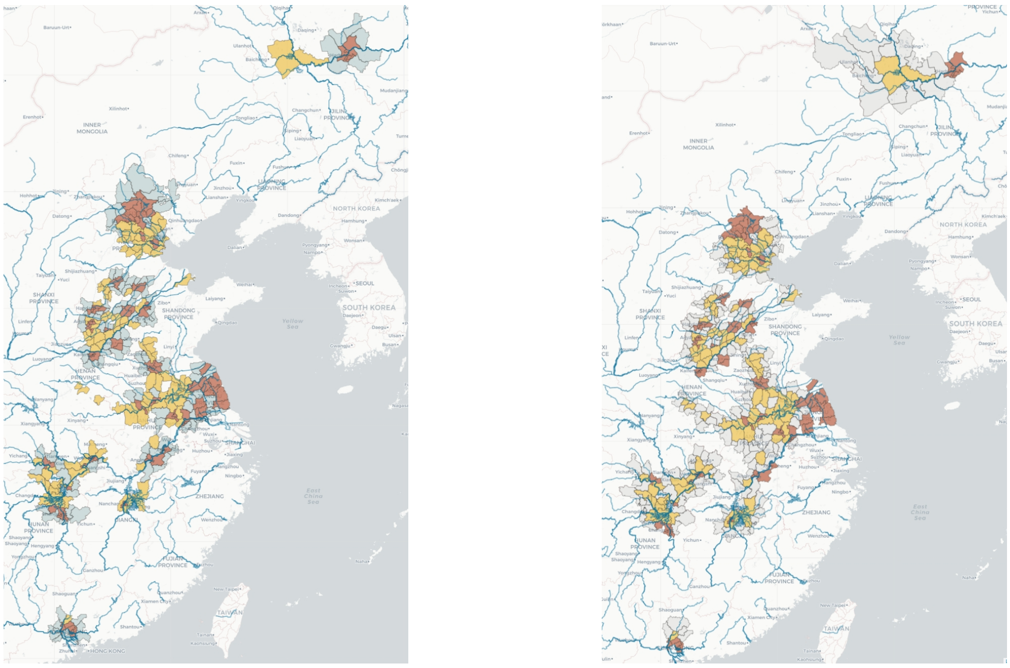


Figure 1.5: Treatment Group, Spillover Group, and Control Group



(a) FDB (& Spillover), and Protected

(b) FDB, and Protected (& Spillover)

Figure 1.6: FDB Counties, FDB-Protected Counties, and Spillover Counties
Note: (1) In Figure a, FDB counties are marked using color yellow, FDB-protected counties are marked using color red, and FDB-Spillover counties are marked using color green; (2) In Figure b, FDB counties are marked using color yellow, FDB-protected counties are marked using color red, and FDB-Spillover counties are marked using color gray;

Table 1.2 presents the main empirical result. Panel A of Table 1.2 presents results using traditional two-way fixed-effect difference-in-differences (TWFE DID) estimates without any controls. In Column (1), we find that county-level nighttime light intensity would decrease by 17.6% if a county is selected into the FDB list. Considering recent discussions on the properties of the staggered DID approach (e.g., Borusyak et al. 2024), potential biases may arise from the weighting problem. Therefore, we separately investigate the impacts of the 2000 and 2010 policy changes in Columns (2) and (3). Column (2) shows that county-level nighttime light intensity would decrease by 17.6% (7.8%) if a county is selected into the 2000 (2010) FDB list, respectively.

Panel B of Table 1.2 reports results using the synthetic difference-in-differences (SDID) approach proposed by Arkhangelsky et al. (2021). We believe SDID is appropriate for our empirical setting for three reasons. First, constructing a counterfactual group using synthetic weights (Abadie et al. 2010) addresses concerns about the weighting problem in traditional TWFE DID. Second, as suggested by Roth et al. (2023), clustering at the unit level is not suitable when the number of treated groups is small. In the 2010 policy change, the Chinese government selected 20 new counties and removed 10 from the list. Given the small size of treated clusters, using bootstrap standard errors offered by the SDID approach is more appropriate. Third, synthetic weight construction helps mitigate potential threats to exogeneity by creating a counterfactual whose pre-treatment outcomes are parallel to the treatment group. Results in Panel B are robust and indicate a negative impact of being selected into the FDB list on nighttime light intensity, with magnitudes similar to those in Panel A.

In Column (4) of both Panels A and B, we focus on the impact of removal from the FDB list in 2000. The results in both panels are not significant, indicating that being removed from the FDB list does not lead to significant economic recovery. We interpret this as a ‘scarring effect,’ where counties once selected into the FDB list struggle to recover even after removal. We consider the result in Column (4) of Panel B to be more credible than that in Panel A, given the small number of counties removed from the list, making SDID more appropriate than TWFE.

Figure 1.7 illustrates the dynamic impacts of the FDB policy on nighttime light intensity using an event-study approach. Before the treatment, there is no significant difference between the treated and control groups. This suggests that the treated and control groups followed similar trends in nighttime light intensity prior to the policy intervention, validating the parallel trend assumption. Immediately after the implementation of the FDB policy, we observe a noticeable and persistent decline in nighttime light intensity for the treated counties. This indicatres both immediate and lasting adverse effects of the FDB policy on economic activity as proxied by nighttime light intensity. We present the SDID event-study results in Figure B1.

1.5.2 Interpreting Effect Size: from Light to GDP

According to column (2) in Panel B of Table 1.2, being selected into the FDB list in 2000 results in a 10.7% decrease in nighttime light intensity. Various studies have examined the elasticity between nighttime light intensity and GDP, allowing us to translate this reduction into a loss in real GDP. Henderson et al. (2012) find that the elasticity of GDP with respect to nighttime lights is 0.277, which is supported by Martinez (2022), who finds an elasticity of 0.296. Additionally, Martinez (2022) notes that elasticity is higher in non-democratic regimes, estimating an elasticity of 0.312 for China. This translates into an annual GDP loss of 2.96%, 3.17%, and 3.34%, respectively. Using real GDP data from Chen et al. (2022), we estimate the GDP loss to be $9.84 billion, $10.54 billion, and $11.13 billion, respectively, based on the elasticities from Henderson et al. (2012) and Martinez (2022). On average, an FDB county tends to lose $0.10-0.12 billion per year due to being selected into the FDB list. To validate these findings, we conducted an interdisciplinary cross-check. Our results align with a hydrological case study by Wang et al. (2021), published in the leading hydrological journal *Journal of Hydrology*, which also reports an annual economic loss of $0.1 billion for an FDB county in Yangtze River.

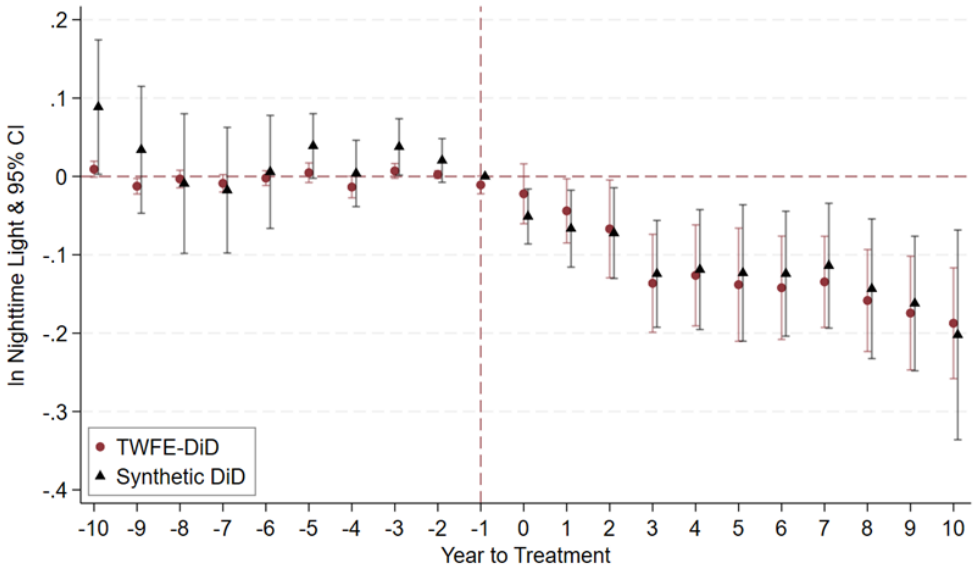


Figure 1.7: Dynamic Impacts of FDB on Nighttime Light Intensity
Note: (1) Black dot represents the policy effect (ATT) estimated using TWFE-DID, while red dot represents the policy effect (ATT) estimated using DiD with synthetic weights; (2) Data: 1990-2020 Nighttime Light Intensity data; (3) 96 counties were selected into the FDB list in 2000, while 20 counties were selected into the FDB list in 2010; (3) The event-study regression includes county and year fixed effects, standard errors are clustered at county level; (4) We report the confidence interval at 95% confidence level.

Table 1.2: Main Results: Impacts of FDB on Nightime Light Intensity

	Selection into FDB			Removal from FDB
(ln)	All	2000 Cohort	2010 Cohort	
Panel A: Method - Traditional TWFE Difference-in-Differences				
	(1)	(2)	(3)	(4)
$\beta^{TWFE}_{Selection}$	-0.176***	-0.137***	-0.078*	
	(0.056)	(0.035)	(0.045)	
$\beta^{TWFE}_{Removal}$				-0.052
				(0.074)
Panel B: Method - Synthetic Difference-in-Differences (Arkhangelsky et al. 2021)				
	(1)	(2)	(3)	(4)
$\beta^{SDID}_{Selection}$	-0.156***	-0.107***	-0.078**	
	(0.025)	(0.015)	(0.039)	
$\beta^{SDID}_{Removal}$				-0.003
				(0.064)
Sample Period	1990-2020	1990-2010	2000-2020	2000-2020
N(obs)	70,463	46,680	47,208	50,148
N(Treated Counties)	106	86	20	10
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) 'Selection into FDB' indicates the treatment of selecting counties into the FDB list in both 2000 and 2010, 'Removal from FDB' indicates the treatment of removing counties from the FDB list, solely in 2010; (2) 'All' includes two treated groups: counties selected into the FDB list in 2000, and in 2010, '2000 Cohort' focuses only on one treated group: counties selected into the FDB list in 2000, '2010 Cohort' focuses only on one treated group: counties selected into the FDB list in 2010; (3) We deliberately select control groups to remove possibly spillover groups and groups that receive other treatments, as indicated in Figure 1.5.

1.5.3 Heterogeneity Analysis

We then examine the heterogeneous impacts of the FDB policy on nighttime light intensity across different FDB classifications established by the Chinese government: Important FDB counties, General FDB counties, and Reserved FDB counties. These classifications are based on each FDB's hydrological capacity to absorb floodwaters. Due to historically high flood risks in China, Important FDBs may have already served as de facto FDBs prior to the policy announcement, while Reserved FDBs are likely regarded as designated areas for floodwater diversion following the policy's implementation.

Our findings in Figure 1.8 and Table 1.6 reveal that nighttime light intensity decreases the least in Important FDB counties (11.6%). On the other hand, light has decreased by 30.8% and 16.6% in Reserved and General FDB counties. The findings suggest that counties historically exposed to frequent flooding, like Important FDBs, have developed better expectations for flood events. As a result, while nighttime light intensity decreases in Important FDB counties, the decline is less significant than in other FDB categories. In contrast, General and Reserved FDB counties, which lack a history of frequent flooding, face a more substantial reduction in light intensity, as the FDB designation introduces an unexpected economic shock. This sudden risk leaves these regions more vulnerable, leading to greater negative impacts on economic activity. The key difference lies in the anticipation effect: Important FDBs, having established flood expectations and adaptive measures, experience a moderated impact, while General and Reserved FDBs suffer more severe economic setbacks due to the policy-induced risks. This analysis indicates our cost estimates may underestimate total costs by not accounting for the costs on important FDBS befor the policy announcement.

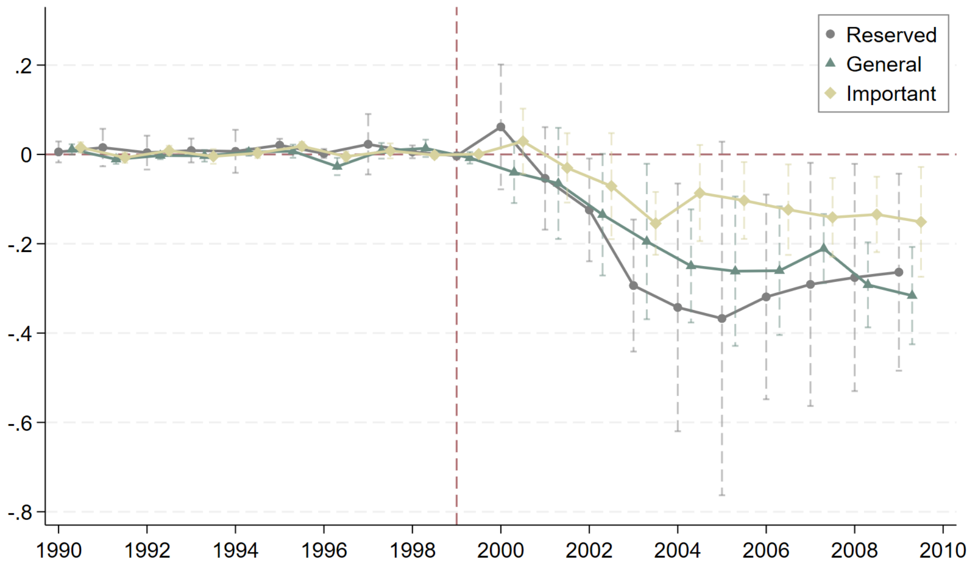


Figure 1.8: Heterogeneous Impact of 2010 Policy Change on Nighttime Light Intensity
Method: SDID (Arkhangelsky et al. 2021)
Note: (1) Each dot represents the policy effect (ATT) estimated using the event-study approach; (2) Data: 1990-2020 Nighttime Light Intensity data; (3) 96 counties were selected into the FDB list in 2000, while 20 counties were selected into the FDB list in 2010; (3) The event-study regression includes county and year fixed effects, standard errors are clustered at county level; (4) We report the confidence interval at 95% confidence level; (5) We classify FDB counties into three categories: Important, General, and Reserved according to the government classification. The likelihood of being flooded is the highest for Important FDBs, and the lowest for Reserved FDBs.

Table 1.3: Heterogeneous Impacts of FDB on Nightime Light Intensity

		Type of FDBs		
	All Sample	Reserved FDB	General FDB	Important FDB
Sample Period: 1900-2020	(1)	(2)	(3)	(4)
$\beta^{SDID}_{Selection}$	-0.156^{***}	-0.308^{***}	-0.166^{***}	-0.116^{***}
	(0.025)	(0.079)	(0.043)	(0.043)
N(obs)	70,463	69,316	69,998	69,657
N(Treated Counties)	106	16	46	44
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) We use the SDID appraoch proposed by Arkhangelsky et al. (2021); (2) Data: 1990-2020 Nighttime Light Intensity data; (3) 96 counties were selected into the FDB list in 2000; (4) Standard Error: Bootstrap; (5) We also report the confidence interval at 95% confidence level; (6) We classify FDB counties into three categories: Important, General, and Reserved according to the government classification. The likelihood of being flooded is the highest for Important FDBs, and the lowest for Reserved FDBs.

1.5.4 Robustness and Placebo

In Figure B2 and Table B2, we report our results using other difference-in-differences methods. Although we believe that synthetic difference-in-differences (Arkhangelsky et al. 2021) is the most suitable method in our setting, we report the event-study results using different methods proposed by De Chaisemartin and d'Haultfoeuille (2020), Gardner (2022), and Callaway and Sant'Anna (2021). The robustness checks demonstrate that our main findings are consistent across these alternative methodologies. Specifically, the results in Table 1.2 are robust in terms of both statistical significance and magnitude when using other difference-in-differences approaches. Overall, the consistency of our findings across multiple methodologies underscores the validity of our results and the robustness of our conclusions.

In Figure B3, we conduct three distinct types of placebo tests: the in-time placebo test, the in-space placebo test, and the mixed placebo test. In the in-time placebo tests, we forward the treatment time by several years, using fake treatment times to assess if our results are driven by temporal trends rather than the actual intervention. This result is consistent with our event-study analysis (Figure 1.7) that we do not find significant evidence that argue against the parallel trend assumption. For the in-space placebo tests, we assign treatment to randomly selected units that did not receive the intervention. By assigning fake treated units, we are able to test the robustness of our findings against spatial confounding factors. Lastly, the mixed placebo tests combine both approaches by randomly assigning fake treatment units and times. The results shown in Figure B3 indicate that our main findings hold up under these placebo tests, as the estimated effects do not show significant deviations from zero, thus confirming the robustness and validity of our original results.

1.5.5 Individual-Level Outcomes

A comprehensive analysis of the costs associated with the FDB policy requires more than just evaluating total outputs, as we demonstrate in this section. To fully assess these costs, it is crucial to account for socio-economic factors affecting individual well-being. Unfortunately, data limitations in China prevent us from conducting a thorough examination of key outcomes such as health and education. To address this gap, we use data from the 2010, 2012, 2014, 2016, 2018, and 2020 waves of the China Family Panel Study (CFPS). Our correlation analysis reveals that, after controlling for city and time fixed effects, residents of FDB counties earn approximately 20% less than those in non-FDB counties. This result further highlights the economic disadvantage faced by individuals in FDB areas. We describe our detailed results in Appendix A.2.2.

1.6 Exploring Mechanisms of Costs on FDB Counties

In this section, we examine the primary factors contributing to economic underdevelopment in FDB counties, focusing on three key channels: (1) migration, (2) agriculture, and (3) manufacturing. Ultimately, we identify firm responses as the main mechanism driving economic underdevelopment in these regions.

1.6.1 Migration Channel

A natural hypothesis is that rational individuals will leave FDB counties, leading to a loss of labor which results in economic underdevelopment. However, as shown in Figure 1.9, we do not find significant evidence of people leaving FDB counties. Although there is a downward (upward) trend of registered population after counties being selected (removed) from the FDB list, we do not find the estimate being neither economically significant nor statistically significant, indicating that migration decision is not sensitive to FDB policy. Extensive literature has demonstrated the difficulty of individuals in developing countries to make rational migration decisions, as summarized in Lagakos (2020). For China specific studies, we would like to propose several possible reasons that people do not migrate in response to FDB policy.

First, according to the seminal work of Zhao (1999), the existing arrangement of land management is a major reason why rural people in China choose not to migrate in spite of the incentive and ability to migrate. In the early 1980s, the Chinese government

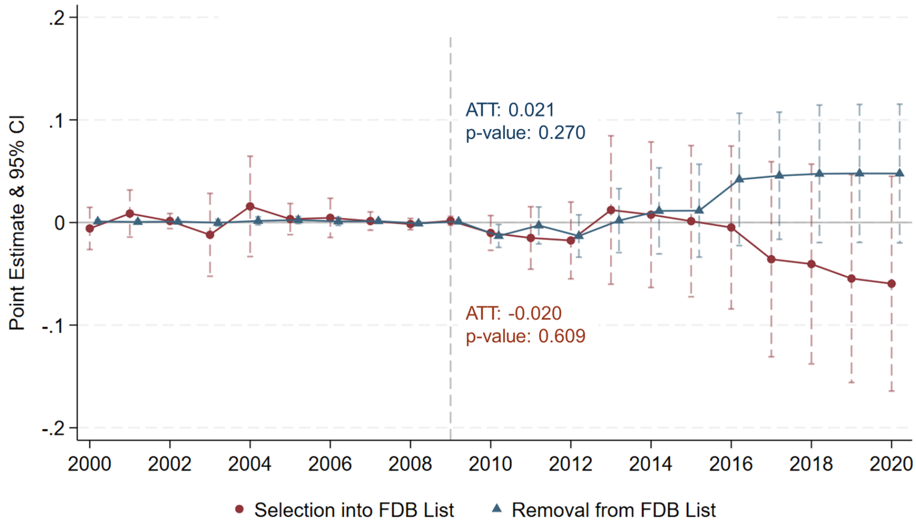


Figure 1.9: Dynamic Impacts of 2010 FDB Policy Change on Registered Population
Note: (1) Each dot represents the policy effect (ATT) estimated using the SDID event-study approach by Arkhangelsky et al. (2021)); (2) Data: 2000-2020 county-level statistical yearbook; (3) 20 counties were selected into the FDB list in 2010, while 10 counties were removed from the FDB list in 2010; (3) The event-study regression includes county and year fixed effects; (4) Standard Error: Bootstrap; (5) 'Registered population' refers to the population who registers as the official resident of the county.

introduced the Household Responsibility System that grants rural households land use rights and income rights over lands. Although land belongs to the village, land allocation within villages was highly egalitarian, resulting in minimal per capita differences in landholdings among households within a village. A recent paper by Adamopoulos et al. (2024) also indicates that the land system is a major friction of rural-urban migration.

Second, the Chinese government has not designed a suitable incentive scheme to motivate FDB residents to leave. According to the latest migration subsidy plan in 2017, the government compensates \$2.4k per person, which is significantly less than the \$8.1k per person provided under *the Relocation for Poverty Alleviation* program and is insufficient to cover migration costs. According to a survey conducted by *the Huai River Regulation Commission of the Ministry of Water Resources*, 93% of residents in the Mengwa Flood Detention Basin are dissatisfied with the migration subsidy provided by the government, and 94% are unhappy with the proposed migration destinations.

Table 1.4: Impacts of 2010 FDB Policy Change on Registered Population

	Selection into FDB List		Removal from FDB List	
Sample Period: 2000-2020	(1)	(2)	(3)	(4)
$\beta^{SDID}_{Selection}$	-0.020	-0.020		
	(0.039)	(0.030)		
$\beta^{SDID}_{Removal}$			0.021	0.021
			(0.019)	(0.052)
Standard Error	Bootstrap	Placebo	Bootstrap	Placebo
N(obs)	43,050	43,050	43,050	43,050
N(Treated Counties)	20	20	10	10
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) We use SDID approach by Arkhangelsky et al. (2021); (2) Data: 2000-2020 county-level statistical yearbook; (3) 20 counties were selected into the FDB list in 2010, while 10 counties were removed from the FDB list in 2010; (3) We use two types of standard errors (bootstrap and placebo), county and year fixed effects are included; (4) 'Registered population' refers to the population who registers as the official resident of the county.

This dissatisfaction reflects broader issues in the policy's design, including inadequate financial support and poorly planned relocation sites, which fail to meet the needs and preferences of the affected residents. Consequently, the lack of proper incentives and satisfactory relocation plans has resulted in non-optimal migration from FDB counties.

1.6.2 Loss in Agriculture or Manufacturing?

We also investigate whether the costs associated with flooding are predominantly caused by its impact on agriculture. Given that FDB counties primarily depend on agriculture, it is plausible that floods would incur significant costs by damaging agricultural crops. However, our findings (Figure 1.10) do not show significant evidence of a decline in agricultural output, with the observed change being minimal (0.3%). This resilience in agricultural output could be possibly attributed to the geographical conditions of China's agricultural land. For instance, in Hunan Province, the quality

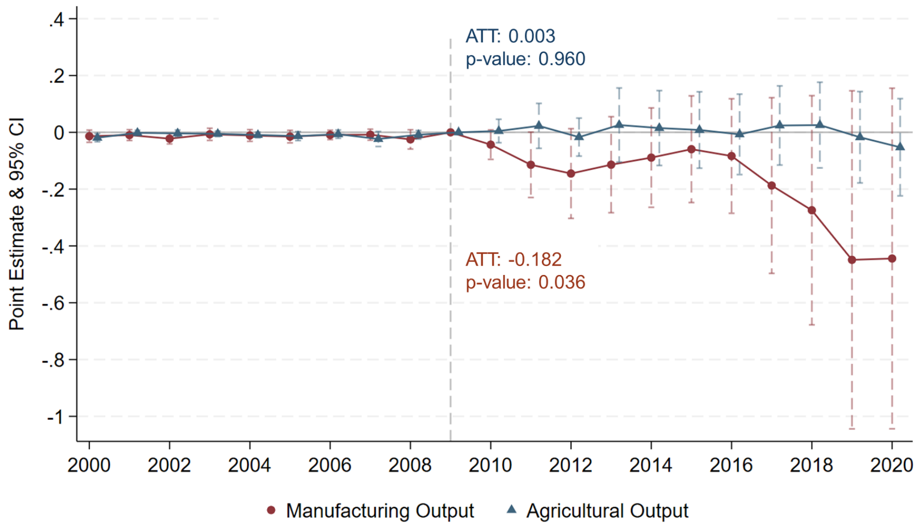


Figure 1.10: Dynamic Impacts of 2010 FDB Policy Change on Manufacturing and Agricultural Output

Note: (1) Each dot represents the policy effect (ATT) estimated using the SDID event-study approach by Arkhangelsky et al. (2021); (2) Data: 2000-2020 county-level statistical yearbook; (3) 20 counties were selected into the FDB list in 2010; (3) The event-study regression includes county and year fixed effects; (4) Standard Error: Bootstrap.

Table 1.5: Impacts of 2010 FDB Policy Change on Agricultural and Manufacturing Output

	ln(Agricultural Output)		ln(Manufacturing Output)	
Sample Period: 2000-2020	(1)	(2)	(3)	(4)
$\beta^{SDID}_{Selection}$	0.003	0.003	-0.182^{***}	-0.182^{***}
	(0.059)	(0.054)	(0.087)	(0.081)
Standard Error	Bootstrap	Placebo	Bootstrap	Placebo
N(obs)	39,354	39,354	39,354	39,354
N(Treated Counties)	20	20	20	20
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) We use SDID approach by Arkhangelsky et al. (2021); (2) Data: 2000-2020 county-level statistical yearbook; (3) 20 counties were selected into the FDB list in 2010; (3) We use two types of standard errors (bootstrap and placebo), county and year fixed effects are included.

of arable land tends to improve after floods, which may mitigate the adverse effects. Additionally, farmers in the southern region can harvest three times a year, so even if they suffer flood damage during the rainy season, they can partially compensate for the losses through winter crops.

In contrast, manufacturing output experiences a substantial and significant decrease of 18.2%. Specifically, there was a sustained output reduction of about 20% during the initial five years (2010-2015), which widened to approximately 40% post-2016. This suggests that the FDB policy has a lasting negative impact on manufacturing activities within FDB counties. This stark decline underscores the lag in structural transformation within FDB counties. While farmers adapt to new policies, they remain largely confined to agriculture due to limited opportunities for transitioning into the manufacturing sector.

1.6.3 Firm Response Effect

We propose the 'firm response effect,' suggesting that firms have less incentive to enter and invest in counties with higher flood risk, leading to an underdeveloped manufacturing sector in FDB counties. This hypothesis has two empirical implications. First, when a county is added to the FDB list, firms are less likely to enter and invest in that county. Second, when a county is removed from the FDB list, firms begin to reenter and invest. In 2010, the Chinese government added 20 counties to the FDB list and removed 10 counties from it, allowing us to empirically test the 'firm response effect' hypothesis.

In this section, we present balanced and symmetric results of three different outcomes that show both the impact of being added to the FDB list and the impact of being removed from the list. By comparing these two scenarios, we can confirm that the FDB policy significantly influences firms' entry and investment decisions. Specifically, we find a decline in firm entry and investment in counties added to the list, and an increase in firm entry and investment in counties removed from the list. These balanced and symmetric findings serve as strong evidence to rule out other possible mechanisms and underscore the exclusive impact of FDB policy on firms' decision making.

It would be ideal for us to study the causal impact of both 2000 policy and 2010 policy, especially the 2000 policy given its importance. However, the unavailability of firm-level data prior to 2000 makes us impossible to construct pre-treatment counterfactual control groups. Hence, we have to restrict our examination to the causal impacts of 2010 policy on various firm level outcome variables.

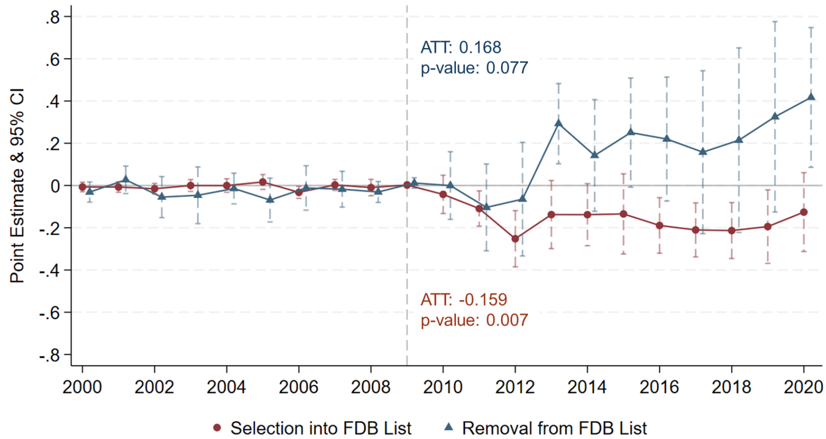
Firm Entry - The increased flood risk in FDB counties necessitates higher expected returns on investment for firms considering entry into these areas. Consequently, firms have less incentive to enter FDB counties. In other words, the increase in flood risk acts as a deterrent for new firm entry. To explore this intuition, we examine the impact of the 2010 FDB policy change on firm entry using the Annual Registration Data of Chinese Enterprises from 2000 to 2020. In Panel A of Figure 1.11, we find balanced and symmetric impacts of selection into and removal from the FDB list. Each dot in the figure represents a point estimate, showing the difference between actual FDB counties and their synthetic counterparts. Prior to 2010, the proximity of these estimates to zero, coupled with their statistical insignificance, confirms that our synthetic group effectively mirrors the counterfactual FDB counties.

The negative impact on firm entry in these counties is immediate and persists over a decade, as evidenced by the consistently negative and significant coefficients observed even in 2020. One year after the policy implementation, in 2011, firm entry in FDB counties decreased by approximately 10.9%. In 2012, this decrease grew to around 25.2%. The negative impact then persists from 2013 to 2021, stabilizing at around 15%. This empirical evidence supports our theory that firms lack incentives to enter counties newly designated as FDB-county. Conversely, we also find that firms begin to reenter counties removed from the FDB list. Although the impact is not immediate, by 2013 we observe a significant increase in firm entry, with a magnitude of 29.2%. This positive impact persists until 2020.

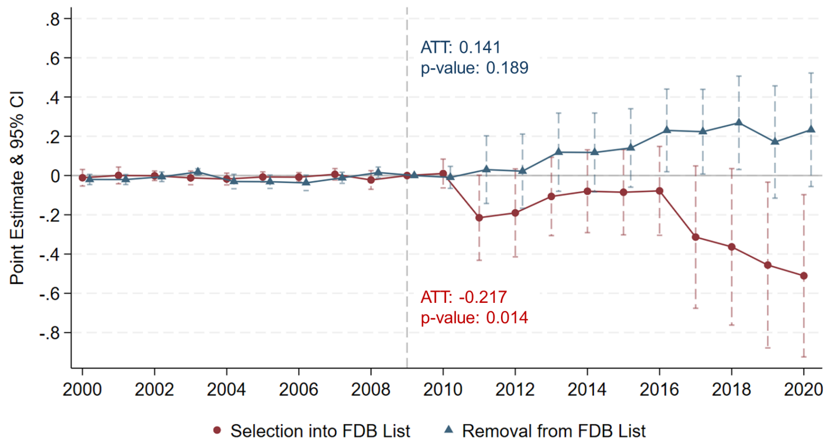
Regarding the average treatment effect, we find that firm entry tends to significantly decrease by 15.9% after a county is selected into the FDB list. This indicates that selection into the FDB list diminishes the county's attractiveness for the entry of manufacturing firms. On the other hand, firm entry tends to significantly increase by 16.8% after a county is removed from the FDB list. The balanced and symmetric result indicate the importance of FDB policy in affecting firms' entry decisions.

Number of Large Manufacturing Firms - In Panel B of Figure 1.11, we present robust evidence that the FDB policy influences firm entry decisions, focusing specifically on the number of larger manufacturing firms. Using county-level statistical yearbook data from 2000 to 2010, we find that the average number of larger manufacturing firms in a county significantly decreases by 21.7% after the county is included in the FDB list in 2010. Conversely, when a county is removed from the FDB list, the number of larger manufacturing firms increases by 14.1%, although this change is not statistically significant. Comparing the results of Panel B with those of Panel A, we observe that the impact of being added to the FDB list is more pronounced for larger manufacturing firms compared to all firms. However, when a county is removed from the FDB list, larger manufacturing firms show more hesitation in re-entering these counties, while all firms tend to respond more sensitively to the policy change. This suggests that larger manufacturing firms are more cautious in their entry decisions, possibly due to their higher position in fixed asset investments.

Combining the findings from Panel A and Panel B, we conclude that: (i) being included in the FDB list tends to decrease a county's attractiveness for firm entry, whereas removal from the list tends to increase it; (ii) larger manufacturing firms, compared to other firms, are more cautious in their entry decisions.



(a) Outcome: ln(Number of Registered Firms)



(b) Outcome: ln(Number of Large Manufacturing Firms)

Figure 1.11: Dynamic Impacts of 2010 FDB Policy Change on Firm Entry
Note: (1) Each dot represents the policy effect (ATT) estimated using the SDID event-study approach by Arkhangelsky et al. (2021)); (2) Panel A Data: 2000-2020 National Enterprise Credit Information Public System (NECIPS); Panel B data: 2000-2020 county level statistical yearbooks; (3) 20 counties were selected into the FDB list in 2010, while 10 counties were removed from the FDB list in 2010; (3) The event-study regression includes county and year fixed effects; (4) Standard Error: Bootstrap; (5) Larger Manufacturing Firms refer to firms whose annual revenue exceeds US$ 3million.

Table 1.6: Impacts of 2010 FDB Policy Change on Firm Entry

	Selection into FDB List		Removal from FDB List	
Sample Period: 2000-2020	(1)	(2)	(3)	(4)
Panel A: Outcome - ln(Number of Registered Firms)				
$\beta^{SDID}_{Selection}$	-0.159^{***}	-0.159^{***}		
	(0.059)	(0.071)		
$\beta^{SDID}_{Removal}$			0.168^{*}	0.168
			(0.095)	(0.138)
Standard Error	Bootstrap	Placebo	Bootstrap	Placebo
N(obs)	58,191	58,191	58,191	58,191
N(Treated Counties)	20	20	10	10
Panel B: Outcome - ln(Number of Larger Manufacturing Firms)				
$\beta^{SDID}_{Selection}$	-0.217^{***}	-0.217^{***}		
	(0.088)	(0.117)		
$\beta^{SDID}_{Removal}$			0.141	0.141
			(0.107)	(0.116)
Standard Error	Bootstrap	Placebo	Bootstrap	Placebo
N(obs)	41,160	41,160	41,160	41,160
N(Treated Counties)	20	20	10	10
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) We use SDID approach by Arkhangelsky et al. (2021); (2) Panel A Data: 2000-2020 National Enterprise Credit Information Public System (NECIPS); Panel B data: 2000-2020 county level statistical yearbooks; (3) 20 counties were selected into the FDB list in 2010, and 10 counties were removed from the FDB list in 2010; (3) We use two types of standard errors (bootstrap and placebo), county and year fixed effects are included; (4) Larger Manufacturing Firms refer to firms whose annual revenue exceeds US$ 3million.

Fixed Assets Investment - By using spatial Regression Discontinuity (SRD), we provide evidence to indicate that the FDB policy affects firms' investment decision. We specifically focus on fixed assets investment because fixed assets are especially prone to suffering from flood damage because they are either immovable or it is highly challenging to relocate them. Given the data availability constraints that prevent tracking post-2013 data, we concentrate on outcomes likely to be immediately influenced by the FDB policy. We hypothesize that the considerable financial costs associated with either repairing or replacing these assets makes entrepreneurs hesitate to invest in fixed assets situated in FDB counties with higher flood risk.

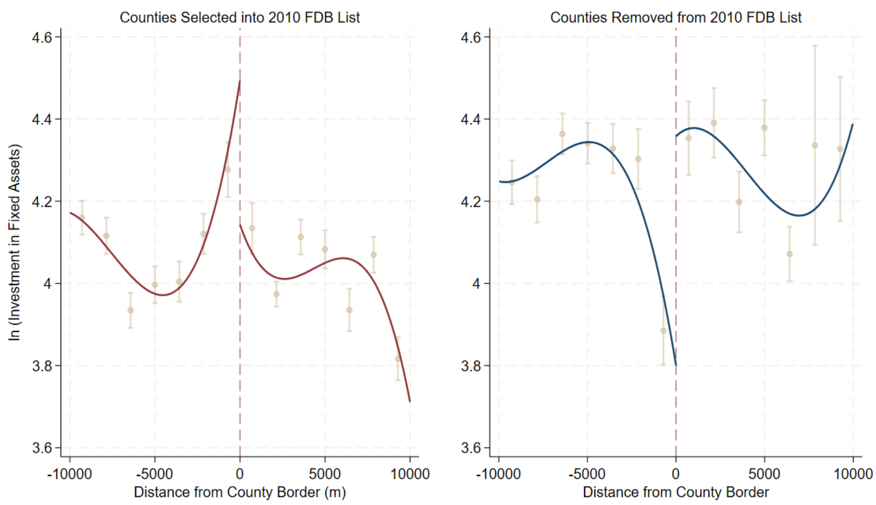
Figure 1.12 displays the logarithm of fixed asset investment, adjusting for both county fixed effects and industry fixed effects, plotted against the distance to the corresponding FDB county boundary. Each point on the graph represents the average logarithmic fixed asset investment for firms within specific distance intervals. And the 95% confidence intervals for these averages are also indicated in the figure. To highlight the policy's impact at the FDB county boundary, a curve fitting these data points is presented on the plot, clearly demonstrating the discontinuity at the boundary of FDB counties.

Panel A of Figure 1.12 presents a regression discontinuity (RD) plot of the residual logarithm of fixed asset investment. In the left sub-figure of Panel A, we explore how being designated as an FDB county influences fixed asset investment. This plot reveals a pronounced decline in fixed asset investment exactly at the boundary of counties newly included in the FDB list. This observation implies that within firms of these newly designated FDB counties, fixed asset investment is substantially lower compared to firms in adjacent counties. Conversely, the right sub-figure of Panel A in Figure 1.12 examines the effects on fixed asset investment following a county's removal from the FDB list. Contrary to Panel A, we observe a significant jump in fixed asset investment right at the boundary of counties recently excluded from the FDB list. This suggests that after being removed from the FDB list, firms in these counties exhibit considerably higher fixed asset investment relative to those in neighboring counties.

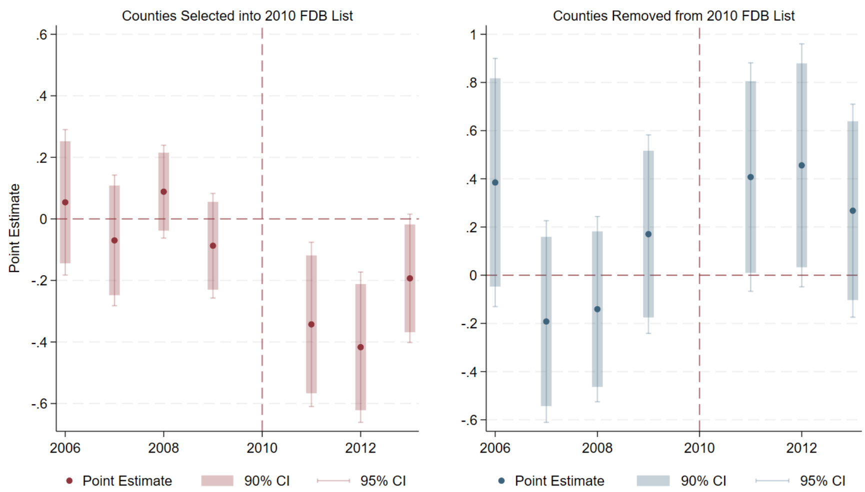
Following the work by He et al. (2020), we investigate the dynamics in fixed assets investment in Panel B of Figure 1.12. This SRD approach hinges on comparing firms located within FDB-designated areas to those in geographically adjacent but non-FDB counties. A critical assumption of SRD is the similarity in pre-treatment outcomes between neighboring FDB and non-FDB counties. For newly-selected FDB counties, we find that the fixed assets discontinuity was close to zero before 2010, but became significantly larger in 2011.[3] This negligible and insignificant effect prior to 2010 supports our foundational assumption: absent the FDB policy, manufacturing firms in FDB and non-FDB counties would have similar trends for fixed asset investment.

Table 1.7 quantifies the graphical evidence depicted in Figure 1.12, examining the impact of counties entering and exiting the FDB list. Panel A presents the SRD analysis without control variables. Columns (1) to (3) show that firms in counties newly included in the FDB list exhibit lower levels of fixed asset investment compared to firms in geographically adjacent counties. Conversely, columns (4) to (6) indicate that firms in counties recently removed from the FDB list demonstrate higher fixed asset investments than their counterparts in neighboring counties. To further validate our findings, we conduct robustness tests in Panel B, incorporating both county and industry fixed effects, and in Panel C, incorporating county-by-industry fixed effects. Panel B assesses differences in fixed asset investment across counties and industries, while Panel C provides a more detailed comparison by evaluating firms within the same industries but located in proximate geographical areas, thus eliminating potential industry-specific confounding factors. Our analyses yield significant results across Panels A, B, and C, with consistent effect sizes in Panels B and C. Additionally, the SRD estimates exhibit strong robustness across various kernel function selections. Findings from Panels B and C underscore the significant influence of the FDB policy on firms' investment decisions.

[3] Due to data availability, unfortunately, we can only track the impact to the year of 2013.



(a) Spatial Regression Discontinuity (Imbens and Wager 2019)



(b) Dynamic Spatial Regression Discontinuity

Figure 1.12: FDB v.s. Neighboring non-FDB Counties: Firm-Level Fixed Assets Investment

Note: (1) A positive distance indicates firms located within FDB counties, while a negative distance indicates firms located outside the border of FDB counties; (2) Industry and county fixed effects are absorbed before plotting the regression discontinuities; (3) FDB counties refer to those selected into the FDB list in 2010.

Table 1.7: Spatial Regression Discontinuity: Fixed Assets Gap

	ln(Gap in Fixed Assets Investment)					
	Selection into FDB List:			Removal from FDB List:		
	(1)	(2)	(3)	(4)	(5)	(5)
Panel A: No Control						
RD	−0.403***	−0.315***	−0.368***	0.553***	0.593***	0.631***
	(0.100)	(0.111)	(0.126)	(0.146)	(0.147)	(0.149)
Bandwidth	4.387	3.707	2.863	4.751	4.435	3.894
Panel B: County FE + Industry FE Absorbed						
RD	−0.217***	−0.166**	−0.179*	0.279**	0.285**	0.257*
	(0.078)	(0.084)	(0.097)	(0.129)	(0.131)	(0.148)
Bandwidth	4.883	4.294	3.360	4.629	4.314	3.516
Panel C: County by Industry FE Absorbed						
RD	−0.190***	−0.203***	−0.197***	0.258**	0.271**	0.276**
	(0.065)	(0.071)	(0.077)	(0.124)	(0.124)	(0.127)
Bandwidth	5.933	5.155	4.189	4.659	4.405	3.834
N(obs)	46,044	46,044	46,044	16,759	16,759	16,759
Kernel	Triangle	Epanech	Uniform	Triangle	Epanech	Uniform

Note: (1) Each coefficient represents a separate RD regression; (2) The running variable is the distance between a firm and the border of a corresponding FDB county, where negative (positive) means firms are located outside (within) FDB counties; (3) Negative coefficients indicate a negative gap between newly selected FDB counties and neighboring counties, positive coefficients indicate a positive gap between newly delisted FDB counties and neighboring counties; (4) The discontinuities are estimated using local linear regressions and MSE-optimal bandwidth proposed by Calonico et al. (2014); (5) Standard errors are clustered at the county level.

Chapter 2

Assessing a Flood Management Policy - A Spatial General Equilibrium Framework

In this chapter, we develop a spatial general equilibrium model that captures trade linkages among FDB counties, protected cities, and other regions. By comparing the actual output to a counterfactual scenario without FDBs, we find that as FDBs absorb more floodwater, the policy's output gains increase; however, this comes at the cost of growing inequality between FDB counties and others. In summary, FDBs may improve economic resilience against floods, but the economic cost is taken disproportionately by rural counties.

2.1 Introduction

Chapter 1 presents empirical evidence on the economic cost of the FDB policy. Most importantly, Chapter 1 reveal two primary findings: 1. the FDB policy has effectively redistributed flood exposures across different regions; 2. the FDB designation has had a negative and persistent impact on the economic development of FDB counties. They further show that the firm-response mechanism is primary reason behind the negative impacts of the policy - firms are reluctant to enter and invest in FDB counties with higher flood risks.

Motivated by the empirical findings in Chapter 1, in this chapter we use a spatial general equilibrium model to quantify the net output gain brought by the FDB policy. We need a general equilibrium model for two reasons. First, it is difficult to empirically identify the impact of the FDB policy on protected cities, as these economically important cities are targeted by numerous policies, with the FDB policy being just one among many. Second, as indicated by Redding and Turner (2015) and Allen and Arkolakis (2022), infrastructure investments (e.g., dams) could reshape the spatial distribution of economic activity and have general equilibrium effects. We need a general equilibrium model to analyze the impact of changing flood water flow on the broader region, so that we can quantify spillover effects and understand whether other counties benefit from protecting the manufacturing sector in economically important cities. Following the approach of Fajgelbaum et al. (2019), manufacturing goods are assumed to be tradable across different regions. Firms of rational expectations make entry decisions prior to flood events. After calibrating the model to fit real-world data, we construct a counterfactual scenario in which FDB counties did not protect urban cities from floods. In this counterfactual scenario, without FDBs, flood risk in FDB counties (protected cities) would decrease (increase). Comparing the counterfactual output with the actual output, we find that as FDBs absorb more flood water, the net output gain would be higher, although the inequality between FDB counties and urban cities would be exacerbated.

Based on the general equilibrium framework, we also conduct another counterfactual practice, in which FDB counties of different productivity levels would be removed from the list, successively. We find that (i) higher-productivity counties contribute minimally to overall output gains; and (ii) lower-productivity and more economically vulnerable counties contribute significantly to output gains but experience greater flood exposure. These findings imply two key policy considerations. First, the Chinese government may be overprotecting urban cities, as similar output gains could be realized by excluding higher-productivity counties from the FDB list. Second, a more equitable compensation scheme that transfers surplus from protected urban areas to FDB counties could significantly improve social equity.

2.2 Spatial General Equilibrium Model to Quantify the Net Output Gain

To quantify the net output gain from the FDB policy, we develop a spatial general equilibrium model where manufacturing firms enter the market and capital owners make optimal investment decisions, both based on their rational expectations of flood risk, and flood risk may also change in response to FDB policies. This general equilibrium framework allows us to systematically analyze how the FDB policy protects urban areas, compare the magnitude of output loss in FDB-treated counties with the output gain in FDB-protected counties, and account for spillover effects and trade flows between different counties.

2.2.1 Model Purpose

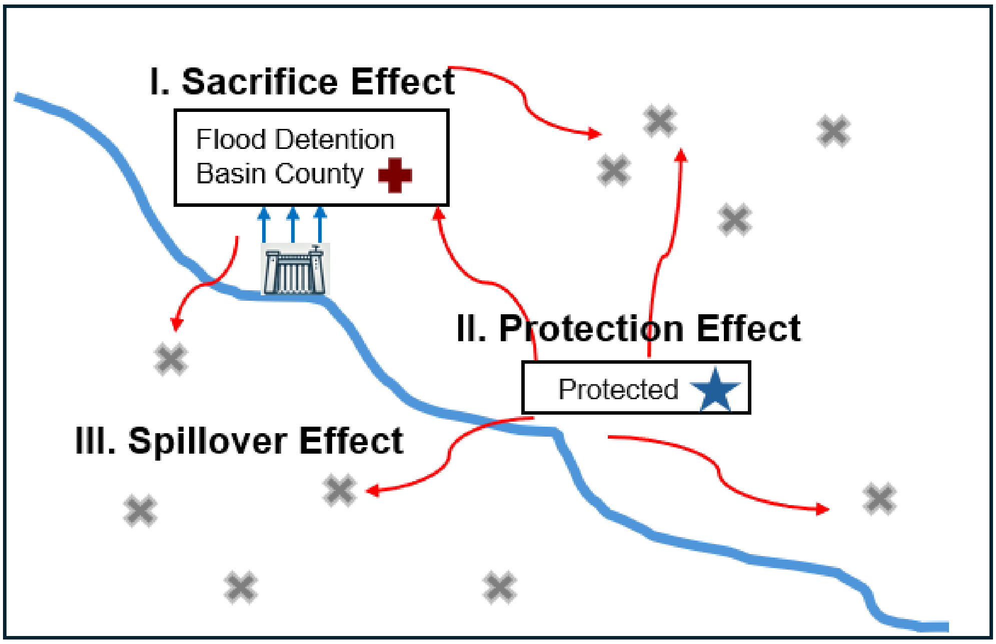
To quantify the aggregate impact of the Flood Detention Basin (FDB) policy on a broader region, we develop a spatial general equilibrium model. This approach is necessary because reduced-form estimations cannot fully capture the aggregate effects of the FDB policy. These effects can be decomposed into three components: (i) the sacrifice effect, (ii) the protection effect, and (iii) the spillover effect.

The sacrifice effect represents the economic costs on FDB counties due to the policy design, or the extent of economic sacrifices made by these counties. Using a difference-in-differences approach, Chapter 1 estimate that nighttime light intensity decreases by approximately 10% in counties selected into the FDB list. The protection effect refers to the benefits urban areas receive from being protected from floods. This effect has two components: the direct protection effect and the indirect protection effect.

The direct protection effect occurs during severe flood events when floodwaters are diverted to FDB counties, thereby reducing damage in protected urban areas. Reduced-form analysis shows that compared to the control group, flood damage in protected counties decreases by around 10%, while flood damage in FDB counties increases by approximately 18%. These findings confirm that FDB-protected counties experience significant direct protection during floods. The indirect protection effect, however, generates from reduced flood risk in protected counties during normal (non-flood) periods. This reduced risk makes these counties more attractive to firms, leading to increased economic activity even outside flood events. Unlike the direct effect, the indirect protection effect cannot be easily estimated through reduced-form approaches because protected counties benefit from various policies, making it difficult to isolate the FDB policy's contribution. Thus, our general equilibrium model is essential to capturing this indirect effect.

Finally, the spillover effect captures the broader regional benefits from trade linkages. Urban areas that gain from the FDB policy can increase their manufacturing output, indirectly benefiting other regions through trade. For instance, higher production in urban areas leads to increased consumption of their goods in neighboring counties. Like the indirect protection effect, this spillover effect is difficult to estimate using reduced-form methods alone. Hence, we also need a general equilibrium framework to evaluate the spillover effect.

Figure 2.1: Three Effects of FDB Policy



2.2.2 Model Environment and Equilibrium Conditions

Model Framework

Consider an economy with N regions, each region $n \in N$ has one representative capital owner who cannot move across regions and makes optimal investment decisions to determine the amount of capital to be used for production. Before the flood events s_j occur, capital owners in each region anticipate future flood risks and decide their optimal investment $a_{n,t+1}$ for the next period. This enables us to capture the mechanism by which higher flood risk in a region leads to reduced investment. The consumption goods in this economy include agricultural goods, manufacturing goods, and service goods. Agricultural and service goods are not tradable, while manufactured goods are tradable (Fajgelbaum et al. 2019) and subject to an iceberg trade cost, d_{ni}, which represents the cost of shipping one unit of goods from region n to destination region i. Firms hire workers to produce goods, and we assume workers are hand-to-mouth and cannot migrate across regions, consistent with our empirical evidence showing no significant migration. Before the realization of the flood event s_j, in each region, manufacturing firms[1] anticipate future flood risks and can decide to enter the market, subject to an entry cost. When firms expect to see a higher future flood risk, they will choose to not enter the market, leading to a reduced number of manufacturing firms. After the flood realization, workers and capital owners choose optimal consumption bundles, and firms maximize their profits accordingly. We will elaborate each agent's decision in detail in the following sections.

Floods

We assume that at every time t, a flood event s_t^j is determined by nature, and some regions may be flooded while others may not (it could also be the case that no regions are flooded, leading to an event with no flooding). Therefore, a flood event $s_t^j = \{f_{1,t}^j, f_{2,t}^j, \ldots, f_{N,t}^j\}$ is a vector of zeros and ones, where zero indicates no flood and one indicates being flooded. Each element $f_{n,t}^j$ describes whether region n is flooded

[1] For simplicity, we assume a single aggregate agricultural sector and a single aggregate service sector, without explicitly modeling potential firm entry and exit in these sectors.

(=1) or not (=0) at time t in event j[2]. We define $S = \{s_t^1, s_t^2, \ldots, s_t^j\}$ as the set of all possible flood events, with each flood event occurring with a probability $pr(s_t^j)$[3].

We assume that, in a flood event s_t^j, if region n is flooded, the flooding will negatively affect the productivity of local manufacturing firms. We model the flood-contingent productivity $z_n^M(s_t^j)$ as:

$$z_n^M(s_t^j) = \bar{z}_n^M exp(-\varepsilon_M f_{n,t}^j) \tag{2.1}$$

where $\bar{z}_n^M$ denotes the region-specific productivity during non-flooding times $f_{n,t}^j = 0$, and ε_M denotes the percentage productivity loss when a region is flooded $f_{n,t}^j = 1$. At any time t, only one specific type of flood event can occur; hence, we suppress the event subscript j, and we will use s_t instead of s_t^j in the following sections.

Workers

In each region n, there is a unit mass of hand-to-mouth workers L_n, who are immobile across regions[4]. Workers supply one unit of labor inelastically in the region where they live. After observing the flood event s_t, workers choose their consumption on $C_n^{w,A}(s_t)$ (agricultural goods), $C_n^{w,M}(s_t)$ (manufacturing goods), and $C_n^{w,S}(s_t)$ (service goods) to maximize their utility, subject to the budget constraint.

$$\begin{aligned} \max_{\{C_n^{w,A}(s_t), C_n^{w,M}(s_t), C_n^{w,S}(s_t)\}} \quad & U(C_n^{w,A}(s_t), C_n^{w,M}(s_t), C_n^{w,S}(s_t)) \\ s.t. \quad & P_n^A(s_t)C_n^{w,A}(s_t) + P_n^M(s_t)C_n^{w,M}(s_t) + P_n^S(s_t)C_n^{w,S}(s_t) = w_n(s_t) \end{aligned} \tag{2.2}$$

The utility $U(\cdot)$ takes a Cobb-Douglas form such that $U(\cdot) = \xi_A log(C_n^{w,A}(s_t)) + (1 - \xi_A - \xi_S) log(C_n^{w,M}(s_t)) + \xi_S log(C_n^{w,S}(s_t))$ where ξ_A is the share of income spent on agricultural goods, ξ_S is the share of income spent on service goods, and $1 - \xi_A - \xi_S$ is the share of income spent on manufacturing goods. $w_n(s_t)$ is the wage rate in region n,

---

[2]If no regions are flooded, the vector will consist entirely of zeros.

[3]In theory, the cardinality of the set is 2^N. However, many flood events are naturally impossible. For example, it is unlikely to have floods in regions located in deserts. Therefore, in the calibration and counterfactual sections, we only consider flood events observed in historical data.

[4]Without loss of generality, we normalize the population such that $\sum_{n=1}^{N} L_n = 1$

and $P_n^A(s_t)$, $P_n^M(s_t)$, and $P_n^S(s_t)$ represent the prices of agricultural goods, manufacturing goods, and service goods, respectively, in region n. All of wage w_n, price P_n, and consumption C_n^w are contingent on flood event s_t because, in different flood events, the equilibrium wage, prices, and people's optimal consumption may change in response to flood shocks.

Capital Owners

During time period t, capital owners in region n decide how much to invest for the next period, $a_{n,t+1}$, before the realization of the flood event s_t. Hence, the asset position decision is independent of s_t, capturing the fact that investment only respond to long-term flood risk changes and is irrelevant to whether a flood occurs in a given period.

$$\begin{aligned} V_n^o(a_{n,t}) = & \max_{\{C_n^{o,A}(s_t), C_n^{o,M}(s_t), C_n^{o,S}(s_t), a_{n,t+1}\}} \mathbb{E}_{s_t} U(C_n^{o,A}(s_t), C_n^{o,M}(s_t), C_n^{o,S}(s_t)) + \beta V_n^o(a_{n,t+1}) \\ & s.t. \quad P_n^A(s_t)C_n^{o,A}(s_t) + P_n^M(s_t)C_n^{o,M}(s_t) + P_n^S(s_t)C_n^{o,S}(s_t) + a_{n,t+1} = (1+r(s_t))a_{n,t} + I_{n,t}\pi_n(s_t) \end{aligned} \tag{2.3}$$

The income of capital owners come from two sources. On the one hand, they get their return from the last period investment $(1+r(s_t))a_{n,t}$, where $r(s_t)$ is the national interest rate. One the other hand, capital owners obtains all the profits of manufacturing firms $I_{n,t}\pi_n(s_t)$, where $I_{n,t}$ is the number of manufacturing firms and $\pi_n(s_t)$ is the the average profit of manufacturing firms in region n. After the realization of the flood event s_t, capital owners optimize their consumption bundles subject to the budget constraint, and their preferences are identical to those of the workers, such that $U(\cdot) = \xi_A log(C_n^{o,A}(s_t)) + (1-\xi_A-\xi_S) log(C_n^{o,M}(s_t)) + \xi_S log(C_n^{o,S}(s_t))$.

Production

In this economy, there are three sectors producing distinct consumption goods: agriculture, manufacturing, and services. These sectors produce agricultural goods $Y_n^A(s_t)$, manufacturing goods $Y_n^M(s_t)$, and service goods $Y_n^S(s_t)$, respectively. The agricultural sector uses labor $l_n^A(s_t)$ as the only input, supplying non-tradable agricultural goods with linear production technology to the local market. It operates in a perfectly competitive way, so the price of agricultural goods equals the local wage. The profit maximization problem for the agricultural sector during flood event s_t is given by:

$$\begin{aligned} \max_{\{l_n^A(s_t)\}} \quad & P_n^A(s_t)Y_n^A(s_t) - w_n(s_t)l_n^A(s_t) \\ s.t. \quad & Y_n^A(s_t) = z_n^A(s_t)l_n^A(s_t) \end{aligned} \tag{2.4}$$

We assume the service sectors also supply non-tradable goods in the local market in a perfectly competitive way. However, unlike the agricultural sector, the service sectors use both labor $l_n^S(s_t)$ and capital $k_n^S(s_t)$ in a Cobb-Douglas production technology, with the factor share of labor denoted by α. The maximization problem for the service sector is given by:

$$\begin{aligned} \max_{\{l_n^S(s_t),k_n^S(s_t)\}} \quad & P_n^S(s_t)Y_n^S(s_t) - w_n(s_t)l_n^S(s_t) - r_n(s_t)k_n^S(s_t) \\ s.t. \quad & Y_n^S(s_t) = z_n^S(s_t)l_n^S(s_t)^\alpha k_n^S(s_t)^{1-\alpha} \end{aligned} \tag{2.5}$$

The manufacturing sector is the key focus of this paper, and therefore, we model this sector in greater detail to better capture the mechanisms identified in the empirical results. Firstly, we describe the demand for manufacturing goods and model consumers in region n as consuming a variety of manufacturing goods produced by heterogeneous firms from different regions, using a CES aggregator:

$$Y_n^M(s_t) = \left[\sum_{i=1}^N I_{i,t} y_{in}^M(s_t)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{2.6}$$

where σ measures the elasticity of substitution across manufacturing goods produced by different regions, $y_{in}^M(s_t)$ is the quantity of manufacturing good produced in region i and sold to region n, and $I_{i,t}$ is the number of manufacturing firms in region i. Denote $P_{in}^M(s_t)$ as the price of manufacturing goods produced by region i and sold to region n. Then, one can easily show that the price index of manufacturing goods sold in region n is $P_n^M(s_t) = \left[\sum_{i=1}^N I_{i,t} P_{in}^M(s_t)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$.

On the supply side, firms in region n hire labor $l_{ni}^M(s_t)$ and capital $k_{ni}^M(s_t)$ to produce manufacturing goods y_{ni} using a Cobb-Douglas technology with productivity $z_n^M(s_t)$.

Manufacturing goods can be traded from region n to region i, subject to the iceberg cost d_{ni}, meaning that to ship one unit of goods, firms need to produce d_{ni} unit. The profit for each firm in region n is given by:

$$\begin{aligned}\pi_n(s_t) = \max_{\{l_{ni}^M(s_t), k_{ni}^M(s_t)\}_{i=1}^N} & \sum_{i=1}^N \left[P_{ni}^M(s_t) y_{ni}^M(s_t) - w_n(s_t) l_{ni}^M(s_t) - r(s_t) k_{ni}^M(s_t) \right] \\ s.t. \qquad & d_{ni} y_{ni}^M(s_t) = z_n^M(s_t) l_{ni}^M(s_t)^\alpha k_{ni}^M(s_t)^{1-\alpha} \qquad \forall i\end{aligned} \tag{2.7}$$

To operate and earn profit $\pi_n(s_t)$ at time t, manufacturing firms must first decide whether to enter the market before the realization of the flood event s_t. We also assume there is a probability η that the manufacturing firm will exit the market in the next period. Therefore, the value of a manufacturing firm in region n is the expected profit in period t plus the discounted value (with discount rate β) of the firm in the next period, conditional on survival:

$$V_{n,t}^s = \mathbb{E}_{s_t} \pi_n(s_t) + \beta(1-\eta) V_{n,t+1}^s \tag{2.8}$$

The free entry condition requires that the value of manufacturing firms should equal to the entry cost c_n^s.

$$V_{n,t}^s = c_n^s \tag{2.9}$$

Market Clearing Conditions

There are three sets of market clearing conditions.

1. National capital market: The flood-event-specific interest rate $r(s_t)$ require asset positions equal flood-event-specific capital demands in all regions:

$$\sum_{n=1}^N I_n \sum_{i=1}^N k_{ni}^M(s_t) + \sum_{n=1}^N k_n^S(s_t) = \sum_{n=1}^N a_{n,t} \tag{2.10}$$

2. Local labor markets: The flood-event-specific wage rates $w_n(s_t)$ require labor

supply equal flood-event-specific labor demands in all regions:

$$l_n^A(s_t) + \sum_{i=1}^{N} l_{ni}^M(s_t) + l_n^S(s_t) = L_n \qquad \forall n \tag{2.11}$$

3. Local final good markets: The final good markets are assumed to be perfectly competitive, so prices $P_n^A(s_t)$, $P_{ni}^M(s_t)$ and $P_n^S(s_t)$ satisfy that the final good demands and supplies are equalized in all regions:

$$L_n C_n^{w,A}(s_t) + C_n^{o,A}(s_t) = Y_n^A(s_t) \qquad \forall n \tag{2.12}$$

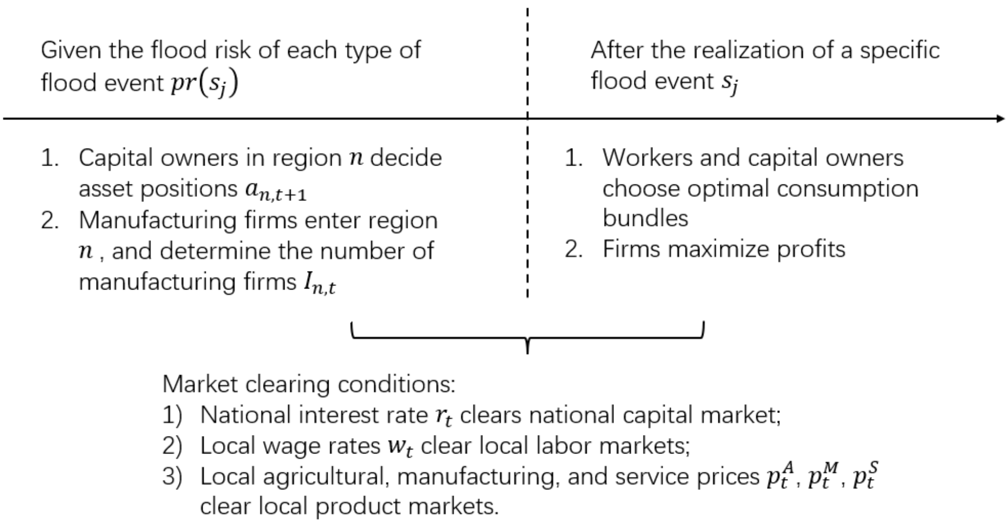
$$L_n C_n^{w,S}(s_t) + C_n^{o,S}(s_t) = Y_n^S(s_t) \qquad \forall n \tag{2.13}$$

$$P_{ni}^M(s_t) = \left[L_i C_i^{w,M}(s_t) + C_i^{o,M}(s_t)\right]^{\frac{1}{\sigma}} P_i^M(s_t) y_{ni}^M(s_t)^{-\frac{1}{\sigma}} \qquad \forall i,n \tag{2.14}$$

$$P_n^M(s_t) = \left[\sum_{i=1}^{N} I_{i,t} P_{in}^M(s_t)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \qquad \forall n \tag{2.15}$$

Model Timeline

The figure below provides an illustration of the model's timeline. It shows the sequence of events and decisions made by capital owners, manufacturing firms, and workers, both before and after the realization of a specific flood event. It also outlines the market clearing conditions for national capital market, local labor markets, and local product markets.


Given the flood risk of each type of flood event $pr(s_j)$
After the realization of a specific flood event s_j
1. Capital owners in region n decide asset positions $a_{n,t+1}$
2. Manufacturing firms enter region n , and determine the number of manufacturing firms $I_{n,t}$
1. Workers and capital owners choose optimal consumption bundles
2. Firms maximize profits
Market clearing conditions:
1) National interest rate r_t clears national capital market;
2) Local wage rates w_t clear local labor markets;
3) Local agricultural, manufacturing, and service prices p_t^A, p_t^M, p_t^S clear local product markets.

Equilibrium

The spatial general equilibrium consists of capital owners' asset positions $\{a_{n,t}\}$ and consumption bundles $\{C_n^{o,A}(s_t),\ C_n^{o,M}(s_t),\ C_n^{o,S}(s_t)\}$, workers' consumption bundles $\{C_n^{w,A}(s_t),\ C_n^{w,M}(s_t),\ C_n^{w,S}(s_t)\}$, sector-specific factor demands and outputs $\{l_n^A(s_t),\ l_n^M(s_t),\ l_n^S(s_t),\ k_n^M(s_t),\ k_n^S(s_t),\ Y_n^A(s_t),\ Y_n^M(s_t),\ Y_n^S(s_t)\}$, manufacturing firms counts $\{I_{n,t}\}$, and prices $\{w_n(s_t),\ r(s_t),\ P_n^A(s_t),\ P_n^M(s_t),\ P_n^S(s_t)\}$, such that given the distribution of workers $\{L_n\}$

1. Before the realization of flood events s_t
   - (i) $\{a_{n,t}\}$ satisfy capital owners' optimal investment decisions in Equation 2.3;
   - (ii) $\{I_{n,t}\}$ satisfy the free entry condition in Equation 2.9;
2. After the realization of flood event s_t
   - (i) $\{C_n^{o,A}(s_t), C_n^{o,M}(s_t), C_n^{o,S}(s_t)\}$ and $\{C_n^{w,A}(s_t), C_n^{w,M}(s_t), C_n^{w,S}(s_t)\}$ satisfy capital owners' and workers' utility maximization problems in Equation 2.2 and 2.3;
   - (ii) $\{l_n^A(s_t),\ l_n^M(s_t),\ l_n^S(s_t),\ k_n^M(s_t),\ k_n^S(s_t),\ Y_n^A(s_t),\ Y_n^M(s_t),\ Y_n^S(s_t)\}$ satisfy sectors' profit maximization problems in Equation 2.4, 2.5, and 2.7;
   - (iii) $\{w_n(s_t), r(s_t), P_n^A(s_t), P_n^M(s_t), P_n^S(s_t)\}$ clear the factor and product markets in Equation 2.10 - 2.15.

2.2.3 Calibration and Simulation

In this section, we calibrate our model to match Chinese counties in Huai River Basin, the basin with the highest river flood risk, between 2000 and 2010.

Exogenously Calibrated Parameters

Panel A of Table 2.1 shows parameter values obtained directly from literature and data. We treat each region as a county, and there are $N = 176$ counties in Huai River Area. We standardize labor force $\bar{L}$ to be 1. Following previous literature (Head et al.

2014 & Jia et al. 2022), we set the elasticity of substitution across varieties, σ, as 5. We choose a discount factor, β, to be 0.95 to generate an aggregate steady-state interest of 5%. We further match the shares of sector-specific consumption with the real data provided by 2000-2010 Chinese National Bureau of Statistics. To be specific, the share of agricultural consumption, ξ_A, is 11.7%, and the share of service consumption, ξ_S, is 42.2%. We choose a factor share of capital, $1-\alpha$, to be 0.5 for both the manufacturing and service industry. This is consistent of the national-level sector specific factor share in China, calculated by Chinese input and output tables and national accounts, sourced from Chinese National Bureau of Statistics.

Transportation Cost - The calculation of transportation costs, d_{ni}, is based on geodesic distances across different counties. For the transportation cost within a county, we adopt a similar approach as existing literature (e.g., Redding and Venables 2004, Au and Henderson 2006, and Balboni 2019). Specifically, we calibrated trade costs by approximating intra-unit trade costs based on the average distance traveled to the center of a circular unit of the same area from evenly distributed points, given by $\frac{2}{3}(\text{area}/\pi)^{1/2}$. We standardize the smallest transportation costs to be 1.

Probability of Each Flood Type - In 2000 and 2010, there were 5 major floods in Huai River Basin, which happened in 2002, 2003, 2005, 2007, and 2010, respectively. The list of counties being affected is different across different events. For example, the 2003 flood caused damages to 61 counties out of 176 counties in Huai river, while the 2010 flood caused damages to 25 counties. Based on the level of precipitation, we divide the monthly-averaged precipitation during flood seasons (June to September) into two categories: (i) $<$ 120 mm; (ii) $>$ 120 mm. We then calculated the region-specific flooding probability based on both historical data on monthly precipitation and actual flood event.

Productivity Loss - We estimate productivity loss in agriculture sector, manufacturing sector and service sector based on the estimation below.

$$Y_{ict} = \alpha + \beta FloodExposure_{ict} + \gamma_t + \lambda_c + \eta_t + \varepsilon_{ict}$$

In this estimation, Y_{ict} represents the average productivity in county i, city c

and year t, which is measured as the ratio of output per worker in an industry. $FloodExposure_{icjt}$ indicates the size-adjusted flood exposure, which is the average days of flood[5] in county i in year t. λ_c and η_t represent city and time fixed effects. Standard errors are clustered at city level. Reduced form results suggest that when the average days of flood in county increases by one day[6], then the productivity in manufacturing sector would decrease by 5.9%.

Internally Calibrated Parameters

In Panel B of Table 2.1, we calibrate the flood-free productivity of agriculture, manufacturing and service industry in different counties, to match county-level data on real outputs and labor force share in different sectors. Although we estimate all parameters jointly, we can pinpoint which parameter influences a specific outcome. For instance, sector-specific real outputs at the county level are influenced by sector-specific productivity, while regional amenities are determined by the labor force in each area. To maintain consistency, we standardize the total national GDP and population to 1 in our baseline calibration, as these factors do not impact our baseline calibration.

[5] the flood data is further processed by excluding permanent water pixels

[6] Note: insert the expression for the flood days to explain what does one flood day mean

Table 2.1: Calibration Targets

Parameter	Numbers	Value	Source/Targeted Moments
Panel A: Exogenously Calibrated Parameters			**Source:**
N - Number of regions	1	176	Number of counties in Huai River Basin
$\bar{L}$ - Labour force	1	1	Standardized to 1
σ - Elasticity of substitution across varieties	1	5	Head et al. (2014)
β - Discount factor	1	0.95	Steady-state interest of 5%
ξ_A - Share of agricultural consumption	1	0.117	Chinese National Bureau of Statistics
ξ_S - Share of service consumption	1	0.422	Chinese National Bureau of Statistics
$pr(s_t)$ - Flooding event probability	7	0.12(0.21)	Precipitation and flood event (2000-2009)
d_{ni} - Transportation costs	N^2	1.23(0.04)	Geodesic distances
$1-\alpha$ - Factor share of capital	1	0.5	Factor shares of manufacturing and service industries
ε_M - Productivity loss when flooded	1	-0.059	Estimation (Table A2)
Panel B: Internally Calibrated Parameters			**Targeted Moments:**
$\bar{z}_n^A$ - County-level agriculture productivity	N	0.83(0.34)	County-level agriculture outputs
$\bar{z}_n^M$ - County-level manufacturing productivity	N	0.29(0.12)	County-level manufacturing outputs
$\bar{z}_n^S$ - County-level service productivity	N	0.21(0.22)	County-level service outputs
B_n - Local amenity	N	5.05(0.23)	County-level labor force share

Note: for flooding event probability, transportation costs, internally calibrated productivity and local amenity, the value in the table indicates the average value across all regions, and the standard error is in the parenthesis

Table 2.2: Comparison of Actual and Model-generated Regression Results

	Actual Data:	Model Simulation:
(in logarithm)	Fixed Assets/Worker (1)	Capital/Worker (2)
FDB	-0.197*** (0.077)	-0.175*** (0.036)
N(obs)	46,044	1,936

Note: (1) Column 1 is extracted from Column (3) in Panel C of our regression discontinuity regression in Table Chapter 1; (2) Column 2 is based on our model prediction; (3) The consistency between those two estimates indicate that our model can well predict the fixed assets per worker.

2.2.4 Model Prediction

In this section, we conduct a comparative analysis to illustrate the consistency between the empirical findings and the predictions of our general equilibrium model. Our objective is to validate the model's capability to accurately reflect the reality of FDB counties, demonstrating its robustness and reliability as a tool for simulating real-world economic scenarios. Column 1 in Table 2.2 reports the regression result we gained in Chapter 1, while Column 2 reports the result we gain based on model simulation. The magnitudes do not differ significantly, and each of them falls within the other's 95% confidence interval, indicating that our model closely matches even the non-targeted moments and achieves a good fit.

2.2.5 Counterfactual Practice 1: FDB-Induced Net Output Gain

In this section, we quantify three different effects: (1) *the sacrifice effect*, representing the cost incurred by FDB counties due to the FDB policy, which we can compare to our reduced-form results; (2) *the protection effect*, capturing the benefits gained by FDB-protected counties from the FDB policy; and (3) *the total output effect*, reflecting the net output gain for the economy as a result of the FDB policy. In the counterfactual scenario, where the FDB policy is absent and FDB counties no longer protect urban cities, flood exposure in FDB counties would decrease, while flood exposure in protected areas would increase. Therefore, an important parameter for constructing the counterfactual scenario is the flood redistribution rate between FDB counties and FDB-protected counties.

Constructing the Counterfactual Practice

We construct the counterfactual scenario, in which FDB counties do not protect urban cities, based on the following steps. First, we calculate the total flood size in each county by aggregating flooded areas (pixels) over flooded days (duration) in each year between 2000 and 2010, indicating the total amount of floodwater in each county. Second, in the counterfactual scenario without the FDB policy, 45% (as estimated from the hydrological analysis in Section Chapter 1) of the floodwater in the current FDB counties is equally redistributed to the currently protected counties. This process allows us to construct a set of counterfactual flood events, $S' = \{s'_1, s'_2, \ldots, s'_J\}$, reflecting the counterfactual distribution of flood risk. In the third step, we translate the changes in flood exposure into changes in manufacturing output. Specifically, under the counterfactual scenario, flood damage would increase in protected urban cities while decreasing in FDB counties compared to the baseline case. Figure 2.2 provides a mind map illustrating how we construct the counterfactual scenario.

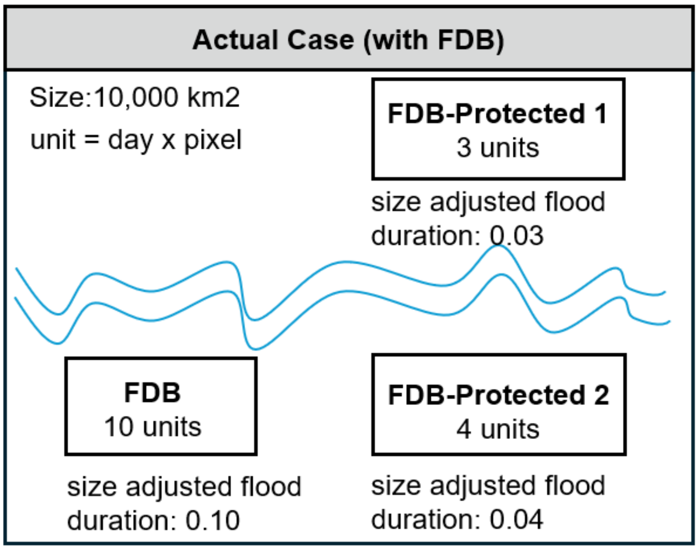
Sacrifice Effect

In Table 2.3, we quantify the sacrifice effect on FDB counties by collecting β_{FDB} in the calibrated case and the counterfactual case from running the regression

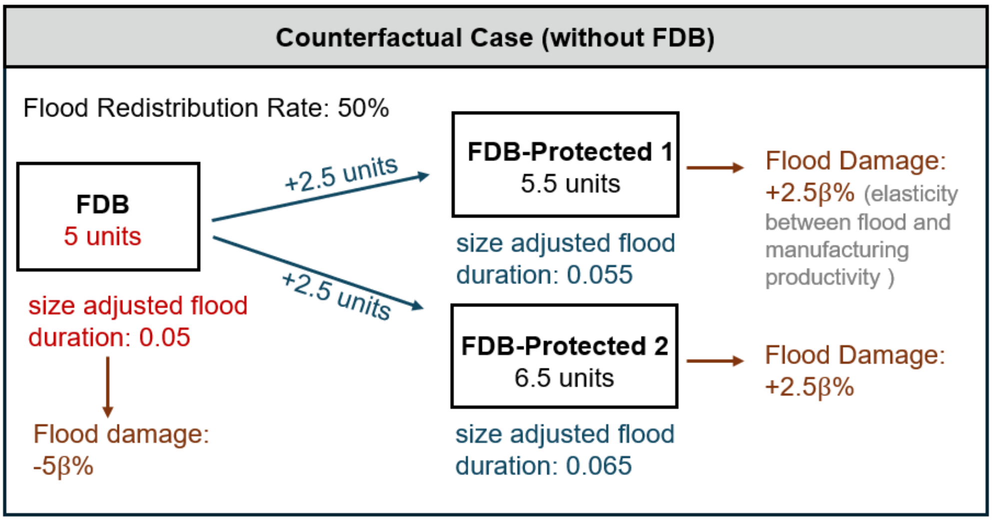
$$lnY_{icpt} = \alpha + \beta_{FDB}\text{FDB}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_c$$

where FDB_{icpt} is a dummy variable that equals 1 if the county i in city c, province p, at time t, is an FDB-county, and 0 if not. γ_{pt} is province-year fixed effect, η_t is time fixed effect, and λ_c is city fixed effect. ε_c is the standard error, which is clustered at the city

Figure 2.2: Mind Map: Constructing the Counterfactual Scenario



(a) Actual Case: With Flood Detention Basin



(b) Counterfactual Case: Without Flood Detention Basins level.

Column 3 reports the magnitude of change in β_{FDB} in the calibrated case and counterfactual case (flood exposure redistribution rate: 45%). We compare the results on total output with the result presented in Table Chapter 1. As shown in Chapter 1, the average treatment effect of FDB policy on nighttime light in FDB counties is around -10%. According to the work of Henderson et al. (2012) on estimating the elasticity between light and GDP, we can then translate this impact to around -3%, which is consistent with the result presented in Column 3 of Table 2.3. This consistency further validates our methods of constructing the counterfactual scenario.

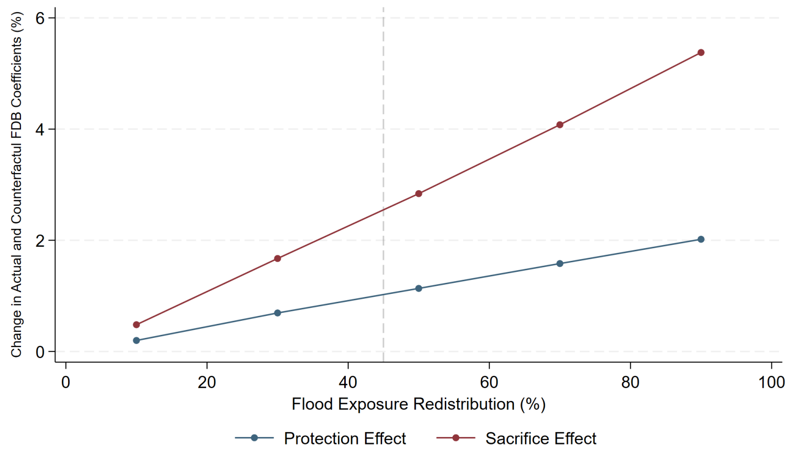
Table 2.3 then helps us to overcome the limitation of data availability and provides us with more results on the sacrifice effect. We find that the manufacturing output, total capital, manufacturing capital, share of manufacturing labor, and wage decreases by 9.62%, 5.11%, 8.49%, 10.86%, and 3.76%, respectively, because of the policy given a flood exposure redistribution rate of 45%. More results on sacrifice effect of different flood exposure redistribution rates are presented in Figure 2.3.

Protection Effect

When examining the impact of the FDB policy on FDB-protected counties, we divide the protection effect to two main sources: (1) a *direct protection effect*, where protected counties experience less damage during flood events; and (2) an *indirect protection effect*, where protected counties benefit from a decreased flood risk. We find that a protected county tends to suffer approximately 10% less damage when hit by floods, while an FDB county tends to suffer around 18% more. This finding indicates that FDB-protected counties are indeed *directly* protected during flood events. However, in our general equilibrium framework, we focus more on the *indirect* protection effect, whereby reduced flood risk encourages firms to enter and invest in these protected counties. Consequently, compared to the counterfactual scenario in which FDB counties do not protect urban cities, manufacturing output in these protected urban areas is higher in reality.

To understand the magnitude of protection effect, in Table 2.4, we quantify the total protection effect on FDB-protected counties by collecting $\beta_{Protected}$ in the calibrated

Figure 2.3: Sacrifice Effect and Protection Effect



case and the counterfactual case from running the regression

$$lnY_{icpt} = \alpha + \beta_{Protected} * \text{Protected}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_c$$

where FDB_{icpt} is a dummy variable that equals 1 if the county i in city c, province p, at time t, is an FDB-protected county, and 0 if not. γ_{pt} is province-year fixed effect, η_t is time fixed effect, and λ_c is city fixed effect. ε_c is the standard error, which is clustered at the city level.

Table 2.4 presents the results on the protection effect. We find that, if we assume that the flood exposure redistribution rate at 45%, then the FDB policy would lead to an increase in total output, manufacturing output, total capital, manufacturing capital, share of manufacturing labor, and wages by 1.74%, 3.92%, 2.51%, 3.30%, 4.40%, and 4.17%, respectively. Additional results on the protection effect across different flood exposure redistribution rates are shown in Figure 2.3

Net Output Gains of the FDB Policy

Finally, in Table 2.5, we quantify the net output gain brought by the FDB policy by comparing the total output in the calibrated case and the counterfactual scenario. Overall, we find a 0.06% increase in total output due to the FDB policy, which equates to an annual net increase in output of around US$3billion in Huai River Basin. According-

Table 2.3: Quantification of Sacrifice Effect (Actual v.s. Counterfactual)

	β_{FDB} :		
	A.Calibration	B.Counterfactual	$\|$Diff/A$\|$(%)
	(1)	(2)	(3)
Output: Total	−0.030***	−0.029***	3.46%
Output: Manufacturing	−0.468***	−0.427***	9.62%
Capital: Total	−0.251***	−0.239***	5.11%
Capital: Manufacturing	−0.373***	−0.344***	8.49%
Share of Manufacturing Labor	−0.052***	−0.048***	10.86%
Wage	−0.373***	−0.359***	3.76%

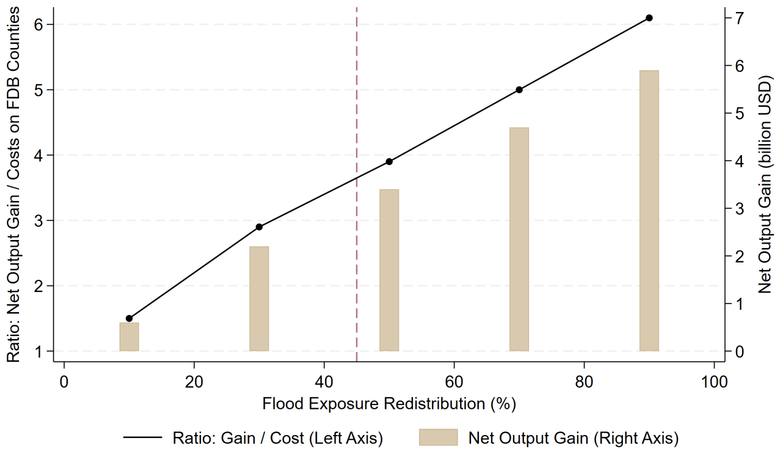
Note: (1) In the counterfactual case, we redistribute 45% of the flood risk to FDB-protected counties; (2) We collect β_{FDB} from running the regression $ln(Output)_{icpt} = \alpha + \beta_{FDB} * \text{FDB}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_{icpt}$, where FDB_{icpt} is a dummy that equals 1 if the county is an FDB-county, and 0 if not, γ_{pt} is province-year fixed effect, η_t is time fixed effect, and λ_c is city fixed effect; (3) The '|Diff/A|(%)' can be interpreted as the 'sacrifice effect', which is the impact of FDB policy on different outcomes in FDB counties.

Table 2.4: Quantification of Protection Effect (Actual v.s. Counterfactual)

	$\beta_{Protected}$:		
	A.Calibration	B.Counterfactual	$\|$Diff/A$\|$(%)
	(1)	(2)	(3)
Output: Total	0.983***	0.967***	1.74%
Output: Manufacturing	1.304***	1.255***	3.92%
Capital: Total	0.750***	0.732***	2.51%
Capital: Manufacturing	1.044***	1.011***	3.30%
Share of Manufacturing Labor	0.138***	0.132***	4.40%
Wage	0.544***	0.522***	4.17%

Note: (1) In the counterfactual case, we redistribute 50% of the flood risk to FDB-protected counties; (2) We collect $\beta_{Protected}$ from running the regression $ln(Output)_{icpt} = \alpha + \beta_{Protected} * \text{Protected}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_{icpt}$, where Protected_{icpt} is a dummy that equals 1 if the county is an FDB-protected county, and 0 if not; (3) The '|Diff/A|(%)' can be interpreted as the 'protection effect', which is the impact of FDB policy on different outcomes in FDB-protected counties.

Figure 2.4: Net Output Gains of the FDB Policy



ing to EM-DAT International Disaster Database, the average flood damage in China is around US$8billion every year in China. Hence, we believe that the FDB policy has substantially mitigated the economic threat posed by floods.

Figure 2.4 illustrates the benefit-to-cost ratio across various flood exposure redistribution rates. Our findings indicate that the benefit-to-cost ratio exceeds 1 at all redistribution rates, suggesting that intentionally flooding certain counties to protect urban areas results in a net gain in output. Moreover, as the redistribution rate increases, the benefit-to-cost ratio also rises, indicating that the net output gain from the policy increases as FDBs absorb more floodwater. However, as shown in Figure 2.3, the cost borne by FDB counties also intensifies with increased floodwater absorption. This highlights a tradeoff in policy design between mitigating flood risks and exacerbating inequality.

We also examine the potential policy implications under two future scenarios with increased flood damages, due to climate change. In these scenarios, we simulate a 50% and 100% increase in flood risk, in which the elasticity between flood and manufacturing productivity would increase by 50% and 100%, respectively. According to Table 2.5, under these projected conditions, the overall total output is expected to rise by 0.08% and 0.11%, respectively. The results also show that the sacrifice effect on FDB counties intensifies, with the gap reaching 4.79% and 5.84% under the 50% and 100% risk increase scenarios, respectively. Conversely, the protection effect for FDB-

Table 2.5: Total Output in Actual and Counterfactual Case

	Current Case:	Future Flood Risk Increases by:	
Actual - Counterfactual:		50%	100%
	(1)	(2)	(3)
Sacrifice Effect on FDB Counties ($\beta_{FDB} < 0$)			
$\Delta(\beta_{FDB})$	3.46%	4.79%	5.84%
Protection Effect on FDB-protected Counties ($\beta_{Protected} > 0$)			
$\Delta(\beta_{Protected})$	1.74%	2.51%	3.15%
Overall Economy:			
Δ(Total Output)	0.06%	0.08%	0.11%

Note: (1) We collect β_{FDB} from running the regression $ln(Output)_{icpt} = \alpha + \beta_{FDB} * \text{FDB}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_{icpt}$, where FDB_{icpt} is a dummy that equals 1 if the county is an FDB-county, and 0 if not, γ_{pt} is province-year fixed effect, η_t is time fixed effect, and λ_c is city fixed effect; (2) We collect $\beta_{Protected}$ from running the regression $ln(Output)_{icpt} = \alpha + \beta_{Protected} * \text{FDB-Protected}_{icpt} + \gamma_{pt} + \eta_t + \lambda_c + \varepsilon_{icpt}$, where $\text{FDB-Protected}_{icpt}$ is a dummy that equals 1 if the county is an FDB-protected county, and 0 if not; (3) The coefficient in Column (1) is the same as the coefficient in Column (3) in Table 2.3 and Table 2.4.

protected counties grows, with output gains of 2.51% and 3.15% in these scenarios. This counterfactual analysis indicates that as the severity of floods increases, FDBs would play an increasingly important role in managing flood damages. However, FDB counties would bear more costs because of the policy design.

2.2.6 Counterfactual Practice 2: Relative Contribution of Different FDB Counties

In the second counterfactual practice, we extend our discussion to think about whether the policy is optimal. It would be ideal for us to provide a list of counties that are most suitable for flood water detention. But we are not able to complete this task, in the current stage, because of hydrological challenges. The optimal design given economic criteria may not be feasible if we take geographical factors into account. Consider an extreme example. Under economic criteria, we may assign a county far away from river as an FDB county. Even we incorporate some geographical factors (e.g., elevation) into an economic model, the result may not be hydrologically feasible.

Despite of the challenge, the discussion on policy optimality is intrinsically important. We take a second-best approach by considering whether the government is over-protecting urban cities by designating too many FDB counties. In the first step, we rank FDB counties in terms of their exposure-standardized productivity, which is consistent with the proposition we have in Chapter 1. In the second step, we successively remove FDB counties of higher productivity from the FDB list and calculate the total output in each counterfactual scenario. In the third step, we calculate the relative contribution of each productivity group by comparing the counterfactual with the actual case.

In Figure 2.5, we present the net output gain of successively adding counties of higher productivity. Overall, we find that the net output gain increases as we add more counties to the list. However, according to Chapter 1, we find that the relative contribution is much higher in lower productivity groups than in higher productivity groups. County groups ranking 0-10%, 10-20%, 25-40%, and 40-50% in terms of productivity contribute more than 10%. Specifically, county group with a rank of 10-25% and 25-40% contribute the most to the net output gain, all above 25%. However, we find that the relative contribution of higher productivity group is low. County group ranking 75-80% and 85-100% contribute 0% and 3%, respectively.

On the one hand, we do not find counter-evidence to indicate that the inclusion of higher productivity counties is imposing negative effects on total outputs as the net output gain is increasing with the number of included FDB counties. On the other hand, however, the relative contribution of adding higher productivity counties is small. In terms of total outputs, it may be cost beneift efficient. However, if considering other non-monetary costs, then it may not be efficient because those counties may experience other costs that we are not able to measure in this study.

Overall, we suggest that the Chinese government is over protecting urban areas from floods by designating too many counties as FDB counties. Removing counties of higher productivity will not cause significant losses in output, but may save those counties from suffering both monetary and non-monetary costs.

Figure 2.5: Counterfactual Outputs with Different Numbers of FDBs

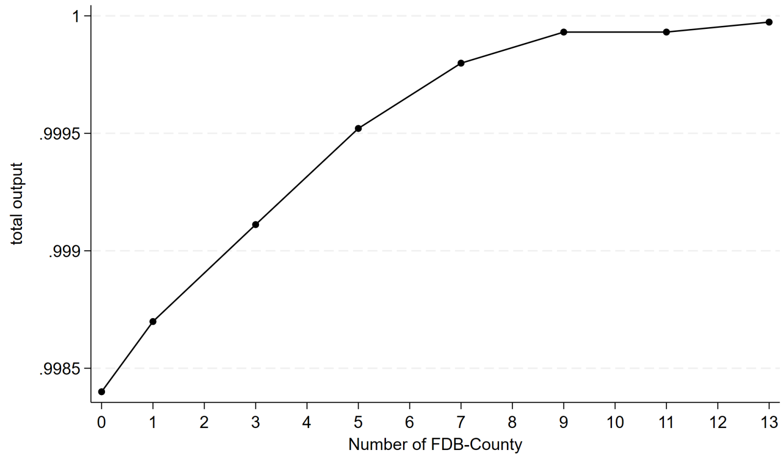
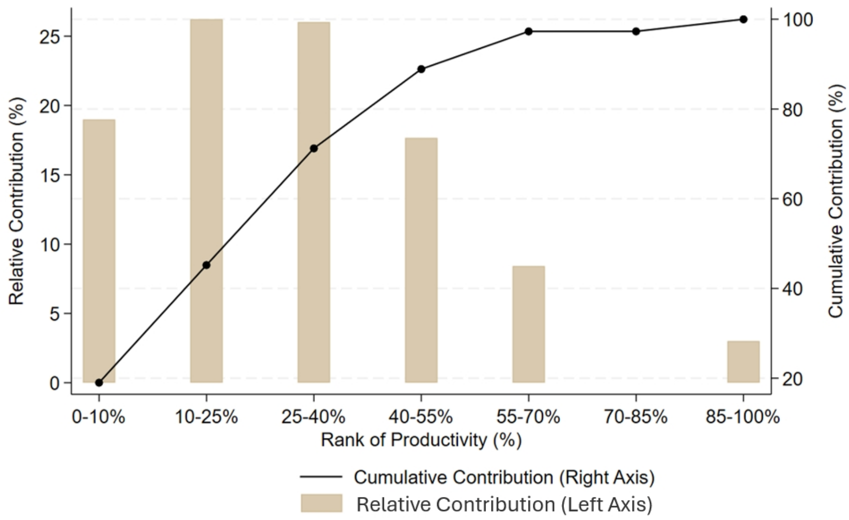


Figure 2.6: Relative Contribution of Different Productivity Groups



2.3 Conclusions

Flood disasters, especially common in developing countries like China and India, have profound impacts on the overall economy. In China, one approach to mitigating severe river floods is the construction of Flood Detention Basins (FDBs). Strategically located in low-lying areas, FDBs are designed to temporarily hold excess floodwaters, thereby protecting downstream regions but increasing flood risk for those within the designated basins. While this policy may increase economic resilience against floods, it requires a closer examination of the economic costs and its uneven distributional impacts.

Chinese government states that residents living in FDB counties have made substantial sacrifice for the greater good. Our study quantitatively examines the economic costs and output gains of the FDB policy. We find that although the policy has improved the economic resilience against floods, it has also induced economic inequality between between FDB counties and their non-FDB counterparts. The empirical results of Chapter 1 show that counties designated as FDB counties by the Chinese government in 2000 experience persistent negative effects on their economic development. In studying the mechanism, Chapter 1 find that firms have less incentives to enter and invest in FDB counties due to their increased flood risks. In this project, we build a general equilibrium model to assess whether the FDB policy has yielded an overall increase in net output. Our counterfactual practice indicates that as FDBs absorb more floodwater, the total output gain brought by the policy would increase, though at the cost of widening inequality between FDB and other counties.

Our research has two major policy implications. First, our research highlights a critical insufficiency in the Chinese government's compensation on FDB counties. Since 2000, many counties has started to absorb floodwaters, thereby protecting other regions from flood damage. The compensation, however, focuses solely on compensating for direct losses caused by flood inundation, such as damage to agricultural crops. Our findings suggest that this compensation falls markedly short of addressing the total economic costs induced by the FDB policy. The substantial long-term economic costs have not been adequately compensated by the Chinese government. Based on our analysis, we recommend Chinese government to transfer the surplus taken by urban cities to rural counties.

Second, the findings of our study on China's Flood Detention Basin (FDB) policy offer insights for other nations contemplating similar flood risk management strategies. The evidence suggests that while such policies can provide broader regional protection from floods, they may come with significant long-term economic costs for the areas designated to absorb flood risks. For countries considering the adoption of similar policies, it is crucial to recognize the potential for creating economic disparities and to weigh these against the intended benefits of reduced flood risk. Policymakers should ensure that compensatory mechanisms are in place to support affected regions, mitigating the economic sacrifices made by FDB-designated areas. In sum, while such policies can be an effective component of a comprehensive flood risk management strategy, they should be implemented with careful consideration of the tradeoff between envrionmental justice and economic efficiency.

Chapter 3

Floods and Geographical Distribution of Patents

This chapter examines whether and how floods reshape the geography of innovation in China—a country highly exposed to flood risk, with a substantial share of its innovation activity concentrated in flood-prone areas. We develop a search-and-matching theoretical framework that emphasizes two key mechanisms: risk sharing and mutual support. Under the risk-sharing mechanism, firms in flooded counties seek collaborative partners in other regions to mitigate future flood risks. Under the mutual-support mechanism, firms with shared flood experiences are more likely to co-develop flood-resilient technologies. Leveraging detailed patent records and satellite-derived flood data from 2008 to 2018, we construct a novel dataset on cross-county patent collaborations. Our empirical findings support both mechanisms: while local innovation declines in flooded counties, firms in these areas increasingly collaborate with partners in other counties. Additionally, flooded counties are more likely to co-innovate on flood-resilient technologies with other similarly affected regions. Together, these results suggest that floods reshape the spatial dynamics of knowledge production by promoting broader inter-regional collaboration.

3.1 Introduction

Floods are among the most disruptive natural disasters globally, affecting more than 1.8 billion people to date and posing rising risks under climate change scenarios (Tellman et al. 2021). By 2050, the frequency of severe flooding events is projected to double across 40% of the world's land area (Arnell and Gosling 2016). Existing literature has thoroughly examined the economic consequences of floods, including direct damages to infrastructure and productivity (e.g., Kocornik-Mina et al. 2020) as well as changes in firms' investment and entry decisions in response to flood risks (e.g., Jia et al. 2022; Balboni et al. 2023; Hsiao 2024). However, much less is known about how natural disasters influence the geography of knowledge production. In particular, it remains an open question whether and how firms adjust their innovation strategies in response to increasing exposure to flood risk.

To address this question, we study China—a country that is not only one of the most flood-exposed globally, with approximately 395 million people at risk each year—but also a major player in global innovation. Many of China's leading innovation hubs are situated in regions with high flood exposure, making the country an ideal case for understanding the interaction between climate shocks and spatial patterns of technological advancement. Drawing on rich patent and satellite-derived flood data, we investigate whether floods reshape where and how innovation occurs, with a particular focus on inter-regional patent collaborations.

Descriptive patterns suggest that such a relationship may indeed exist. As illustrated in Figure 3.1, a large share of patenting activity in China is concentrated in flood-prone areas. However, Panel A of Figure 3.2 shows a marked inland shift in the spatial distribution of patent applications between 2008 and 2018. Initially concentrated in coastal provinces, patent activity gradually expanded to inland regions. Panel B reveals a concurrent increase in cross-county collaborations during the same period, with collaborative networks growing outward from a few coastal hubs to include a broader set of inland regions. These trends suggest that while floods may harm local innovation, they may also prompt firms to collaborate across regions—potentially altering the spatial distribution of innovation in the long run.

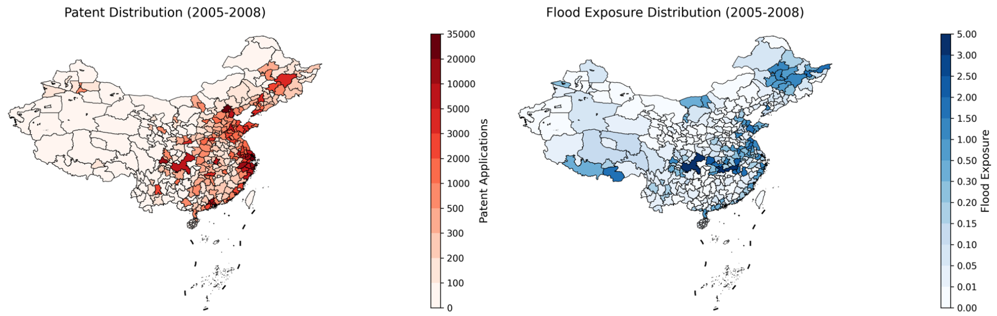
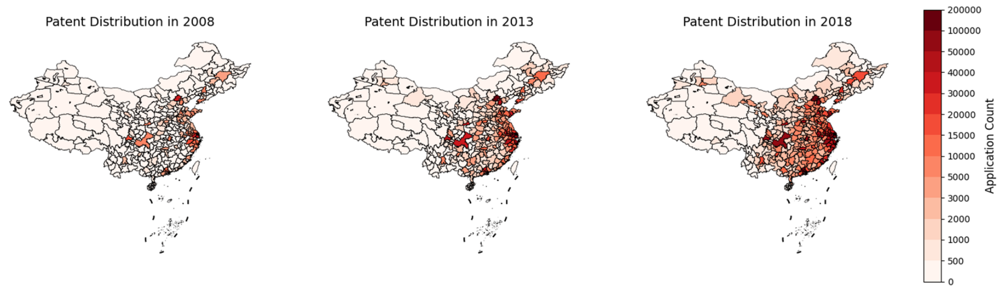
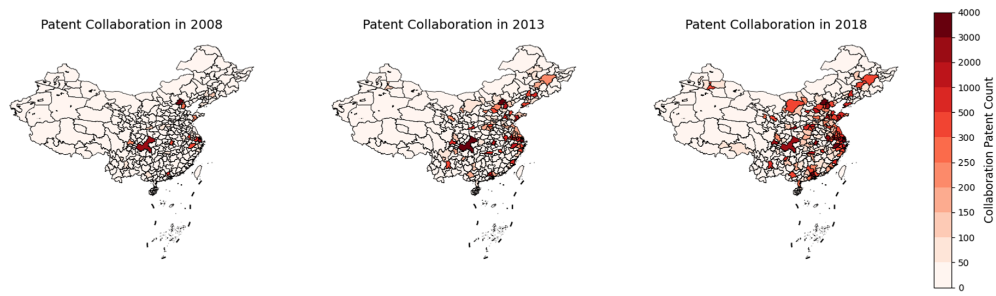


Figure 3.1: China's patent activities are more concentrated in flood zones.



(a) Patents are gradually shifting inlands.



(b) Patent collaborations are increasingly prevalent over time.

Figure 3.2: Spatial Distribution of Patents and Patent Collaborations in 2008, 2013, and 2018

Following these motivating facts, we ask: Do floods affect the geography of innovation in China? If so, through what mechanisms? To answer these questions, we construct a novel dataset that combines flood exposure measures from the Global Flood Database (GFD) with detailed patent records from the China National Intellectual Property Administration (CNIPA) and firm location data from Tianyancha. We use two complementary measures of flood exposure: Cumulative Flood Duration (CFD), which captures the total number of flood-experienced days per county, and Pixel-Adjusted

Flood Duration (PFD), which adjusts for county size and normalizes flood intensity. We follow the existing literature (e.g., Wu et al. 2022, and König et al. 2022) and focus on a unique episode of rapid expansions of China's patenting activities from 2008 to 2018. Figure A2 shows that China's total number of patent applications increased from 828,328 to 4323,112 during that period. For innovation outcomes, we identify 452,734 collaborative patents that involve 29,620 unique county pairs during the 2008–2018 period.

We develop a search-and-matching model to provide a theoretical foundation for understanding how firms respond to flood events in their innovation decisions. In our model, firms can invest either in general-purpose technologies or in flood-resilience innovations. After a flood realization, firms decide whether to collaborate with local or non-local partners. Two key mechanisms emerge from the model. The risk-sharing hypothesis posits that floods act as negative productivity shocks, reducing the expected returns to local innovation and prompting firms to seek external partners as a form of insurance against future disruptions. The mutual-support hypothesis, in contrast, suggests that firms in flooded counties may have stronger incentives to partner with others who share similar experiences in order to co-develop flood-resilient technologies. Together, these mechanisms offer a framework for understanding how floods shape collaboration networks and influence the geography of knowledge production.

We empirically test the two central hypotheses—risk-sharing and mutual-support—using a two-way fixed effects (TWFE) design that controls for both county and year fixed effects. This empirical strategy allows us to isolate the effect of flood exposure on innovation outcomes while accounting for time-invariant county characteristics and nationwide temporal shocks. Our analysis begins by examining the direct effect of floods on local innovation activity. We find that floods significantly hinder innovation within affected counties. Specifically, a one-day increase in Pixel-Adjusted Flood Duration (PFD) is associated with a 9.7% decline in county-level patent applications. And a standard deviation increase in PFD is associated with a 1.7% decline in county-level patent applications. This result underscores the disruptive nature of flooding, which interrupts ongoing R&D processes.

Next, we test the risk-sharing hypothesis, which suggests that firms respond to flood-induced productivity shocks by engaging in cross-regional collaborations to hedge against future disruptions. We find robust evidence supporting this mechanism. We find that a standard deviation increase in Pixel-Adjusted Flood Duration (PFD) is associated with a 2.1% increase in the number of county-pair level collaborative patent applications. Compared to the impact of floods on overall patenting activity, the positive effect on collaboration is notably large. While a standard deviation increase in Pixel-Adjusted Flood Duration (PFD) is associated with a 1.7% decline in county-level patent applications, it is also associated with a 2.1% increase in county-pair collaborative patenting. This suggests that counties not only respond to flood shocks by maintaining innovation through collaboration, but that the compensatory effect via collaboration may even outweigh the direct negative impact on local patenting activity.

To explore whether the magnitude of flood exposure amplifies this effect, we conduct a heterogeneity analysis that focuses on more severe flood events. We define a county-pair as experiencing a "major flood" if at least one county within the pair records a Pixel-Adjusted Flood Duration (PFD) above the 60th percentile of the national distribution. Under this definition, we observe that such county-pairs experience an 8% increase in collaborative patenting in the aftermath of major floods. This larger effect suggests that experiencing major floods more strongly motivates firms to seek collaborative partners. We then investigate the role of expectations of future flood risk in affecting collaborations. Using historical flood patterns, we decompose observed flood exposure into expected and unexpected components. The results suggest a standard deviation increase in expected Pixel-Adjusted Flood Duration (PFD) is associated with an at least 4.5% increase in cross-county collaborations, while unexpected flood shocks are found to have a negative and significant effect. This highlights the dominant role of flood expectations in shaping cross-regional collaborations: firms decide whether to collaborate based on their anticipated flood risks, which are informed by their historical exposure to flooding.

We then turn to the mutual-support hypothesis, which posits that shared flood experiences between regions encourage deeper and more targeted collaborations aimed at enhancing flood resilience. To test this mechanism, we classify county-pairs into three categories based on their flood exposure: (1) both counties experienced floods,

(2) only one county experienced a flood, and (3) neither county experienced a flood. We find that the collaborative response is significantly stronger among county pairs in which both counties have been exposed to flooding. Specifically, for these pairs, a one standard deviation increase in the average Pixel-Adjusted Flood Duration (PFD) is associated with a 1.2% increase in disaster-related collaborative patent applications. While this effect is smaller than the corresponding estimate for all patent types, it remains economically meaningful. In contrast, we find no significant increase in disaster-related collaborative patents among county pairs where only one county experienced flooding. These findings provide support for the mutual-support mechanism: only when both counties share similar flood experiences do they have strong incentives to jointly develop disaster-resilient technologies.

In summary, this paper demonstrates that floods could reshape the geography of innovation. On one hand, they reduce localized innovation in directly affected areas. On the other hand, they simultaneously stimulate inter-regional collaboration. These patterns are consistent with the dual mechanisms proposed in our theoretical framework: the risk-sharing mechanism, whereby firms hedge against localized shocks by collaborating with other counties, and the mutual-support mechanism, whereby shared experiences of floods tend to motivate cooperative innovation targeted at improving disaster resilience. Overall, our study contributes to a growing literature on the intersection of climate risk and innovation, and provides a framework for understanding the role of floods on reshaping the geographical distribution of innovations.

The remainder of the paper is organized as follows. Section 3.2 outlines the research background. Section 3.3 develops a theoretical framework to examine the role of floods. Section 3.4 describes the data sources and measurement strategies. Section 3.5 presents the empirical findings that test the two hypotheses. Finally, Section 3.6 concludes the paper.

3.2 Research Background

3.2.1 Floods in China

China ranks among the top countries globally for flood risk, due to its large population exposed to both coastal and river flooding. According to the Aqueduct Global Flood Risk Country Rankings, China ranks third in the world for the absolute number of people exposed to flood risks, with approximately 395 million people at risk annually. This places China among the most flood-exposed countries, alongside India and Bangladesh. About 27.5% of China's population is vulnerable to flooding, driven by river floods in the Yangtze, Huai, and Yellow River basins, as well as coastal areas prone to typhoons and rising sea levels. From 2000 to 2017, floods caused economic damage exceeding $150 billion, according to the EM-DAT International Disaster Database. Furthermore, Arnell and Gosling (2016) predicts that the likelihood of a 100-year flood occurring in China could increase by 33-67% by 2050.

A key feature of China's floods is their disproportionate impact on economically important regions. Jiangsu Province, for instance, ranks second in GDP among China's provinces, yet faces severe flood risks due to its location along the Yangtze River and Huai River. Fegions with higher flood risks are also more economically significant. For instance, the Yangtze River Basin, home to one-third of China's population, is a crucial economic hub. Frequent flooding, exacerbated by seasonal rainfall and extreme weather events, poses significant risks to infrastructure and livelihoods in these areas. Similarly, the Huai River Basin, another key region, faces recurring flood threats. Flooding in these economically vital regions could hinder China's overall economic growth, making flood management a critical concern for the government.

Due to rapid urbanization, urban populations in major cities (e.g., Beijing, Wuhan, and Nanjing) are increasingly exposed to severe flood risks. The urbanization rate surged to 64.72% in 2021, up from 36.00% in 2000, which has significantly heightened the vulnerability of urban areas to flooding. For instance, the 2012 Beijing flood, triggered by extreme rainfall, resulted in over 79 fatalities and caused approximately $2 billion in economic damage. The 2021 Zhengzhou flood led to over 350 deaths and caused around \$6 billion in economic losses. This underscores the severe impact of urban flooding on densely populated areas.

3.2.2 Patents in China

China has sustained a fairly long period of rapid economic growth over the past decade, to which the contributions from innovation have become increasingly important. Among numerous measurements, patents are often considered an indicator of technological change, which play a central role in research on innovation (Basberg, 1987; He et al., 2018) and attract considerable attention from researchers. The Early phase of innovation development was marked by Chinese representatives attending all international IPR conventions in the 1980s and 1990s[1]. And in April 1985, *Patent Law of China* was officially implemented, marking the beginning of patent modernization.

Since then, it took about 15 years for China to reach the first one million patent applications, but merely 1 year and 4 months to reach the fifth million patent applications in 2008 from the fourth million. After that period, China's patent applications began to experience an extraordinary surge. Specifically, China overtook the U.S. and Japan in terms of patent applications in 2009 and 2010, respectively (Hu et al., 2017). Furthermore, China in 2019 surpassed the U.S. to become the leading source of international patent applications filed with the WIPO[2].In 2023, the number of patent applications in China exceeded 5.5 million[3].

Figure 3.3 illustrates a more detailed picture of patent application. Among China's provinces and municipalities directly under the central government, the number of patent applications in Guangdong province has always ranked the top 1, and the number of patent applications has reached 0.96 million in 2023. Jiangsu, Zhejiang, Shandong, and Shenzhen follow closely. In addition, Beijing and Shanghai are another two municipalities directly under the central government ranking in the Top 10.

[1]For example, in June 1980, the Chinese government rode the wave of reform and became a member of the World Intellectual Property Organization (WIPO)

[2]See https://www.wipo.int/pressroom/en/articles/2020/article_0005.html for more details.

[3]See https://www.cnipa.gov.cn/tjxx/jianbao/year2023/a.html for more details.

Both the quantity and quality of patent applications in China have significantly increased recently, while a single entity often faces problems such as high R&D investment, high risks, and a low achievement transformation rate under the fierce technological competition. In order to sustain the robust growth in patents, we need to further understand the importance of collaboration (Anderson and Richards-Shubik, 2022), which can stimulate knowledge creation and innovation. Patent collaboration not only helps to integrates advantageous resources, but also promotes the efficient transformation and utilization of intellectual property rights. Therefore, conducting deep research on patent collaboration mechanism is an important aspect to achieve high-quality patent development.

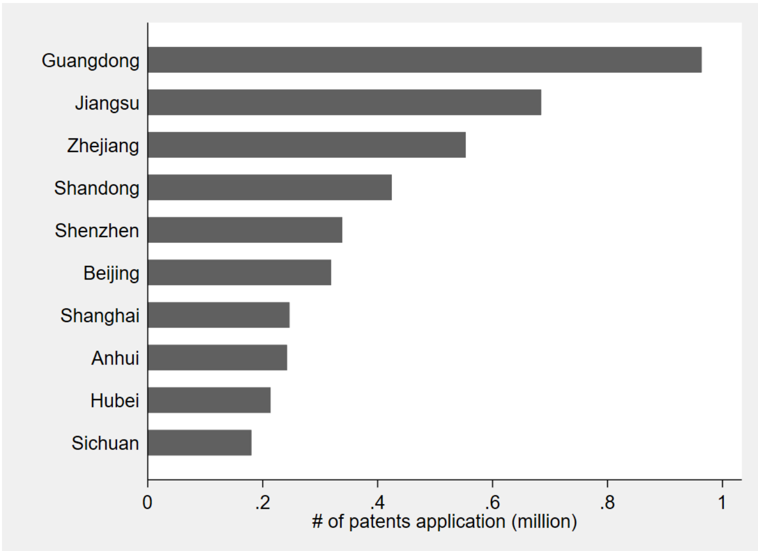


Figure 3.3: Patent Applications of Top 10 Provinces or Municipalities in 2023.

3.3 Theoretical Framework

3.3.1 Model Intuition

The timeline in Figure 3.4 illustrates the sequence of events in a typical period t for a representative firm. At the beginning of the period, the firm is endowed with two key attributes: its productivity level $z_{i,t}$ and its flood resilience level $\theta_{i,t}$.

1. Flood Realization. At the start of the period, the firm observes whether a flood occurs in its county. If a flood event takes place, it affects the firm's productivity through a damage function. Firms with higher flood resilience $\theta_{i,t}$ suffer less from the flood.

2. Production. Given its post-shock productivity $\hat{z}_{i,t}$, the firm produces output and earns profit $\pi_{i,t}$ during the production stage.

3. Innovation Decision. Following production, the firm decides whether to invest in *general technology* (to enhance productivity) or in *flood resilience* (to mitigate future flood damage). It also chooses the region c' in which to search for a collaboration partner.

4. Matching. The firm enters a search-and-matching process with potential partners in county c'. The likelihood of finding a collaborator depends on the size of the search pools in both counties and is given by the matching rates $f^{c,c'}_{gen,t}$ for general innovation and $f^{c,c'}_{flood,t}$ for flood resilience innovation.

5. Innovation Outcome. Conditional on matching, the probability of successful innovation depends on the characteristics of both the firm and its partner. For general innovation, success is increasing in the realized productivities of both parties, modeled by the function $g(\hat{z}_{i,t}, \hat{z}_{j,t})$. For flood resilience innovation, the probability of success depends on whether both firms have experienced flooding, as captured by $g(I^{Flood}_{i,t}, I^{Flood}_{j,t})$. Here, we believe that a commond experience of floods will shape the foundation of collaboration on developing flood resilient technologies.

6. State Update. At the end of the period, the firm updates its state variables. If general innovation is successful, its productivity increases to $z_{i,t+1}$. If flood innovation is successful, its resilience improves to $\theta_{i,t+1}$. These updated states become the inputs for the following period.

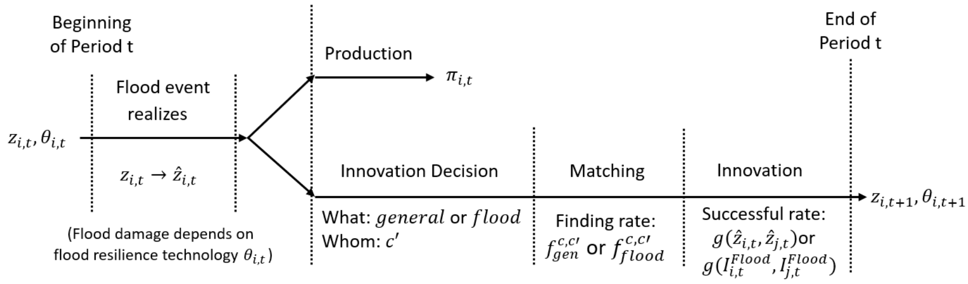


Figure 3.4: Model Timeline

3.3.2 Model Setup

Consider an economy with three counties, indexed by $c = 1,2,3$, which differ in both local demand (A_c) and flood risk (η_c). We impose two key assumptions. First, only counties 1 and 2 face a positive risk of flooding, such that $\eta_1, \eta_2 > 0$, while county 3 is flood-free with $\eta_3 = 0$. Second, counties 1 and 2 have stronger local demand relative to county 3, i.e., $A_1 = A_2 > A_3$. In each county, there is a unit mass of firms indexed by $i \in [0,1]$. Firms within the same county are heterogeneous in terms of productivity $z_{i,t}$ and flood resilience $\theta_{i,t}$. Let $G_c(z,\theta)$ denote the joint cumulative distribution function (CDF) of productivity and resilience among firms in county c. Then, the distribution satisfies the normalization condition:

$$\iint dG_c(z,\theta)\, dz d\theta = 1. \tag{3.1}$$

3.3.2.1 Firm Behavior and Profit Maximization

The static profit maximization problem of firm i in county c is given by:

$$\pi(\hat{z}_{i,t}) = \max_q \left[A_c q^{1-\sigma} - \frac{q}{\hat{z}_{i,t}} \right], \tag{3.2}$$

where A_c denotes the local demand shifter, $\sigma > 1$ is the elasticity parameter, and $\hat{z}_{i,t}$ represents the firm's actual productivity at time t after accounting for any flood-related damages. Actual productivity is defined as $\hat{z}_{i,t} = z_{i,t}e^{-\tau\theta_{i,t}\mathbb{I}^{Flood}_{c,t}}$, where $z_{i,t}$ is the firm's intrinsic productivity, $\tau > 0$ captures the proportional damage to productivity caused by a flood, and $\mathbb{I}^{Flood}_{c,t}$ is an indicator variable equal to 1 if a flood occurs in county c at time t, which happens with probability η_c. The term $\theta_{i,t} \in [0,1]$ measures the firm's flood resilience, such that higher $\theta_{i,t}$ corresponds to greater damage mitigation. Thus, firms with higher resilience suffer less productivity loss when a flood occurs.

3.3.2.2 Technology Evolution and Innovation

The evolution of firm i's general technology-driven productivity is governed by the following process:

$$z_{i,t} = (1 - \delta + \rho_1 \mathbb{I}^{Inn}_{i,t})z_{i,t-1}, \tag{3.3}$$

where $\delta \in (0,1)$ denotes the obsolescence rate of existing technology, and $\rho_1 > 0$ represents the proportional productivity gain from a successful innovation. The binary indicator $\mathbb{I}^{Inn}_{i,t} = 1$ if the firm successfully innovates in period t, and 0 otherwise. Innovation occurs through collaboration: when firm i successfully forms a partnership with another firm j, the probability of a successful innovation is given by the function $g(\hat{z}_{i,t}, \hat{z}_{j,t})$, which depends on the actual productivity levels of both firms at time t. This specification captures the idea that more productive firms are more likely to generate valuable innovations when collaborating.

3.3.2.3 Search and Matching

Let $\Omega^{c,c'}_{gen,t}$ denote the set of firms in county c that are searching for general technology innovation partners in county c' at time t. The process of forming innovation collaborations between counties c and c' is governed by a matching function $M(|\Omega^{c,c'}_{gen,t}|, |\Omega^{c',c}_{gen,t}|)$, which depends on the number of firms searching in both directions. The probability that a firm in county c successfully matches with a firm from county c' is then given by:

$$f^{c,c'}_{gen,t} = \frac{M(|\Omega^{c,c'}_{gen,t}|, |\Omega^{c',c}_{gen,t}|)}{|\Omega^{c,c'}_{gen,t}|}. \tag{3.4}$$

Conditional on a successful match, the probability that firm i in county c successfully innovates is determined by the quality of the match. Specifically, the probability of a successful innovation is given by:

$$\Pr(\mathbb{I}_{i,t}^{Inn} = 1) = f_{gen,t}^{c,c'} \int_{\Omega_{gen,t}^{c',c}} g(\hat{z}_{i,t}, \hat{z}_{c',t})\, dG_{c',t}, \tag{3.5}$$

where $g(\hat{z}_{i,t}, \hat{z}_{c',t})$ captures the likelihood of innovation success based on the productivity of firm i and the matched partner from county c', and $G_{c',t}$ is the distribution of actual productivity in the pool of potential partners from county c'.

Let $\Omega_{flood,t}^{c,c'}$ denote the set of firms in county c that are actively searching for partners in county c' to collaborate on flood resilience innovation at time t. The formation of such collaborations is governed by a matching function $M(|\Omega_{flood,t}^{c,c'}|, |\Omega_{flood,t}^{c',c}|)$, which depends on the number of firms searching in both counties. The probability that a firm in county c successfully matches with a partner from county c' is given by:

$$f_{flood,t}^{c,c'} = \frac{M(|\Omega_{flood,t}^{c,c'}|, |\Omega_{flood,t}^{c',c}|)}{|\Omega_{flood,t}^{c,c'}|}. \tag{3.6}$$

Conditional on a successful match, the probability that firm i achieves a successful flood resilience innovation is determined by the extent to which both it and its partner from county c' have experienced flooding. Formally, this probability is given by:

$$\Pr(\mathbb{I}_{i,t}^{Inn\_flood} = 1) = f_{flood,t}^{c,c'} \int_{\Omega_{flood,t}^{c',c}} g(\mathbb{I}_{i,t}^{Flood}, \mathbb{I}_{j,t}^{Flood})\, dG_{c',t}, \tag{3.7}$$

where $g(\cdot,\cdot)$ captures the probability of successful innovation as a function of both firms' flood exposure, and $G_{c',t}$ is the distribution of potential partners' characteristics in county c'.

3.3.2.4 Dynamic Optimization Problem

Firm i in county c chooses between two types of collaboration—general technology innovation or flood resilience innovation—to maximize its expected value given the current flood realization $Flood_t$. The firm's value function is defined as:

$$V(z_{i,t}, \theta_{i,t}, Flood_t) = \max\left\{V_{gen}^{c,c'}(z_{i,t}, \theta_{i,t}, Flood_t), V_{flood}^{c,c'}(z_{i,t}, \theta_{i,t}, Flood_t)\right\}, \quad (3.8)$$

where $V_{gen}^{c,c'}$ is the continuation value associated with collaboration on general innovation, and $V_{flood}^{c,c'}$ represents the value of flood resilience collaboration.

General Innovation Value Function

$$\begin{aligned} V_{gen}^{c,c'}(z_{i,t}, \theta_{i,t}, Flood_t) = & \pi(\hat{z}_{i,t}) - \kappa_{c,c'} + \beta(1-s)\mathbb{E}V(z_{i,t+1}, \theta_{i,t+1}, Flood_{t+1}) \\ \text{s.t.} \quad & \hat{z}_{i,t} = z_{i,t}e^{-\tau\theta_{i,t}\mathbb{I}_{c,t}^{Flood}}, \\ & \Pr(\mathbb{I}_{i,t}^{Inn} = 1) = f_{gen,t}^{c,c'} \int_{\Omega_{gen,t}^{c',c}} g(\hat{z}_{i,t}, \hat{z}_{c',t})\, dG_{c',t}, \\ & z_{i,t+1} = (1 - \delta + \rho_1 \mathbb{I}_{i,t}^{Inn})z_{i,t}, \\ & \theta_{i,t+1} = \theta_{i,t}. \end{aligned} \quad (3.9)$$

Flood Innovation Value Function

$$\begin{aligned} V_{flood}^{c,c'}(z_{i,t}, \theta_{i,t}, Flood_t) = & \pi(\hat{z}_{i,t}) - \kappa_{c,c'} + \beta(1-s)\mathbb{E}V(z_{i,t+1}, \theta_{i,t+1}, Flood_{t+1}) \\ \text{s.t.} \quad & \hat{z}_{i,t} = z_{i,t}e^{-\tau\theta_{i,t}\mathbb{I}_{c,t}^{Flood}}, \\ & z_{i,t+1} = (1-\delta)z_{i,t}, \\ & \Pr(\mathbb{I}_{i,t}^{Inn\_flood} = 1) = f_{flood,t}^{c,c'} \int_{\Omega_{flood,t}^{c',c}} g(\mathbb{I}_{i,t}^{Flood}, \mathbb{I}_{j,t}^{Flood})\, dG_{c',t}, \\ & \theta_{i,t+1} = (1 - \rho_2 \mathbb{I}_{i,t}^{Inn\_flood})\theta_{i,t}. \end{aligned} \quad (3.10)$$

3.3.3 Hypotheses

Based on the model set-up, we then propose two key mechanisms drive firms' collaborative behavior in response to flood events: the risk sharing channel and the mutual support channel. Correspondingly, we propose two hypotheses.

Hypothesis 1 (Risk Sharing Mechanism): Under the risk-sharing mechanism, when county c experiences a flood, firms in that county suffer from reduced realized productivity, which lowers the probability of successful general innovation within the county, i.e., $g(\hat{z}_{c,t}, \hat{z}_{c,t})$. This reduction in innovation success lowers the expected value from local general innovation collaboration, $V^{c,c}_{gen}$, and thereby increases the incentive for firms to seek general innovation partners in other counties, leading to an expansion in the cross-county collaboration set, $\Omega^{c,c'}_{gen}$.

Hypothesis 2 (Mutual Support Mechanism): Under the mutual-support mechanism, When county c is flooded, the value of collaborating with firms that have also experienced flooding rises, as the probability of a successful flood resilience innovation increases, i.e., $g(\mathbb{I}^{Flood}_{i,t}, \mathbb{I}^{Flood}_{j,t})$. However, the expected continuation value of resilience innovation, $V^{c,c'}_{flood}$, may fall due to immediate damages. This trade-off encourages a higher propensity for firms to search for flood resilience partners in other counties, increasing the size of the set $\Omega^{c,c'}_{flood}$. Together, these mechanisms explain how flood events dynamically shape inter-county collaboration patterns in both general and resilience-focused innovation.

3.3.4 An Illustrative Example

To illustrate the risk-sharing mechanism in our theoretical framework, we consider a simplified example in which the productivity distribution in county c follows a log-normal distribution. County c' serves as a potential collaboration partner, with its characteristics held constant throughout the analysis: the number of collaborators and their productivity levels are fixed. This setup ensures that any change in firm behavior stems solely from productivity shocks in county c, rather than changes in external conditions.

Figure 3.5 displays the value functions of firms across different productivity levels under two scenarios: a benchmark scenario without floods (Panel A) and a flood scenario (Panel B). The value function is normalized relative to the value of the no-innovation option. The x-axis represents firm productivity, and the y-axis represents the corresponding value function. The three lines represent different innovation strategies: no innovation (flat yellow line), within-county collaboration (red line), and cross-county collaboration (blue line).

Panel A: Benchmark (No Flood) Scenario We observe a non-linear relationship between firm productivity and collaboration choices:

(i) **Low productivity**: Firms choose not to establish collaborations or only collaborate within county c because their productivity is too low to offset the fixed cost of innovation, denoted κ.

(ii) **Medium productivity**: Firms choose to collaborate with partners from county c', seeking higher-productivity collaborators to increase the likelihood of successful innovation.

(iii) **High productivity**: Firms prefer to collaborate within their own county. Since their own productivity is already high, the success probability of innovation is sufficiently large without needing to incur the additional search cost of finding external partners.

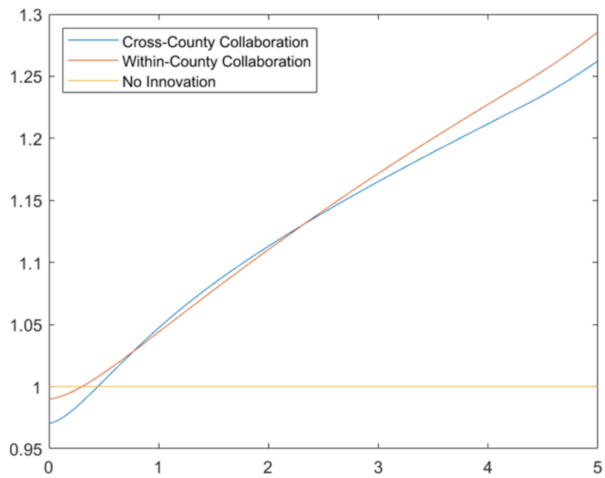
This pattern is consistent with empirical observations: for example, firms located in coastal regions—often with higher productivity—tend to collaborate less across counties, as shown in Figure 3.1.

Panel B: Flood Scenario When a flood hits county c, it reduces the realized productivity of firms. As a result:

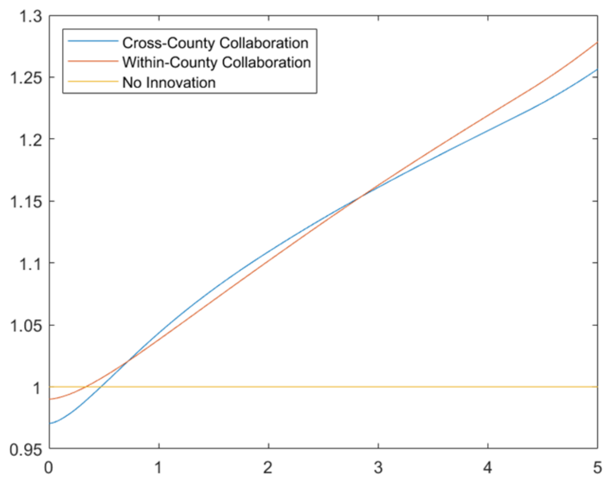
- Firms that previously belonged to the high-productivity group now find themselves in the medium-productivity range. To maintain innovation success, they switch from within-county to cross-county collaboration.

- Firms with low initial productivity must seek even more productive external collaborators to ensure innovation succeeds, which implies bearing higher search costs.

This behavioral shift reflects the **risk-sharing mechanism** posited by the model: after a negative productivity shock, firms seek external collaboration to hedge against local shocks and maintain innovation performance.



(a) Benchmark Scenario: without Floods



(b) Flood Scenario

Figure 3.5: Innovation Decisions in Benchmark Scenario (without Floods) and Flood Scenario

Note: (1) The Y-axis represents the firm's value function, while the X-axis represents firms' productivity levels; (2) The blue line represents the value function if firms in county c collaborate firms in county c', the orange line represents the value function if firms in county c only collaborate with firms within county c, while the yellow line represents the value function if firms do not collaborate; (3) Compared to the *Benchmark Scenario*, firms in the *Flood Scenario* experience a negative productivity shock caused by flooding.

3.4 Data

3.4.1 Measuring Floods

We gather data on flood events from the Global Flood Database (GFD), which provides comprehensive tracking of floods in China from 2000 to 2018. This database documents a total of 189 flood events across the country. Since the GFD offers satellite-derived flood maps at the county level, we are able to collect data on the duration of flooding for each 30m × 30m pixel. Additionally, the database identifies whether a pixel contains permanent water bodies, which are defined by the GFD as pixels that "are consistently identified with the presence of surface water for the majority of observations in 2000-2018 at 30-meter resolution, which was resampled to 250m resolution in Google Earth Engine using nearest neighbor resampling." Using this dataset, we construct two county-level proxies to quantify flood exposure:

Cumulative Flood Duration (CFD) captures the extent of flood exposure in a given county. It quantifies the total number of flood-experienced days across all pixels within a county in a given year. Specifically, CFD is measured in pixel-days, meaning that it accounts for both the spatial distribution and temporal persistence of floods. A higher CFD value indicates a more severe and prolonged presence of floods. It is constructed using the following equation:

$$CFD_{it} = \sum_{j \in A_i} FloodDuration_{jt} \tag{3.11}$$

where CFD_{it} denotes the cumulative flood duration for county i in year t. A_i represents the set of non-permanent water pixels, excluding permanent water bodies such as lakes and reservoirs. Hence, $FloodDuration_{jt}$ refers to the number of days that non-permanent water pixel j experiences flooding in year t.

Pixel-Adjusted Flood Duration (PFD) captures the average flood exposure per pixel within a given county. It is calculated as the cumulative flood duration (CFD) divided by the total number of non-permanent water pixels in the county. Unlike CFD, which reflects the total flood exposure in pixel-days, PFD normalizes this measure by county size and provides a standardized metric of flood exposure that is comparable across regions. A higher PFD value indicates that, on average, each pixel in a county experiences flooding for a greater number of days. It is constructed using the following equation:

$$PFD_{it} = \frac{CFD_{it}}{|A_i|} = \frac{\sum_{j \in A_i} FloodDuration_{jt}}{|A_i|} \tag{3.12}$$

where PFD_{it} indicates the pixel-adjusted flood exposure of county i at year t, $|A_i|$ represents the total number of non-permanent water pixels in county i. By adjust the total sum of flood duration by the size of non-permanent water pixels, we are able to measure the average number of flooded days experienced by non-permanent water pixels in county i at year t.

3.4.2 Measuring Collaborative Patents

Our primary data source is the China National Intellectual Property Administration (CNIPA). Patents are typically categorized into three types: invention, design patents, and utility models[4]. Invention patents are called "invention applications" initially, and some of them are approved eventually, which are called "invention grants". In our study, we focus on invention patent applications while using design patents and utility models as robustness checks. In the subsequent empirical results of this paper, we will refer "patent applications" as "invention patents application" for short. For each patent, we record a unique patent identifier number, along with the dates of application and publication, International Patent Classification (IPC) codes, the names of applicant(s) and inventor(s), and an associated address. We use the application date because it more accurately reflects when new knowledge was created and formalized (Moretti, 2021). Moreover, since inventors are mainly individuals while applicants can be firms, research institutes, individuals, or other types of entities, and patents assigned to firms usually take up a dominant share in recent decades (Akcigit et al., 2022), we finally adopt

[4]According to the definition of CNIPA, an invention patent is a new technical solution related to a product, a process, or an improvement; An utility model is a new technical solution related to a product's shape, structure, or a combination thereof fit for practical use; A design patent is a new design of the shape, pattern, or a combination thereof. More detailed information about patents can be found at https://english.cnipa.gov.cn/col/col2995/index.html

"applicants" as the proxy for innovators in our empirical analysis (Koh et al., 2024).[5]

To identify cross-county patent collaborations, a central challenge with our analysis is to identify the county in which all applicants are located. The CNIPA database only reports the first-listed applicant's exact address, and the address of other applicants is not reported. Hence, we are not able to identify the address of all applicants using only CNIPA database. To solve this issue, we constructed a search database following the work of Koh et al. (2024) and used Tianyancha, a firm-level database with detailed firm location data, to manually cross-check our identification. Specifically, we address this issue using two approaches following Koh et al. (2024).

Firstly, we construct a search database that includes accurate county-level addresses for applicants of solo patent applications and for first-listed applicants of collaborative patents. Although our sample period covers 2008-2018, we can retrieve patent addresses from the CNIPA database for the years 2000 to 2022 to extend our search database. To prevent the misidentification of distinct entities that may share identical applicant names, we limit the extension of the time range. Then, we standardize all applicant names in a consistent format and get approximately 600,000 unique names. And we search patents in both earlier and later years[6].

Secondly, we obtain approximately 60,000 firms' addresses from the Tianyancha database, which we use both to match firms that are not identified in the first step and to double-check the addresses provided by the patent database. Finally, we have addresses for 452,734 collaborative patents and 29,620 county-pairs.

The primary outcome of interest is patent collaborations over time and across regions. We classify a patent with multiple applicants as a "collaborated patent." Our key outcome variables are: (1) $CollabPatents_{ij,t}$, the count of patents in which applicants are located in both paired counties in a given year t for the county level and county-pair level data. For example, suppose a patent has 3 applicants located in coun-

[5] We would like to note that for each patent, we exclude individual applicants from the applicant list using both automated and manual methods. That is because it is difficult to obtain the address of an individual applicant. For example, if a patent application lists two applicants - one is a company and another is an individual - we treat it as a single application. And if a patent only has an individual applicant, then we remove this observation from our database.

[6] For example, we have a patent in 2008, we search sequentially in 2008, 2007,...,2000, then 2011, 2012,..., up to 2022.

ties A, B, C, respectively, in year 2008. Then this patent would be counted in the variable $CollabPatents$ for the county pairs (A, B), (B, C), (A, C) in year 2008; (2) $D\_CollabPatents_{ij,t}$, a dummy variable for the patent level data that equals 1 if counties i and j engage in collaboration in year t, and 0 otherwise.

3.4.3 Descriptive Statistics

Our sample covers the period from 2008 to 2018. Table 3.1 reports the descriptive statistics for the main variables used in our analysis. This table presents three datasets used in this paper: county level, county pair level, and patent level data, respectively. Besides patents applications and relevant measurements of flood above, there are other county-level control variables. County characteristics include the logarithm of the population, GDP, loan balance of financial institutions, the number of industrial firms above state designated scale, and the value-added of the tertiary industry.

Table 3.1 shows that, on average, each county has approximately 155 patent applications per year, and 140 of them only have 1 non-individual applicant. Among all the patents, there are about 4 disaster mitigation patents. In terms of patent collaborations, the county pairs in the sample produce about 8 patents together each year, and 1.8% of them are disaster mitigation patents. And in those collaborative patents, 76% of them are cross-county collaborations. We also show the time pattern of patent applications and collaborations in Figure B3.

Table 3.1: Descriptive Statistics

County Level						
Patent Applications	28,867	155.377	651.047	0.000	8.000	723.000
Single Applications	28,867	139.067	585.210	0.000	7.000	639.000
Disaster Patents	28,867	3.742	21.474	0.000	0.000	15.000
$\ln(L_1(\text{CFD}))$	28,867	0.834	2.191	0.000	0.000	6.483
$L_1(\text{PFD})$	28,867	0.014	0.179	0.000	0.000	0.010
ln(Population)	28,867	3.591	0.883	1.809	3.717	4.745
ln(GDP)	28,867	13.807	1.179	11.716	13.888	15.618
ln(Loan)	28,867	13.074	1.389	10.778	13.122	15.282
ln(# of Firms)	28,867	3.989	1.422	1.386	4.094	6.127
ln(Third)	28,867	12.753	1.238	10.665	12.769	14.795
County Pair Level						
Collaborative Patents	65,828	7.771	47.303	1.000	2.000	22.000
Disaster Patents	65,828	0.140	1.556	0.000	0.000	1.000
$\ln(L_1(\text{CFD}))$	65,828	0.788	2.098	0.000	0.000	6.492
$L_1(\text{PFD})$	65,828	0.012	0.097	0.000	0.000	0.036
Patent Level						
Collaborative Dummy	452,734	0.764	0.425	0.000	1.000	1.000

Note: This table reports the summary statistics, including county level, county pair level, and patent level variables. Detailed variable definitions are presented in Table A1.

3.5 Empirical Results

Building on the theoretical framework outlined in Section 3.3, we present empirical evidence to test the corresponding hypotheses. First, we examine whether floods negatively affect local patenting activity. Second, we investigate the risk-sharing mechanism by analyzing whether firms are more likely to form collaborations with partners in other counties. Third, we assess the mutual-support mechanism by testing whether firms tend to co-develop flood-resilient patents with partners from counties that have also experienced floods.

3.5.1 Impacts of Floods on Patent Applications

We first conduct a county-level analysis to understand the impact of floods on county-level patent applications. The regression is as follows:

$$Y_{it} = \beta_1 + \beta_2 FloodProxy_{i,t-1} + \lambda_t + \eta_i + \varepsilon_{it} \tag{3.13}$$

where Y_{it} represents the logarithm of the number of patent applications in county i. We examine three types of patents: all patents, patents with single applicant, and disaster mitigation patents. $FloodProxy_{i,t-1}$ denotes either $ln(CFD)i,t-1$, the lagged cumulative flood duration of county i, or $ln(PFD)i,t-1$, the lagged pixel-adjusted flood duration of county i. λ_t represents time fixed effects, η_i denotes county fixed effects, and ε_{it} is the standard error, clustered at the county level. Accordingly, β_2 captures the impact of floods on county-level patent applications.

Table 3.2 presents the estimated impacts of floods on patent applications at the county level, with a focus on different types of patents. Independent variables in this table are logarithm cumulative flood duration (CFD) and pixel-adjusted flood duration (PFD). Across all columns, year and county fixed effects are included to control for time-invariant characteristics and temporal trends. The results indicate that flood exposure leads to a decline in number of patent applications. Columns (1) and (2) report the effects on total patent applications, while columns (3) and (4) focus on patents filed by single applicants. The estimates suggest that a 1% increase in Cumulative Flood Duration (CFD) reduces the number of total patents by 1.7% and single-applicant patents by 2.8%, respectively. These findings suggest the negative effect of floods on local innovation.

Columns (5) and (6) provide evidence of a compensatory effect in disaster-related innovation. We use the *China Meteorological Disaster Yearbook* and the *National Emergency Response Plan* to manually extract a total of 34 keywords from the sections that cover flood-related content and general disaster prevention and control. Figure C4 provide the keywords of flood patents. Then, we search through the abstracts of individual patents to determine whether they contain these keywords. If a patent abstract includes any of the keywords, it is classified as a disaster mitigation patent. While total patenting declines, patents specifically related to disaster mitigation increase after a county experiences floods. Importantly, these results also serve as a placebo test for the main findings, which verify that the observed negative impact of flooding on overall patent applications is not simply a result of general changes in innovation trends or unobserved confounding factors. If floods were simply discouraging all forms of innovation, we would expect to see a uniform decline across all patent categories, including disaster-related patents. However, the fact that disaster mitigation patents increase suggests that the observed decline in total patents is not driven by a general downturn in patenting activity but rather by a reallocation of innovation efforts.

Overall, these findings highlight how firms respond to natural disasters by reallocating their innovation efforts. While floods lead to a decline in overall patenting activity, they simultaneously stimulate innovation in disaster mitigation technologies.

Table 3.2: Impacts of Floods on Number of Patent Applications
- County Level Analysis

	ln(Number of Patent Applications)					
Type of Patents:	All		Single Applicant		Disaster Mitigation	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_1(\text{CFD}))$	-0.017^{***}		-0.028^{***}		0.003^{***}	
	(0.003)		(0.004)		(0.002)	
$L_1(\text{PFD})$		-0.097^{***}		-0.095^{***}		0.024^{***}
		(0.035)		(0.034)		(0.011)
R-squared	0.879	0.878	0.871	0.871	0.747	0.747
N(obs)	28,867	28,867	28,867	28,867	28,867	28,867
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y	Y

Note: (1) 'CFD' denotes Cumulative Flood Duration, and 'PFD' denotes Pixel-Adjusted Flood Duration (see Section 3.4.1 for details). (2) The regression specification is: $Y_{it} = \beta_1 + \beta_2 FloodProxy_{i,t-1} + \lambda_t + \eta_i + \varepsilon_{it}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{CFD}))$, the logarithm of lagged CFD (lagged by one year), or L_1(PFD), the lagged PFD (lagged by one year). (3) 'All' includes all patents, 'Single Applicant' refers to patents filed by applicants with only one application. 'Disaster Mitigation' patents are those related to disaster alleviation and are identified using *China Meteorological Disaster Yearbook* and the *National Emergency Response Plan*; (4) Standard errors are clustered at the county level.

3.5.2 Testing the Risk-Sharing Hypothesis

We then test the risk-sharing mechanism by examining whether firms in flood-affected counties are more likely to collaborate with partners in other counties to distribute the flood risks.

3.5.2.1 Baseline Results

We first conduct a patent-level analysis to understand the impact of floods on cross-county patent collaborations. The regression is as follows:

$$DCollab_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt} \tag{3.14}$$

where $DCollab_{ij,t}$ is a dummy variable that equals 1 if county i and county j have formed patent collaborations at time t, and 0 if not. $FloodProxy_{ij_t-1}$ represents either $ln(L_1CFD)_{ij,t-1}$, the lagged average cumulative flood duration of county i and county j, or $L_1PFD_{ij,t-1}$, the lagged average pixel-adjusted flood duration of county i and county j. Additionally, we incorporate time fixed effects λ_t, county fixed effects η_i and π_j into the regression. Standard errors are clustered at the county-pair level.

Table 3.3 and Table 3.4 examine the relationship between flood exposure and patent collaborations at the patent and county-pair level, respectively. In Table 3.3, the outcome variable is a dummy indicating whether a patent was filed collaboratively. Column (1) estimates the effect of cumulative flood duration (CFD), while column (2) presents results using pixel-adjusted flood duration (PFD). The estimates suggest that flood exposure is positively associated with an increase in patent collaborations. Specifically, a 1% increase in cumulative flood exposure raises the likelihood of a patent being collaborative by 1.3%, while one day increase in pixel-adjusted flood exposure increases this likelihood by 3.7%. These findings suggest that firms or inventors respond to flood-induced disruptions by engaging in more cross regional collaborative innovation efforts. We also present a robustness check in Table C2.

Table 3.3: Impacts of Floods on Patent Collaborations - Patent Level Analysis

Dependent Variable:	Collaborative Patent Dummy	
	(1)	(2)
$\ln(L_1(\text{Cumulative Flood Duration}))$	0.013***	
	(0.001)	
$L_1(\text{Pixel-Adjusted Flood Duration})$		0.037***
		(0.011)
R-squared	0.395	0.393
N(obs)	452,734	452,734
Fixed Effects		
Year	Y	Y
County 1	Y	Y
County 2	Y	Y
Patent Type	Y	Y

Note: (1) Detailed definition of Cumulative Flood Duration and Pixel-Adjusted Flood Duration can be found in Section 3.4.1; (2) The regression specification is: $DCollab_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{Cumulative Flood Duration}))$, the logarithm of lagged CFD (lagged by one year), or $L_1(\text{Pixel-Adjusted Flood Duration})$, the lagged PFD (lagged by one year). (3) Standard errors are clustered at the county level.

We then conduct a county-pair analysis to understand the impact of floods on cross-county patent collaborations. And the results remain robust when conducting a similar regression at the county-pair level. The regression is as follows:

$$Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt} \tag{3.15}$$

where $Y_{ij,t}$ denotes the logarithm of the number of patent applications between county i and county j at time t. $FloodProxy_{ij_t-1}$ represents either $ln(CFD)_{ij,t-1}$, the lagged average cumulative flood duration of county i and county j, or $PFD_{ij,t-1}$, the lagged average pixel-adjusted flood duration of county i and county j. Additionally, we incorporate time fixed effects λ_t, county-pair fixed effects γ_{ij}, county fixed effects η_i and π_j, and time fixed effects λ_t. Standard errors are clustered at the county-pair level.

In Table 3.4, the outcome variable is the logarithm of the number of patents jointly filed by inventors from two different counties. Column (1) reports estimates using cumulative flood duration (CFD), while column (2) relies on pixel-adjusted flood duration (PFD). The findings indicate a positive relationship between flood exposure and cross-county patent collaborations. Specifically, a 1% increase in cumulative flood exposure is associated with a 1.1% increase in county-pair collaborative patents, while an additional standard deviation increase of pixel-adjusted flood duration (PFD) corresponds to a 2.1% increase. These results suggest that, although floods negatively impact local innovation activities (Table 3.2), they also stimulate interregional collaboration in innovation.

Table 3.4: Impacts of Floods on Patent Collaborations - County Pair Level Analysis

	ln (Number of County-Pair Collaborative Patents)	
	(1)	(2)
$\ln(L_1$(Cumulative Flood Duration)$)$	0.011***	
	(0.003)	
L_1(Pixel-Adjusted Flood Duration)		0.120**
		(0.059)
R-squared	0.434	0.434
N(obs)	65,828	65,828
Fixed Effects		
Year	Y	Y
County-Pair	Y	Y
County 1	Y	Y
County 2	Y	Y

Note: (1) Detailed definition of Cumulative Flood Duration and Pixel-Adjusted Flood Duration can be found in Section 3.4.1; (2) The regression specification is: $Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1$(Cumulative Flood Duration)$)$, the logarithm of lagged CFD (lagged by one year), or L_1(Pixel-Adjusted Flood Duration), the lagged PFD (lagged by one year); (3) Standard errors are clustered at the county-pair level.

3.5.2.2 Heterogeneity Analysis: the Impact of Larger Floods

We then extend our analysis to identify the impacts of larger floods on cross-county patent collaborations using a specification as below:

$$Y_{ij,t} = \alpha + \beta_1 Flood_{ij,t} + \beta_2 X_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ij,t} \quad (3.16)$$

where $Y_{ij,t}$ denotes the logarithm of the number of patent applications between county i and county j at time t. Our treatment variable, $Flood_{ij,t}$ is a binary indicator that equals 1 if, at time t, the Pixel-Adjusted Flood Duration (PFD) in either county i or county j exceeds 60% of the nation's historical record, and 0 otherwise. By defining the treatment in this way, we ensure that our analysis captures only large flood events that are likely to cause substantial impacts on economic activities. This threshold helps distinguish severe floods from minor inundations but we also conduct robustness checks using other thresholds. The control variable is represented by $X_{ij,t}$. Additionally, we incorporate county-pair fixed effects γ_{ij}, county fixed effects η_i and π_j, and time fixed effects λ_t. Standard errors are clustered at the county-pair level.

To illustrate the time trend of the impacts of floods, we also conduct an event study following this specification:

$$Y_{ij,t} = \sum_{n=1}^{n=3} \alpha_n Flood_{ij,t-n} + \sum_{m=0}^{m=3} \alpha_m Flood_{ij,t+m} + \alpha_2 X_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ij,t} \quad (3.17)$$

$Flood_{ij,t-n}$ represents the nth lag, while $Flood_{ij,t+m}$ denotes the mth lead. This specification enables us to examine the impact of floods on innovation collaboration both before and after their occurrence. It allows us to investigate the dynamic effects of floods over time. Additionally, it allows us to test the parallel trends assumption by evaluating whether pre-flood trends in innovation activity differ significantly between treated and untreated county pairs.

Table 3.5 examines the impact of severe flooding on county-pair patent collaborations, using two flood intensity thresholds: when the Pixel-Adjusted Flood Duration (PFD) exceeds 60% and 70% of its historical record. The results show that flooding increases patent collaborations by 8.4% and 10.1% at these thresholds, respectively, with both effects statistically significant at the 1% level. This suggests that severe floods prompts inventors to seek collaborations across counties. The stronger effect at the 70% threshold implies that extreme disasters have a greater impact on innovation networks.

Table 3.5: Causal Impacts of Floods on Patent Collaborations - County Pair Level Analysis

	ln (Number of County-Pair Collaborative Patents)	
Flood Threshold:	PFD > 60%	PFD > 70%
	(1)	(2)
Estimated Flood Impact ($\hat{\beta}_1$)	0.084***	0.101***
	(0.030)	(0.033)
R-squared	0.413	0.413
N(obs)	65,828	65,828
Fixed Effects		
Year	Y	Y
County-Pair	Y	Y
County 1	Y	Y
County 2	Y	Y

Note: (1) 'PFD' refers to Pixel-Adjusted Flood Duration, and the detailed description can be found in Section 3.4.1; (2) The difference-in-differences regression specification is: $Y_{ij,t} = \alpha + \beta_1 Flood_{ij,t} + \beta_2 X_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ij,t}$ where $Flood_{ij,t}$ is a binary indicator that equals 1 if, at time t, the Pixel-Adjusted Flood Duration (PFD) in either county i or county j exceeds 60% or 70% of the nation's historical record, and 0 otherwise; (3) Standard errors are clustered at the county-pair level.

The event study analysis, presented in two panels in Figure 3.6, examines the impact of severe flooding on county-pair patent collaborations, using two flood intensity thresholds: PFD > 60% (left panel) and PFD > 70% (right panel). Both figures show similar patterns, with no significant differences in patent collaboration trends before the flood, supporting the parallel trends assumption in the specification. In both cases, collaborations increase starting in the year of the flood (Year 0) and persist for at least three years post-flood, indicating a long-term effect. The similarity in results across the two thresholds demonstrates the robustness of the findings.

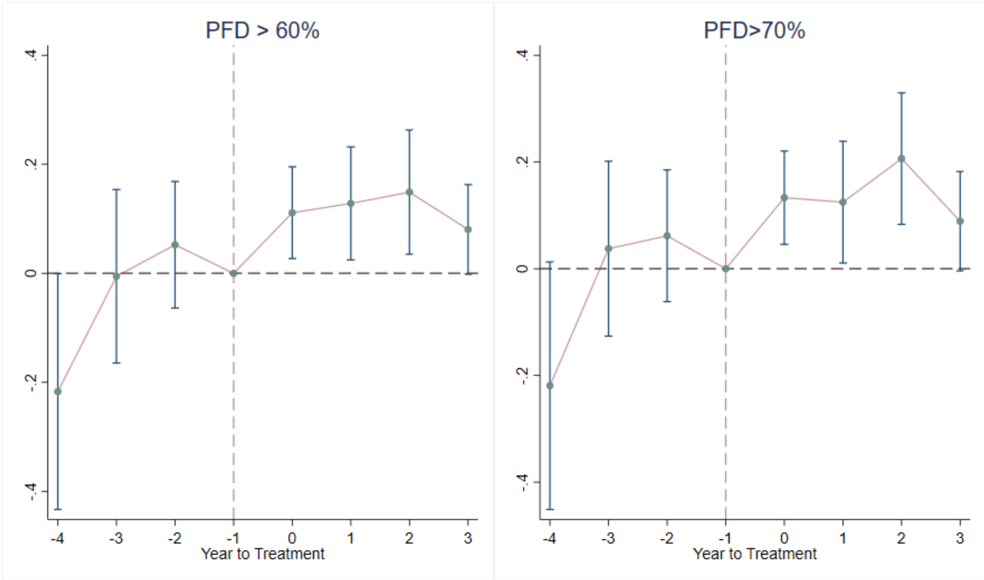


Figure 3.6: Dynamic Impacts of Floods on Cross-County Patent Collaborations
Note: (1) The event-study regression is specified as:$Y_{ij,t} = \sum_{n=1}^{3} \alpha_n Flood_{ij,t-n} + \sum_{m=0}^{3} \alpha_m Flood_{ij,t+m} + \alpha_2 X_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ij,t}$ The black dots represent the policy effect (ATT), corresponding to α_n and α_m; (2) In the left figure, the treatment variable $Flood_{ij,t}$ equals 1 if, at time t, the Pixel-Adjusted Flood Duration (PFD) in either county i or county j exceeds 60%. In the right figure, the threshold is set at 70%; (3) The confidence intervals are reported at the 95% confidence level.

3.5.2.3 Mechanism Analysis: the Role of Flood Expectation

We then investigate the role of flood expectations in shaping cross-region collaboration patterns. Our hypothesis is that flood events can motivate firms to engage in collaborative activities as a form of adaptive response. Specifically, when firms experience floods, they may seek to reduce their exposure to future risks by forming partnerships with firms in other counties—effectively sharing risk across geographic areas. As such, the expectation of future flooding, shaped by a county's historical flood experiences, becomes a critical factor in driving these collaborations. If firms anticipate a higher likelihood of future flooding based on recent trends, they may be more proactive in building collaborative networks that enhance resilience. Therefore, we expect that a stronger history of flooding, particularly recent flooding, will be associated with a greater propensity to form collaborative links with external partners.

Decomposing Floods: Expected and Unexpected Component

To investigate the role of floods, we first tend to decmpose floods into the expected part and the unexpected part. In each year, we use high-frequency data on flood risk to estimate, for each county, a, *expected flood duration* as the predicted flood shock:

$$\hat{F}_{it} = f\left(F_{i,t-5}, \ldots, F_{i,t-1}\right) \tag{3.18}$$

We then decompose the actual flood duration in year t, denoted by F_{it}, into a predictable component—the *expected flood duration* $\hat{F}_{it}$—and an *unpredictable flood shock*, defined as the difference between the observed and expected flood durations:

$$\text{Flood Shock}_{it} = F_{it} - \hat{F}_{it}. \tag{3.19}$$

where $f(\cdot)$ denotes a general prediction function that can take various forms. In our benchmark analysis, the prediction function is specified as either the unweighted or weighted average of flood durations over the previous five years. Specifically, for the unweighted case, we define the expected flood duration as

$$\hat{F}_{it,\text{unweighted}} = \frac{F_{i,t-5} + F_{i,t-4} + F_{i,t-3} + F_{i,t-2} + F_{i,t-1}}{5}. \tag{3.20}$$

The economic intuition behind this specification is that all past flood events are assumed to have an equal influence in shaping expectations. For the weighted case, we define

$$\hat{F}_{it,\text{weighted}} = \frac{0.03125F_{i,t-5} + 0.0625F_{i,t-4} + 0.125F_{i,t-3} + 0.25F_{i,t-2} + 0.5F_{i,t-1}}{5}. \tag{3.21}$$

Here, the economic intuition is that more recent flood events exert a stronger influence on expectations, while older events have progressively less impact. We then decompose the actual flood duration in year t, F_{it}, into a predictable component, or *expected flood duration*, $\hat{F}_{it}$, and an *unpredictable flood shock*, defined as the difference between the observed flood duration and expected flood duration:

$$S_i = F_{it} - \hat{F}_{it} \tag{3.22}$$

Since we have two proxies to measure floods, Cumulative Flood Duration (CFD) and Pixel-Adjusted Flood Duration (PFD), following the approach above, we will have expected and unexpected CFD and PDF, respectively.

Flood Expectation and County-Level Patent Applications

To understand the impact of flood expectation on patent collaborations, we then conduct a regression similar to Equation 3.13:

$$Y_{it} = \beta_1 + \beta_2 ExpectedFloodProxy_{i,t-1} + \gamma_t + \lambda_i + \varepsilon_{it} \tag{3.23}$$

where $ExpectedFloodProxy_{i,t-1}$ denotes either $ln(ExpectedCFD)_{i,t-1}$, the lagged expected cumulative flood duration of county i, or $ExpectedPFD_{i,t-1}$, the lagged expected pixel-adjusted flood duration of county i. Other notations are the same as Equation 3.14.

Table 3.6 reports the impact of expected flood duration on county-level patent applications. Panel A presents results using Expected Cumulative Flood Duration (ECFD) as the independent variable, while Panel B uses Expected Pixel-Adjusted Flood Duration (EPFD). The estimates indicate that greater expected flood exposure is negatively associated with county-level patent applications. Specifically, a 1% increase in ECFD is associated with a 2.9% to 3.2% decline in patent applications, as shown in columns (1)

Table 3.6: Impacts of Flood Expectation on Patent Applications - County Level Analysis

	ln(Number of Patent Applications)			
	(1)	(2)	(3)	(4)
Panel A: Independent Variable - Expected Cumulative Flood Duration (ECFD)				
ln(ECFD)	-0.032^{***}			
	(0.004)			
ln(Weighted ECFD)		-0.035^{***}		
		(0.004)		
Panel B: Independent Variable - Expected Pixel-Adjusted Flood Duration (EPFD)				
EPFD			-0.572^{***}	
			(0.106)	
Weighted EPFD				-0.290^{***}
				(0.086)
R-squared	0.879	0.879	0.878	0.878
N(obs)	28,867	28,867	28,867	28,867
Fixed Effects				
Year	Y	Y	Y	Y
County	Y	Y	Y	Y

Note: (1) 'ECFD' refers to Expected Cumulative Flood Duration, and 'EPFD' refers to Expected Pixel-Adjusted Flood Duration, and the detailed construction can be found in Section 3.5.2.3; (2) The regression takes the form of $Y_{it} = \beta_1 + \beta_2 ExpectedFloodProxy_{i,t-1} + \gamma_t + \lambda_i + \varepsilon_{it}$ where $ExpectedFloodProxy$ is either Expected Cumulative Flood Duration or Pixel-Adjusted Flood Duration; (3) standard errors are clustered at the county level.

and (2). And one standard deviation increase in expected pixel-adjusted flood duration reduces patent applications by at least 5%. These results suggest that anticipated flood risks discourage innovation at the county level, likely due to firms adjusting their R&D investments and resource allocations in response to expected climate disruptions.

Flood Expectation and Patent Collaborations

We then conduct regressions similar to Equation 3.25 and Equation 3.16 to understand the impact of flood expectations on cross-county patent collaborations. At the patent level, the regression is similar to Equation 3.25 and takes the following form:

$$DCollab_{ijt} = \alpha + \beta_1 ExpectedFloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt} \tag{3.24}$$

where $ExpectedFloodProxy_{ij,t-1}$ denotes either $ln(ExpectedCFD)_{ij,t-1}$, the lagged expected average cumulative flood duration of county i and county j, or $ExpectedPFD_{i,t-1}$, the lagged expected average pixel-adjusted flood duration of county i and county j. Other notations remain the same as Equation 3.14.

The results in Tables 3.7 and Table 3.2 provide strong evidence that expected flood exposure increases interregional patent collaborations across counties. At the patent level (Table 3.7), one standard deviation increase in Expected Pixel-Adjusted Flood Duration (EPFD) is associated with an even larger 2.9% to 3.8% increase. Similarly, at the county-pair level (Table 3.2), one standard deviation increase in EPFD is associated with at least a 4.5% increase. These findings suggest that firms anticipating floods are more likely to collaborate across regions.

While previous results (e.g., Table 3.6) show that expected floods reduce overall innovation, the evidence here suggests that firms adjust their innovation strategies by forming collaborations across counties. The increase in cross-county patenting reflects how firms adapt to floods, where firms and inventors seek partners beyond their local areas to mitigate the negative effects of floods. This highlights how flood risks influence the geographic distribution of innovation. In response to floods, firms increasingly engage in interregional cooperation.

The county-pair analysis is as follows:

$$Y_{ijt} = \alpha + \beta_1 ExpectedFloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt} \tag{3.25}$$

where $ExpectedFloodProxy_{ij,t-1}$ denotes either $ln(ExpectedCFD)_{ij,t-1}$, the lagged expected average cumulative flood duration of county i and county j, or

$ExpectedPFD_{i,t-1}$, the lagged expected average pixel-adjusted flood duration of county i and county j. Other notations remain the same as Equation 3.16

Placebo: Unexpected Flood Shock and Patent Collaborations

The results in Table 3.9 and Table 3.10 serve as a placebo test to verify that only expected floods will foster interregional collaborations cross counties, whereas unexpected floods have a negative effect on patent collaborations. This placebo test is crucial because if the observed increase in collaboration were driven by external factors unrelated to flood anticipation—such as general innovation trends or random shocks—then both expected and unexpected floods should yield similar results. However, the findings indicate a stark contrast: while expected floods increase collaboration (as shown in Table 3.6 and Table 3.8), unexpected floods has a negative direct effect to reduce interregional collaborations. This confirms that firms shape flood expectations and adapt strategically in accordance to the expectation.

At the patent level (Table 3.9) and at the county-pair level (Table 3.10), we all find a much smaller and less significant impacts of unexpected floods on collaborative patents. These results suggest that unexpected floods have a direct and negative impact on local innovation networks. Unlike expected floods, which allow firms to proactively establish partnerships to mitigate risks, unexpected floods occur as sudden shocks that disrupt economic activities and make it more difficult for firms to coordinate joint innovation efforts. The finding that only expected floods lead to increased collaboration further supports the idea that firms adjust their innovation strategies in anticipation of flood risks, rather than reacting to sudden, unpredictable events. The contrast between expected and unexpected floods supports the argument that expected flood risk, rather than the occurrence of unexpected floods, is the primary driver of interregional innovation collaboration.

Table 3.7: Impacts of Flood Expectation on Patent Collaborations - Patent Level Analysis

	Collaborative Patent Dummy			
	(1)	(2)	(3)	(4)
Panel A: Independent Variable - Expected Cumulative Flood Duration (ECFD)				
ln(ECFD)	0.031***			
	(0.001)			
ln(Weighted ECFD)		0.034***		
		(0.001)		
Panel B: Independent Variable - Expected Pixel-Adjusted Flood Duration (EPFD)				
EPFD			0.221***	
			(0.032)	
Weighted EPFD				0.164***
				(0.025)
R-squared	0.408	0.407	0.393	0.393
N(obs)	452,734	452,734	452,734	452,734
Fixed Effects				
Year	Y	Y	Y	Y
County 1	Y	Y	Y	Y
County 2	Y	Y	Y	Y
Patent Type	Y	Y	Y	Y

Note: (1) 'ECFD' refers to Expected Cumulative Flood Duration, and 'EPFD' refers to Expected Pixel-Adjusted Flood Duration, and the detailed construction can be found in Section 3.5.2.3; (2) The regression takes the form of $DCollab_{ijt} = \alpha + \beta_1 ExpectedFloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ where $ExpectedFloodProxy$ is either Expected Cumulative Flood Duration or Pixel-Adjusted Flood Duration; (3) standard errors are clustered at the county level.

Table 3.8: Impacts of Flood Expectation on Patent Collaborations - County Pair Level Analysis

	ln(Number of Collaborative Patents)			
	(1)	(2)	(3)	(4)
Panel A: Independent Variable - Expected Cumulative Flood Duration (ECFD)				
ln(ECFD)	0.008***			
	(0.003)			
ln(Weighted ECFD)		0.010***		
		(0.003)		
Panel B: Independent Variable - Expected Pixel-Adjusted Flood Duration (EPFD)				
EPFD			0.197	
			(0.149)	
Weighted EPFD				0.253***
				(0.105)
R-squared	0.434	0.434	0.434	0.434
N(obs)	65,828	65,828	65,828	65,828
Fixed Effects				
Year	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y
County 1	Y	Y	Y	Y
County 2	Y	Y	Y	Y

Note: (1) 'ECFD' refers to Expected Cumulative Flood Duration, and 'EPFD' refers to Expected Pixel-Adjusted Flood Duration, and the detailed construction can be found in Section 3.5.2.3; (2) The regression takes the form of $Y_{ijt} = \alpha + \beta_1 ExpectedFloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $ExpectedFloodProxy$ is either Expected Cumulative Flood Duration or Pixel-Adjusted Flood Duration; (3) standard errors are clustered at the county-pair level.

Table 3.9: Impacts of Unexpected Floods on Patent Collaborations - Patent Level Analysis

	Collaborative Patent Dummy			
	(1)	(2)	(3)	(4)
Panel A: Independent Variable - Unexpected Cumulative Flood Duration (UCFD)				
ln(UCFD)	−0.010***			
	(0.000)			
ln(Weighted UCFD)		−0.009***		
		(0.000)		
Panel B: Independent Variable - Unexpected Pixel-Adjusted Flood Duration (UPFD)				
UPFD			−0.015*	
			(0.009)	
Weighted UPFD				−0.015*
				(0.008)
R-squared	0.396	0.396	0.393	0.393
N(obs)	452,734	452,734	452,734	452,734
Fixed Effects				
Year	Y	Y	Y	Y
County 1	Y	Y	Y	Y
County 2	Y	Y	Y	Y
Patent Type	Y	Y	Y	Y

Note: (1) 'UCFD' refers to Unexpected Cumulative Flood Duration, and 'UPFD' refers to Unexpected Pixel-Adjusted Flood Duration, and the detailed construction can be found in Section 3.5.2.3; (2) The regression takes the form of $DCollab_{ijt} = \alpha + \beta_1 UnexpectedFloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ where $UnexpectedFloodProxy$ is either Unexpected Cumulative Flood Duration or Unexpected Pixel-Adjusted Flood Duration; (3) standard errors are clustered at the county level.

Table 3.10: Impacts of Unexpected Floods on Patent Collaborations - County Pair Level Analysis

	ln(Number of Collaborative Patents)			
	(1)	(2)	(3)	(4)
Panel A: Independent Variable - Unexpected Cumulative Flood Duration (UCFD)				
ln(UCFD)	−0.003**			
	(0.002)			
ln(Weighted UCFD)		−0.004**		
		(0.002)		
Panel B: Independent Variable - Unexpected Pixel-Adjusted Flood Duration (UPFD)				
UPFD			−0.006	
			(0.049)	
Weighted UPFD				−0.032
				(0.040)
R-squared	0.434	0.434	0.434	0.434
N(obs)	65,828	65,828	65,828	65,828
Fixed Effects				
Year	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y
County 1	Y	Y	Y	Y
County 2	Y	Y	Y	Y

Note: (1) ‘ECFD’ refers to Unxpected Cumulative Flood Duration, and ‘EPFD’ refers to Unxpected Pixel-Adjusted Flood Duration, and the detailed construction can be found in Section 3.5.2.3; (2) The regression takes the form of $Y_{ijt} = \alpha + \beta_1 UnexpectedFloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $UnexpectedFloodProxy$ is either Unexpected Cumulative Flood Duration or Unexpected Pixel-Adjusted Flood Duration; (3) standard errors are clustered at the county-pair level.

3.5.3 Testing the Mutual-Support Hypothesis

We then focus exclusively on counties that have engaged in cross-county collaborations, which can be classified into three types: Type A (flooded and flooded), Type B (flooded and non-flooded), and Type C (non-flooded and non-flooded). We use county pairs in Type C as the control group, as they provide a benchmark for collaboration patterns in the absence of flood exposure.

In Columns (1) and (2), we compare the collaborative patterns of county pairs in which at least one county has experienced floods (Types A and B) with those in the control group (Type C). We find that a standard deviation increase in pixel-adjusted flood duration is associated with a 2.1% increase in collaborative patents, and a 3.5% increase in collaborative patents among counties that have all experienced floods. Finally, in Columns (5) and (6), we analyze the collaborative patterns of county pairs in which only one county has experienced floods (Type B), again using Type C as the control group. The smaller and less significant coefficients in these columns suggest that flood exposure does not substantially increase cross-county collaboration when only one county in the pair has experienced floods.

Table 3.11: Heterogeneous Impacts of Floods on Patent Collaborations - County Pair Level Analysis

	ln (Number of County-Pair Collaborative Patents)					
Sample:	Type A+B+C		Type A+C		Type B+C	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(CFD)	0.011***		0.018***		0.008**	
	(0.003)		(0.004)		(0.003)	
L_1(PFD)		0.120**		0.193**		0.089
		(0.059)		(0.086)		(0.086)
R-squared	0.434	0.434	0.438	0.438	0.438	0.438
N(obs)	65,828	65,828	58,760	58,760	62,410	62,410
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y	Y	Y
County 1	Y	Y	Y	Y	Y	Y
County 2	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of CFD (Cumulative Flood Duration) and PFD (Pixel-Adjusted Flood Duration) can be found in Section 3.4.1; (2) The regression specification is: $Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either ln(L_1(CFD))), the logarithm of lagged CFD (lagged by one year), or L_1(PFD), the lagged PFD (lagged by one year); (3) 'Type A' refers to collaborations between flooded counties, 'Type B' refers to collaborations between flooded and non-flooded counties, while 'Type C' refers to collaborations between non-flooded counties; (4) Through Column (1) - (6), the control group is C Type cross-county collaborations; (5) Standard errors are clustered at the county-pair level.

The results in Table 3.12 further reinforce the finding that shared flood experiences play a crucial role in fostering inter-county collaboration on patents. Across all specifications, we find that increased flood exposure is associated with a higher likelihood of patent collaborations. Notably, in columns (3) and (4), where we focus on collaborations exclusively among flooded counties, the effect is particularly strong — an additional standard deviation increase of pixel-adjusted flood exposure is associated with a 8.4% increase. This indicates that counties facing similar flood-related challenges are significantly more likely to engage in joint innovation efforts. In contrast, when we examine collaborations between flooded and non-flooded counties in columns (5) and (6), we still find positive and significant effects, though they are relatively smaller in magnitude. It suggests thatt while floods can facilitate collaboration across different types of counties, the strongest cooperative responses occur when both counties have directly experienced flooding.

Overall, we find that collaboration is more pronounced between counties that have both experienced floods, highlighting the critical role of shared flood experiences in fostering inter-county cooperation. Counties in which both members of a pair have faced flooding (Type A) exhibit significantly stronger collaborative responses compared to those where only one county has been affected (Type B). This pattern suggests that mutual exposure to flood risks creates a stronger foundation for joint collaboration. These findings indicate that floods have reshaped the geographical distribution of patent collaborations, particularly among counties with similar flood histories.

Table 3.12: Heterogeneous Impacts of Floods on Patent Collaborations - Patent Level Analysis

	Collaborative Patent Dummy					
Type of Collaboration:	No Restriction		among Flooded		Flooded and non-Flooded	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_1(\text{CFD}))$	0.013***		0.037***		0.067***	
	(0.001)		(0.001)		(0.001)	
$L_1(\text{PFD})$		0.058***		0.471***		0.229***
		(0.014)		(0.027)		(0.023)
R-squared	0.395	0.393	0.466	0.421	0.414	0.297
N(obs)	452,734	452,734	452,734	452,734	452,734	452,734
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County 1	Y	Y	Y	Y	Y	Y
County 2	Y	Y	Y	Y	Y	Y
Patent Type	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of CFD (Cumulative Flood Duration) and PFD (Pixel-Adjusted Flood Duration) can be found in Section 3.4.1; (2) The regression specification is: $DCollab_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{Cumulative Flood Duration}))$, the logarithm of lagged CFD (lagged by one year), or L_1(Pixel-Adjusted Flood Duration), the lagged PFD (lagged by one year); (3) 'among Flooded' refers to collaborations between flooded counties, 'Flooded and non-Flooded' refers to collaborations between flooded and non-flooded counties; (4) Standard errors are clustered at the county level.

The results in Table 3.13 suggest that counties tend to collaborate on disaster mitigation patents following flood events. When we restrict our analysis to Sample A, where we exclude collaborations between flooded and non-flooded counties, the effect on disaster mitigation patents becomes even more pronounced. In this subset, as shown in column (4), a standard deviation increase in cumulative flood duration corresponds to a roughly 1.2% increase in county-pair collaborative disaster-mitigation patents. These results suggest that counties with similar flood experiences are more likely to engage in joint innovation efforts to develop disaster-mitigation technologies. However, we do not find significant evidence that flood exposure fosters collaborations between flooded and non-flooded counties.

Table 3.13: Impact of Floods on Disaster Mitigation Patents
- County Pair Level Analysis

	ln (Number of County-Pair Collaborative Disaster-Mitigation Patents)					
Sample:	Type A+B+C		Type A+C		Type B+C	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(CFD)	0.004***		0.007***		0.002	
	(0.001)		(0.002)		(0.002)	
L_1(PFD)		0.015		0.067***		−0.027
		(0.018)		(0.022)		(0.028)
R-squared	0.006	0.005	0.031	0.030	−0.007	−0.007
N(obs)	65,828	65,828	58,760	58,760	62,410	62,410
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y	Y	Y
County 1	Y	Y	Y	Y	Y	Y
County 2	Y	Y	Y	Y	Y	Y

Note: (1) Disaster mitigation patents are those related to disaster alleviation and are identified using *China Meteorological Disaster Yearbook* and the *National Emergency Response Plan*; (2) Detailed definition of CFD (Cumulative Flood Duration) and PFD (Pixel-Adjusted Flood Duration) can be found in Section 3.4.1; (3) Disater-mitigation patents refer to patents that target disaster alleivation; (4) The regression specification is: $Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where Y_{ijt} represents the logarithm Number of County-Pair Collaborative Disaster-Mitigation Patents, $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{CFD})))$, the logarithm of lagged CFD (lagged by one year), or L_1(PFD), the lagged PFD (lagged by one year); (5) 'Type A' refers to collaborations between flooded counties, 'Type B' refers to collaborations between flooded and non-flooded counties, while 'Type C' refers to collaborations between non-flooded counties; (6) Through Column (1) - (6), the control group is C Type cross-county collaborations; (7) Standard errors are clustered at the county-pair level.

3.6 Conclusion

This paper demonstrates that floods can significantly reshape the geography of innovation. While severe floods reduce localized innovation in directly affected areas, they simultaneously stimulate inter-regional collaborations. Using detailed data on patent applications and satellite-derived flood exposure from 2008 to 2018 in China, we find strong evidence of a dual response: firms reduce local R&D activities but increasingly turn to external partners.

Our theoretical framework, grounded in a search-and-matching model, highlights two mechanisms driving this response. First, under the risk-sharing mechanism, firms collaborate across regions to hedge against localized productivity shocks caused by flooding. Second, the mutual-support mechanism shows that shared flood experiences foster cooperation in developing disaster-resilient technologies.

Empirical results validate the hypotheses generated by the theoretical framework. First, we find that cross-county collaborations increase significantly following flood events, especially among county pairs with similar flood histories. Importantly, this response is driven by expectations of future flood risk rather than by sudden, unexpected flood shocks. Second, we find that firms located in flooded counties tend to collaborate with each other to develop flood resilient patents.

Overall, our findings contribute to a growing literature on climate adaptation and innovation. They suggest that environmental shocks—while disruptive—can also trigger productive reallocation in innovation networks. By encouraging firms to collaborate more broadly across geographic boundaries, floods may lead to a spatially redistributed innovation networks.

Appendix A

Appendix of Chapter 1

A.1 Supplementary Materials of Research Background

A.1.1 List of Flood Detention Basins

国家法律法规数据库

欢迎您访问国家法律法规数据库

蓄滞洪区运用补偿暂行办法

扫码下

法律效力位阶：行政法规

制定机关：国务院

时效性：有效

施行日期：

公布日期：2000-05-27

WPS版本 公报原版

100%

附：

国家蓄滞洪区名录

长江流域：围堤湖、六角山、九垸、西官垸、安澧垸、澧南垸、安昌垸、安化垸、南顶垸、和康垸、南汉垸、民主垸、共双茶、城西垸、屈原农场、义和垸、北湖垸、集成安合、钱粮湖、建设垸、建新农场、君山农场、大通湖东、江南陆城、荆江分洪区、宛市扩建区、虎西备蓄区、人民大垸、洪湖分洪区、杜家台、西凉湖、东西湖、武湖、张渡湖、白潭湖、康山圩、珠湖圩、黄湖圩、方洲斜塘、华阳河。(共40个)

黄河流域：北金堤、东平湖、北展宽区、南展宽区、大功。(共5个)

海河流域：永定河泛区、小清河分洪区、东淀、文安洼、贾口洼、兰沟洼、宁晋泊、大陆泽、良相坡、长虹渠、白寺坡、大名泛区、恩县洼、盛庄洼、青甸洼、黄庄洼、大黄铺洼、三角淀、白洋淀、小滩坡、任固坡、共渠西、广润坡、团泊洼、永年洼、献县泛区。(共26个)

淮河流域：蒙洼、城西湖、城东湖、瓦埠湖、老汪湖、泥河洼、老王坡、蛟停湖、黄墩湖、南润段、邱家湖、姜家湖、唐垛湖、寿西湖、董峰湖、上六坊堤、下六坊堤、石姚湾、洛河洼、汤渔湖、荆山湖、方邱湖、临北段、花园湖、香浮段、潘村洼。(共26个)

- 8 -

版权所有 © 全国人大常委会办公厅

联系我们 flk@npc.gov.cn

建议使用谷歌Chrome65.0、火狐Firefox47.0、IE10及以上版本浏览器

说明

Figure A1: Flood Detention Basins in 2000 (Original Policy Document)

政府信息公开指南
政府信息公开制度
政府信息公开年报
法定主动公开内容

索 引 号：	111000/2010-00628	信息所属单位：	防御司
发布机构：	水利部	成文日期：	2010年01月07日
名　　称：	国家蓄滞洪区修订名录		
文　　号：	水汛[2010]14号	发布日期：	2010年01月12日

国家蓄滞洪区修订名录

字体：[大 中 小]

根据《蓄滞洪区运用补偿暂行办法》（2000年5月27日中华人民共和国国务院令第286号发布）规定,我部商财政部提出了国家蓄滞洪区名录修订意见并上报国务院。经国务院同意，现将《国家蓄滞洪区修订名录（2010年1月7日）》予以公布。

请你省（直辖市）和流域管理机构按照《蓄滞洪区运用补偿暂行办法》和有关规定，认真做好蓄滞洪区运用补偿的各项工作。

附件：

国家蓄滞洪区修订名录

(2010年1月7日)

长江流域：围堤湖、六角山、九垸、西官垸、安澧垸、澧南垸、安昌垸、安化垸、南顶垸、和康垸、南汉垸、民主垸、共双茶、城西垸、屈原农场、义和垸、北湖垸、集成安合、钱粮湖、建设垸、建新农场、君山农场、大通湖东、江南陆城、荆江分洪区、宛市扩大区、虎西备蓄区、人民大垸、洪湖分洪区、杜家台、西凉湖、东西湖、武湖、张渡湖、白潭湖、康山圩、珠湖圩、黄湖圩、方洲斜塘、华阳河、荒草二圩、荒草三圩、汪波东荡、蒿子圩。（共44处）

Figure A2: Flood Detention Basins in 2010 (Original Policy Document)

长江流域：围堤湖、六角山、九垸、西官垸、安澧垸、澧南垸、安昌垸、安化垸、南顶垸、和康垸、南汉垸、民主垸、共双茶、城西垸、屈原农场、义和垸、北湖垸、集成安合、钱粮湖、建设垸、建新农场、君山农场、大通湖东、江南陆城、荆江分洪区、宛市扩大区、虎西备蓄区、人民大垸、洪湖分洪区、杜家台、西凉湖、东西湖、武湖、张渡湖、白潭湖、康山圩、珠湖圩、黄湖圩、方洲斜塘、华阳河、荒草二圩、荒草三圩、汪波东荡、蒿子圩。（共44处）

黄河流域：北金堤、东平湖。（共2处）

淮河流域：蒙洼、城西湖、城东湖、瓦埠湖、老汪湖、泥河洼、老王坡、蛟停湖、黄墩湖、南润段、邱家湖、姜唐湖、寿西湖、董峰湖、汤渔湖、荆山湖、花园湖、杨庄、洪泽湖周边(含鲍集圩）、南四湖湖东、大逍遥。（共21处）

海河流域：永定河泛区、小清河分洪区、东淀、文安洼、贾口洼、兰沟洼、宁晋泊、大陆泽、良相坡、长虹渠、柳围坡、白寺坡、大名泛区、恩县洼、盛庄洼、青甸洼、黄庄洼、大黄铺洼、三角淀、白洋淀、小滩坡、任固坡、共渠西、广润坡、团泊洼、永年洼、献县泛区、崔家桥。（共28处）

松花江流域：月亮泡、胖头泡。（共2处）

珠江流域：湛江。（1处）

以上合计共98处。

此外，淮河流域的上六坊堤、下六坊堤、石姚湾、洛河洼、方邱湖、临北段、香浮段、潘村洼等8处蓄滞洪区虽不再列入国家蓄滞洪区名录，但在规划工程完工前，遇大洪水时若分洪运用，仍参照《蓄滞洪区运用补偿暂行办法》给予补偿。

Figure A3: Flood Detention Basins in 2010 (Original Policy Document), continued

Table A1: Flood Detention Basins in the Main River Basins of China (2000)

River Basin	Number of FDBs	Affected Population (million)	Total Area (km^2)	Storage Capacity (billion m^3)
Yangtze	40	6.12	11,959	63.6
Yellow	5	3.18	5,212	12.9
Hai	26	4.40	9,597	17.2
Huai	26	1.61	3,674	14.1
Total	97	15.3	30,443	107.7
% of China		*1.1%*	*0.3%*	

Note: (1) This table reports the number of FDBs, affected population, total FDB areas, and the storage capacity of FDBs in 2003; (2) '% of China' refers to the percentage of affected population to the whole population in China and the percentage of total area to the total area of China.

Table A2: Number of FDBs under 2000 and 2010 Policy

		FDBs Located in:			
Rivers	N(FDBs)	N(Provinces)	N(Municipalities*)	N(Cities)	N(Counties)
2000 Policy					
Yangtze	40	4	0	10	28
Hai	26	3	2	11	37
Huai	26	2	0	9	19
Yellow	5	2	0	6	12
Total	*97*	*8*	*2*	*36*	*96*
2010 Policy					
Yangtze	44	5	0	11	31
Hai	28	3	2	11	39
Huai	21	3	0	14	24
Yellow	2	2	0	5	8
Songhua	2	1	0	2	3
Zhu	1	1	0	1	1
Total	*98*	*11*	*2*	*44*	*106*
Δ(*2010-2000*)	*1*	*3*	*0*	*8*	*10*

Note: (1) The term '2000 Policy' refers to the National Flood Control Law implemented by China's Ministry of Water Resources in 2000, and '2010 Policy' to its subsequent update in 2010; (2) The 'Total' number might differ from the sum because some basins span multiple provinces, cities, and counties; (3) The term 'Municipalities*' denotes municipalities directly governed by China's Central Government, specifically Beijing and Tianjin in this study; (4) Under the 2000 Policy, provinces designated as Flood Detention Basin (FDB) regions included Hunan, Hubei, Anhui, Henan, Hebei, Shandong, Jiangxi, and Jiangsu. The 2010 Policy expanded this list to include Heilongjiang, Jilin, and Guangdong.

Table A3: Descriptive Statistics: FDB Counties and non-FDB Counties

Mean	Unit	FDB Counties	non-FDB Counties
N(Counties)		116	2,363
N(obs)		2,709	55,729
Geographical Factors:			
Slope		6.14	12.46
Elevation		45.24	561.28
N(Permanent Water Pixels)		1136.33	388.77
Floods:			
Size-Adjusted Flood Exposure	days	0.126	0.020
Size of Flood Inundation	pixels	5,024.44	679.98
Socio-Economic Variables:			
Population	thousands	853.41	632.80
Nighttime Light Intensity		1,676,066	1,259,737
Number of Firms		5,669.49	5,496.63

Note: (1) We use a county panel of 20 years (2000 - 2020); (2) Detailed introduction of data used in this research can be found in Section 1.3.1; (3) From 2000 to 2020, a total of 116 counties have been designated as FDB counties. In 2000, the government selected 96 FDB counties. In 2010, the government selected another 20 counties into the FDB list, but removed 10 from the list.

A.1.2 Example of FDB Implementation (Mengwa FDB)

To illustrate the function of FDBs, we look at flood management in the Huai River Basin (HRB). Located in the transition zone between the southern and northern climates of China, the Huai River Basin experiences dramatic climate changes, resulting in precipitation that varies both spatially and temporally. 70% of the precipitation is concentrated in the flood season from June to September. Due to the unique geographical condition of the HRB, flooding is frequent. For example, the HRB has seen floods in six years in the 1990s.

In 2007, a high-intensity rainfall hit the HRB and the average rainfall reached 465 mm. The precipitation led to multi-peak flooding in the Huaihe River and threatened the downstream areas of the Flood Detention Basin. When the water level reached 29.3m on July 10, the government raised the flood severity level to the highest and operated the Wangjiaba Dentention Basin. The basin diverted water for 46 hours and stored flood with a volume of 250 million cubic meters. Even though the downstream

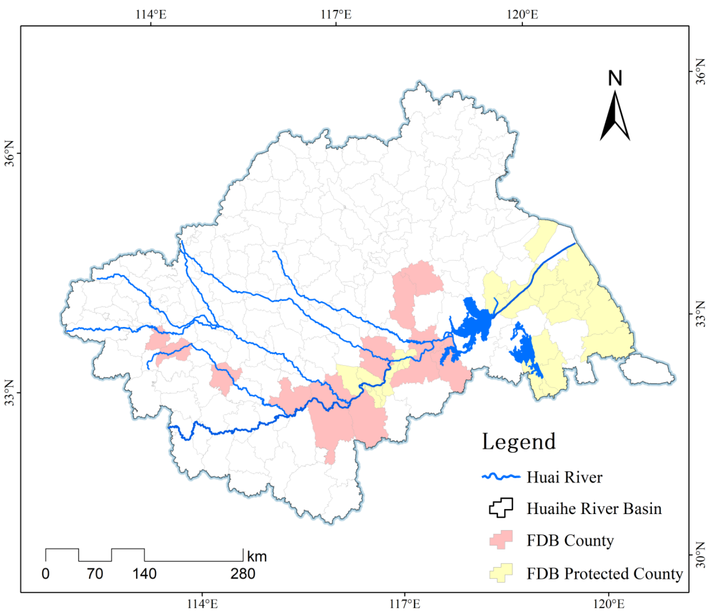


Figure A4: FDB Counties and FDB-Protected Districts in Huai River Basin

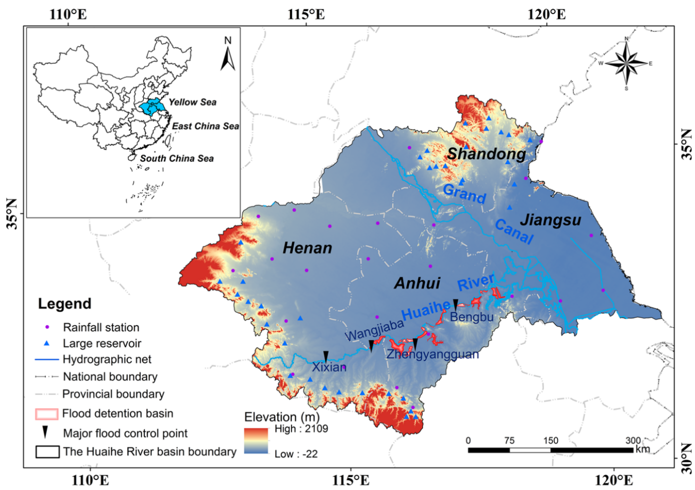


Figure A5: Wangjiaba Location (Source: Zhang and Song 2014)

land is protected, the use of Mengwa resulted in a forced migration of more than 3,000 people, an inundation of more than 12,000 hectares of farmland, and destruction of all Wangjiaba infrastructure. According to Chinese government, the 2007 flood affected around 2.5 million hectares of crops and caused a direct economic loss of around 2.5 billion USD, which is around 50 % less than the flood loss in 1991. The decrease in economic loss is largely contributed to the operation of FDBs.

Figure A6: Before and After Flood Water Diversion of Mengwa Flood Detention Basin

A.1.3 Empirical Analysis of FDB Selection

To understand determinants of FDB selections, we run a linear probability regression model:

$$FDB_{ict} = \alpha + \beta_1 Geo_{ict} + \beta_2 ln(Light)_{ict} + \gamma_c + \lambda_t + \varepsilon_{ict}$$

where FDB_{icut} is a dummy variable that equals 1 if the county i in city c is designated as an FDB county in 2000, and 0 otherwise. Geo_{ict} represents geographical controls (elevation, gradient, and precipitation), which are key factors that affect floods. $ln(Light)_{ict}$ represents the logarithm nighttime light intensity. γ_c, λ_t are city and time fixed effects, respectively. ε_{ict} is the standard error, that clustered at city level.

According to the Chinese government, FDBs are located in low-lying lands that are hydrologically feasible to collect flood water. Table A4 provides supportive evidence that the FDB selection is mainly based on geographical factors, especially, elevation. However, we do not find evidence that FDB selection is significantly correlated with economic factors, for example, nighttime light intensity.

Table A4: FDB Selection Criteria: Linear Probability Model

(in logarithm)	(1)	(2)	(3)	(4)	(5)
Elevation	−0.059***				−0.052***
	(0.017)				(0.015)
Gradient		−0.043***			0.011
		(0.010)			(0.025)
Precipitation			−0.003		0.000
			(0.005)		(0.005)
Nighttime Light				0.006*	0.007
				(0.004)	(0.004)
N(obs)	48,280	48,280	48,280	48,280	48,280
R^2	0.350	0.358	0.343	0.344	0.365
Fixed Effects					
Year	Y	Y	Y	Y	Y
City	Y	Y	Y	Y	Y

Note: (1) We use a county panel of 10 years (1990-2000); (2) The dependent variable is a dummy FDB_i that equals 1 if the county i has a Flood Detention Basin, and equals 0 if not; (3) All regressions control for city fixed effects and year fixed effect; (4) Standard errors are clustered at the city level.

A.1.4 Compensation

According to *Temporary Measures for the Use of Compensation in Flood Storage and Detention Areas*, for crops, specialized farming, and economic forests, compensation will be provided at 50-70%, 40-50%, and 40-50% of the average annual output value over the three years prior to the flood detention, respectively. For housing, compensation will be provided at 70% of the flood damage loss. For household agricultural machinery, draft animals, and major durable household goods, compensation will be provided at 50% of the flood damage loss. However, if the total registered value of household agricultural machinery, draft animals, and major durable household goods is less than 2,000 yuan, compensation will be provided at 100% of the flood damage loss. If the flood damage loss exceeds 2,000 yuan but is less than 4,000 yuan, compensation will be provided at 2,000 yuan.

However, compensation will not be provided if satisfying either conditions: (i)losses from flood damage caused by refusal to abandon farmland that should be abandoned, refusal to relocate when required by national regulations, or losses resulting from unauthorized farming or returning after abandoning farmland or relocation; (ii) losses from flood damage to housing built in violation of safety construction plans or schemes for the flood detention area; (iii) kosses from flood damage to household agricultural machinery, draft animals, and major durable household goods that could have been transferred according to relocation orders but were not.


国家法律法规数据库
欢迎您访问国家法律法规数据库

蓄滞洪区运用补偿暂行办法

扫码下

法律效力位阶：行政法规
制定机关：国务院
时效性：有效

施行日期：
公布日期：2000-05-27

WPS版本 公报原版

100%

(一)根据国家有关规定，应当退田而拒不退田，应当迁出而拒不迁出，或者退田、迁出后擅自返耕、返迁造成的水毁损失；

(二)违反蓄滞洪区安全建设规划或者方案建造的住房水毁损失；

(三)按照转移命令能转移而未转移的家庭农业生产机械和役畜以及家庭主要耐用消费品水毁损失。

第十二条 蓄滞洪区运用后，按照下列标准给予补偿：

(一)农作物、专业养殖和经济林，分别按照蓄滞洪前三

-3-

年平均年产值的50%—70%、40%—50%、40%—50%补偿，具体补偿标准由蓄滞洪区所在地的省级人民政府根据蓄滞洪后的实际水毁情况在上述规定的幅度内确定。

(二)住房，按照水毁损失的70%补偿。

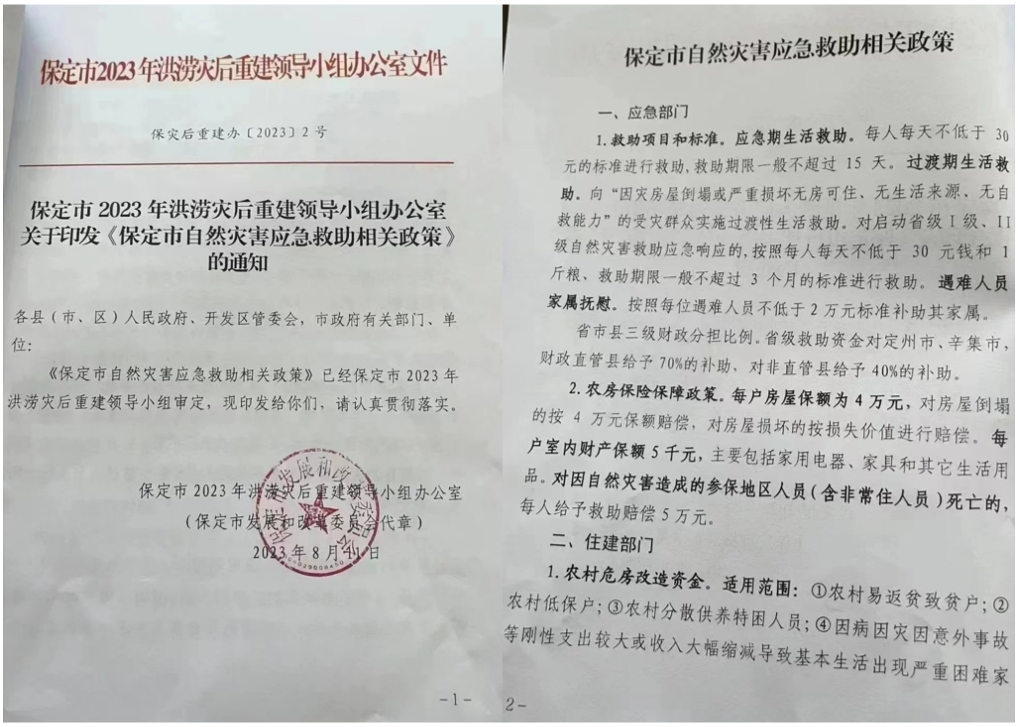
(三)家庭农业生产机械和役畜以及家庭主要耐用消费品，按照水毁损失的50%补偿。但是，家庭农业生产机械和役畜以及家庭主要耐用消费品的登记总价值在2000元以下的，按照水毁损失的100%补偿；水毁损失超过2000元不足4000元的，按照2000元补偿。

第十三条 已下达蓄滞洪转移命令，因情况变化未实施蓄滞洪造成损失的，给予适当补偿。

Figure A7: FDB Compensation According to *Temporary Measures for the Use of Compensation in Flood Storage and Detention Areas* (original policy document)

Zhuozhou was used for flood water diversion in 2023. According to the compensation regulation, each person will receive no less than 30 RMB (5 USD) per day for basic living assistance during the emergency period, which will last no more than 15 days. For those unable to meet their basic living needs due to a disaster, each person will receive no less than 30 RMB (5 USD) per day, for a period not exceeding 3 months. For those who need temporary relocation, each person will receive no less than 2,000 RMB

(300 USD) as standard assistance during the period of resettlement. For agricultural households, 70% of the cost will be compensated, and for non-agricultural households, 40% of the cost will be compensated. In the case of death due to a disaster in designated flood storage areas (including regular residents), each affected household will receive a compensation of 20,000 RMB (3,000 USD).



保定市2023年洪涝灾后重建领导小组办公室文件

保灾后重建办〔2023〕2号

保定市2023年洪涝灾后重建领导小组办公室关于印发《保定市自然灾害应急救助相关政策》的通知

各县（市、区）人民政府、开发区管委会，市政府有关部门、单位：

《保定市自然灾害应急救助相关政策》已经保定市2023年洪涝灾后重建领导小组审定，现印发给你们，请认真贯彻落实。

保定市2023年洪涝灾后重建领导小组办公室
（保定市发展和改革委员会代章）
2023年8月11日

-1-

保定市自然灾害应急救助相关政策

一、应急部门

1. **救助项目和标准。应急期生活救助。**每人每天不低于30元的标准进行救助，救助期限一般不超过15天。**过渡期生活救助。**向"因灾房屋倒塌或严重损坏无房可住、无生活来源、无自救能力"的受灾群众实施过渡性生活救助。对启动省级I级、II级自然灾害救助应急响应的，按照每人每天不低于30元钱和1斤粮、救助期限一般不超过3个月的标准进行救助。**遇难人员家属抚慰。**按照每位遇难人员不低于2万元标准补助其家属。

省市县三级财政分担比例。省级救助资金对定州市、辛集市，财政直管县给予70%的补助，对非直管县给予40%的补助。

2. **农房保险保障政策。每户房屋保额为4万元，**对房屋倒塌的按4万元保额赔偿，对房屋损坏的按损失价值进行赔偿。**每户室内财产保额5千元，**主要包括家用电器、家具和其它生活用品。**对因自然灾害造成的参保地区人员（含非常住人员）死亡的，**每人给予救助赔偿5万元。

二、住建部门

1. **农村危房改造资金。适用范围：**①农村易返贫致贫户；②农村低保户；③农村分散供养特困人员；④因病因灾因意外事故等刚性支出较大或收入大幅缩减导致基本生活出现严重困难家

2-

Figure A8: Actual FDB Compensation for Flood Detention in Baoding, Hebei Province in 2023 (original policy document)

A.2 Supplementary Materials of Economic Costs on FDB Counties

A.2.1 Supplementary Materials of Synthetic-DiD Results

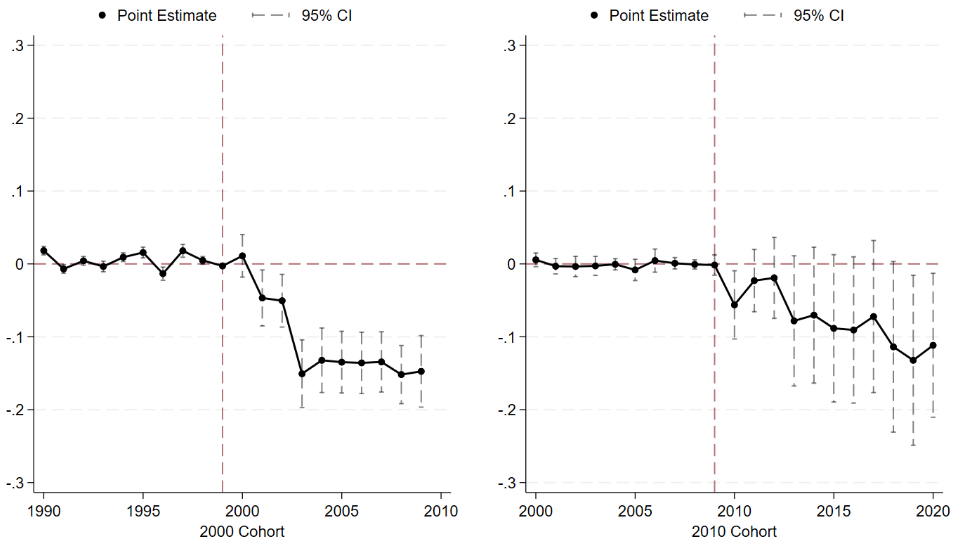


Figure B1: Dynamic Impacts of 2000 and 2010 FDB Policy Change on Light Intensity
Note: (1) Each dot represents the policy effect (ATT) estimated using the SDID event-study approach by Arkhangelsky et al. (2021)); (2) Data: 1990-2020 Nighttime Light Intensity data; (3) 96 counties were selected into the FDB list in 2000, while 20 counties were selected into the FDB list in 2010; (3) The event-study regression includes county and year fixed effects; (4) Standard Error: Bootstrap.

A.2.2 Individual Outcomes

A.2.2.1 Data Source: China Family Panel Studies (CFPS)

The China Family Panel Studies (CFPS) is a nationally representative, biennial longitudinal survey initiated in 2010 by the Institute of Social Science Survey (ISSS) at Peking University. This survey is designed to capture individual-, family-, and community-level data across a broad range of topics in contemporary China. It provides rich insights into both economic and non-economic aspects of well-being, covering areas such as economic activities, education outcomes, family dynamics, migration, and health. Funded by the Chinese government through Peking University, the CFPS aims to offer the academic community one of the most comprehensive and high-quality datasets available on modern China.

A.2.2.2 Empirical Strategy

To compare the individual income in FDB and non-FDB counties, we conduct the regression below:

$$ln(income)_{icjt} = \alpha + \beta_1 FDB_{icjt} + \gamma_j + \lambda_t + \varepsilon_j$$

where $ln(income)_{icjt}$ indicates the logarithm income of individual i residing in county c and city j, in year t, FDB_{icjt} is a dummy variable that equals 1 if the county c is an FDB county in year t, and 0 if not, $gamma_j$ is city fixed effect, λ_t is time fixed effect, standard errors are clustered at the city level.

Here, β_1 measures the difference in individual income in FDB counties and other counties. If it is negative, then individual income in FDB counties is lower than other counties, holding city and year constant. Note that we are not presenting a casual result because we do not have data before 2010 (the treatment year).

A.2.2.3 Result

Table B1 shows that individual income is lower in FDB counties, further supporting the argument that these counties bear long-term economic costs, as we present in Section 1.5. Specifically, Columns (2) and (4) indicate that, after controlling for key socio-economic factors, individuals in FDB counties earn approximately 18% less than those in other counties within the same city and year. However, due to data limitations, our analysis is based on residents from only six FDB counties. A comprehensive understanding of the living condition of residents in FDB counties require better individual-level data.

Table B1: Impacts of FDB Policy on Individual Income

	Sample Selection			
	All Counties		Excluded Spillover Counties	
Outcome: ln(Income)	(1)	(2)	(3)	(4)
FDB	-0.249^{***}	-0.176^{***}	-0.249^{***}	-0.175^{***}
	(0.007)	(0.006)	(0.008)	(0.006)
N(Obs)	70,652	70,652	68,853	68,853
N(FDB Residents)	3,012	3,012	3,012	3,012
N(Provinces)	25	25	25	25
N(Cities)	127	127	123	123
N(Counties)	162	162	158	158
N(FDB Counties)	6	6	6	6
Controls	N	Y	N	Y
Fixed Effects				
Year	Y	Y	Y	Y
City	Y	Y	Y	Y

Note: (1) Data source: 2010, 2012, 2014, 2016, 2018 and 2020 China Family Panel Studies (CFPS); (2) This table presents results of fixed-effect regression: $ln(income)_{icjt} = \alpha + \beta_1 FDB_{icjt} + \gamma_j + \lambda_t + \varepsilon_j$, $ln(income)_{icjt}$ indicates the logarithm income of individual i residing in county c and city j, in year t, FDB_{icjt} is a dummy variable that equals 1 if the county c is an FDB county in year t, and 0 if not, $gamma_j$ is city fixed effect, λ_t is time fixed effect, standard errors are clustered at the city level; (3) 'Spillover Counties' refers to those counties geographically adjacent to FDB counties; (4) Controls: age, married, gender, year of education, and urban status.

A.2.3 Robustness and Placebos

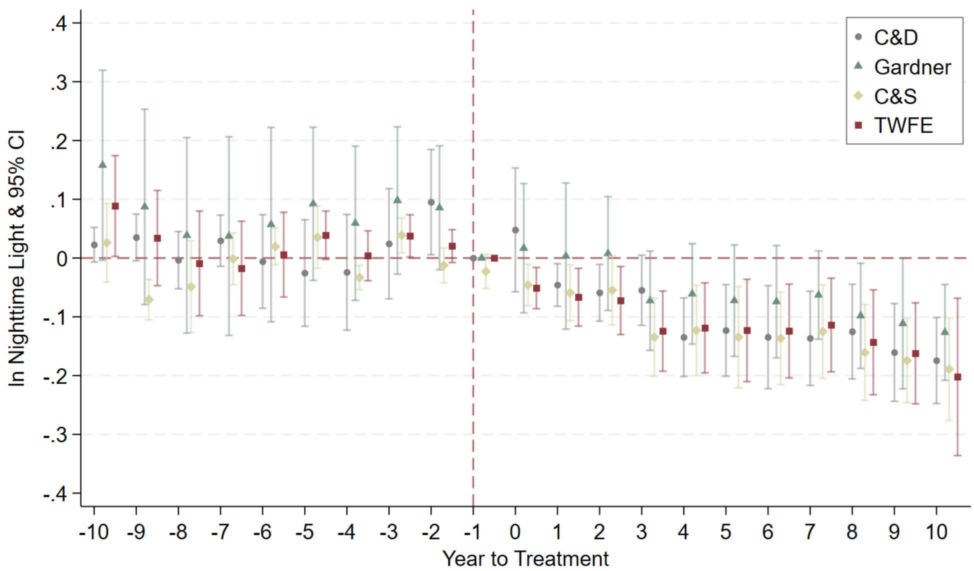


Figure B2: Event Study Robustness Check

Note: (1) Each dot represents the policy effect (ATT) estimated using different event-study approach; (2) 'TWFE' represents the traditional two-way-fixed-effects approach, 'C&D' refers to the two-way fixed effects estimators with heterogeneous treatment effects proposed by de Chaisemartin and D'Haultfœuille (2020), 'Gardner' refers to the two-stage DID approach by Gardner (2022), 'C&S' refers to the DID with multiple time periods by Callaway and Sant'Anna (2021); (3) Data: 1990-2020 Nighttime Light Intensity data; (4) 96 counties were selected into the FDB list in 2000, while 20 counties were selected into the FDB list in 2010; (5) The event-study regression includes county and year fixed effects, standard errors are clustered at county level; (6) We report the confidence interval at 95% confidence level.

Table B2: Robustness Check using Different DID Methods

	TWFE (1)	SDID (2)	C&D (3)	Gardner (4)	C&S (5)
FDB	-0.176^{***} (0.056)	-0.156^{***} (0.025)	-0.115^{***} (0.030)	-0.182^{***} (0.064)	-0.147^{***} (0.040)
N(obs)	70,463	70,463	70,463	70,463	70,463
Fixed Effects					
Year	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y

Note: (1) Each point estimate represents the policy effect (ATT) estimated using different difference-in-differences (DID) approach, 'TWFE' represents the traditional two-way-fixed-effects approach, 'SDID' refers to the synthetic DID proposed by Arkhangelsky et al. (2021), 'C&D' refers to the two-way fixed effects estimators with heterogeneous treatment effects proposed by de Chaisemartin and D'Haultfœuille (2020), 'Gardner' refers to the two-stage DID approach by Gardner (2022), 'C&S' refers to the DID with multiple time periods by Callaway and Sant'Anna (2021); (2) Data: 1990-2020 Nighttime Light Intensity data; (3) 96 counties were selected into the FDB list in 2000, while 20 counties were selected into the FDB list in 2010; (3) All regressions includes county and year fixed effects, standard errors are clustered at county level in Column (1), and (3) - (5), standard errors in Column (2) is set to be bootstrap; (4) The selection of control group is consistent with Column (1) in Table 1.2.

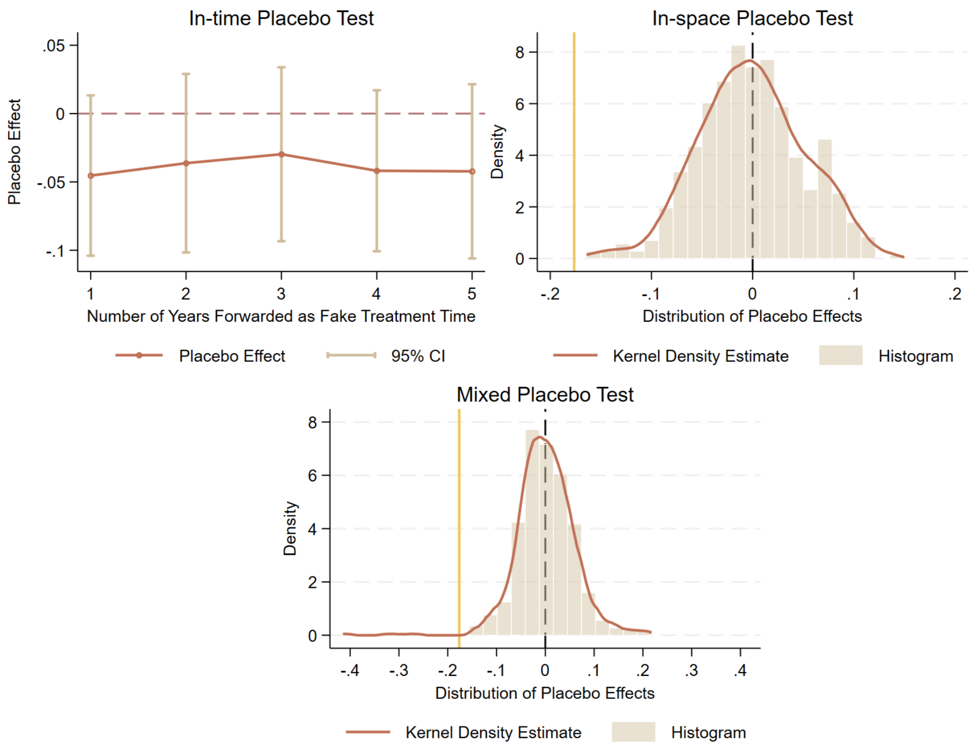


Figure B3: Placebo Test

Note: (1) This figure presents results of three distinct types of placebo tests of the traditional TWFE DID: the in-time placebo test, the in-space placebo test, and the mixed placebo test; (2) In the in-time placebo tests, we forward the treatment time by several years, using fake treatment times to assess if our results are driven by temporal trends rather than the actual intervention; (3) For the in-space placebo tests, we assign treatment to randomly selected units that did not receive the intervention, testing the robustness of our findings against spatial confounding factors; (4) The mixed placebo tests combine both approaches by randomly assigning fake treatment units and times.

Appendix B

Appendix of Chapter 2

B.1 Supplementary Materials of General Equilibrium Framework

B.1.1 An Illustrative Partial Equilibrium Model

We begin by concretizing the 'firm response effect' using an illustrative partial equilibrium model. While our comprehensive general equilibrium model accounts for interactions between different counties by including the flow of capital and manufacturing goods, this simpler model offers more straightforward economic intuitions regarding the trade-off between equality and efficiency in designing this flood risk redistribution policy. We then extend our analysis to the full general equilibrium model, which we use for counterfactual scenarios and to assess the benefit to cost ratio of FDB policies.

Flood Risk and Firm Investment Decision

In a two-period model, we assume that there are two types of counties, $i = s, p$. County s represents FDB counties that are sacrificed for protecting other counties, county p represents counties that are protected by FDB counties.

In period 1, the risk-neutral investor is endowed with an initial wealth, W, that can be used for consumption, investments in different counties, and investment in bonds. In period 2, investors consume the investment returns from the first period. The optimization problem is characterized below.

$$\max_{c_0,c_1,a_s,a_p,b} c_0 + \beta \mathbb{E}_\mu c_1$$

$$s.t. \quad c_0 + \sum_{i=s,p} a_i + b = W$$

$$c_1 = \sum_{i=s,p} (1+r_i)a_i + (1+r_f)b$$

Here, c_0 and c_1 represent the consumption at period 1 and period 2, respectively. a_s and a_p represent investors' period-one investment in sacrificed county and in protected county. b represents the bond investment. r_i is the return of assets, or the marginal benefit of investing in assets. r_f is the risk-free interest rate.

The production problem is characterized as:

$$\max_{k_i} \quad z_i k_i^\alpha - \bar{r}_i k_i$$

Here, k_i is the capital input in county i. z_i represents the productivity in county i. $\bar{r}_i$ represents the effective cost of investment in county i.

At each flood event, $\mu = \{\tau_s, \tau_p\}$, where τ_i is a dummy that equals 1 if the county is flooded at the flood event, and 0 if not. We consider flood as independent event in two types of counties. The flood probability of each county is $Pr(\tau_i = 1) = p_i$. Flood event will create a wedge between return of asset, r_i, and effective cost of investment, $\bar{r}_i$ such that

$$r_i = \bar{r}_i - \tau_i d$$

where d is the damage per asset caused by flood. Here, we assume that flood will cause proportional damages per asset that are identical across sacrificed and protected county.

The market clearing condition requires $r_{i,t}$ to clear the local capital market such that:

$$k_i = a_i$$

Following the above conditions, the optimal investment can be characterized as below:

$$\alpha z_i a_i^{\alpha-1} - r_f = p_i d$$

We can also consider the optimally condition as the characterization of flood risk premium. Here, the marginal product of capital is $MPK_i = \alpha z_i a_i^{\alpha-1}$. Hence, the difference between MPK_i and r_f can be interpreted as the flood risk premium, which equals the expected damage caused to the county i. Hence, the optimal investment a_i is determined by the flood probability p_i. Specifically, when flood probability increases, the amount of investment will decrease.

Impact of FDB Policy

We believe that the key function of FDB policy is to redistribute flood risk. To be more specific, the FDB policy aims to increase the flood risk in sacrificed county by Δp and decrease the flood risk in protected county by Δp. Hence, in sacrificed county, the FDB-adjusted flood probability will be $p'_s = p_s + \Delta p$. And in protected county, In protected county, the FDB-adjusted flood probability will be: $p'_p = p_p - \Delta p$. In Section 1.4, we find empirical evidence to confirm the validity of this assumption. Holding geographical conditions constant, we find that flood inundation area in sacrificed FDB counties is more than 50% higher, and the size adjusted flood exposure is around 5% higher in FDB counties (see Table 1.1).

Proposition 1 (Trade-off in Equality and Efficiency) *Assume* $\frac{z_p}{(p_p d + r_f)^{2-\alpha}} > \frac{z_s}{(p_s d + r_f)^{2-\alpha}}$, *then we have:* $\frac{d(a_p + a_s)}{dp} > 0$ *and* $\frac{d(a_p - a_s)}{dp} < 0$.

$\frac{z_p}{(p_p d + r_f)^{2-\alpha}} > \frac{z_s}{(p_s d + r_f)^{2-\alpha}}$ indicates that the damage standardized productivity in protected county is higher than that in sacrificed county. In other words, it specifies that a government that prioritizes efficiency has correctly identified counties worth to be protected.

The implication of this proposition are twofolds. First, FDB policy will bring an increase in total investment and will improve the economic resilience towards floods. The flood risk redistribution from protected to sacrificed counties will increase the total investment $a_p + a_s$. Second, FDB policy will also bring the inequality between sacrificed counties and protected counties because the investment gap $a_p - a_s$ will increase as well. We provide proof of this proposition in the Appendix B.1.2.

B.1.2 Proof of Proposition 1

Given flood event $\mu = \{\tau_s, \tau_p\}$, we can rewrite the investor's optimization problem in state-contingent form:

$$\begin{aligned} \max_{c_0, a_s, a_p, b, c_1(\mu)} \quad & c_0 + \beta \mathbb{E}_\mu c_1(\mu) \\ s.t. \quad & c_0 + \sum_{i=s,p} a_i + b = W \\ & c_1(\mu) = \sum_{i=s,p} (1 + r_i(\mu)) a_i + (1 + r_f) b \end{aligned}$$

The first-order conditions of the optimization problem yields the optimal asset positions $\{a_i\}_{i=s,p}$:

$$\sum_\mu Pr(\mu)[1 + r_i(\mu)] = 1 + r_f$$

where the actual investment returns $r_i(\mu)$ are determined by intrinsic capital productivity in the local area $\bar{r}_i$ and flood damage under event μ:

$$r_i(\mu) = \bar{r}_i - FloodDamage(\mu)$$

Plugging the actual investment return expressions into the Euler equation yields:

$$\bar{r}_i - r_f = \sum_\mu Pr(\mu) FloodDamage(\mu)$$

Assume that the county-specific events τ_i are independently distributed, then we get the pricing functions for county-specific assets $\{a_i\}_{i=s,p}$ are given by:

$$\bar{r}_i - r_f = p_i d \tag{B.1}$$

The intrinsic capital productivity of county i is given by the following optimization problem:

$$\max_{k_i} \quad z_i k_i^{\alpha} - \bar{r}_i k_i$$

Combined with market clearing conditions $k_i = a_i$, the intrinsic capital return $\bar{r}_i$ is given by:

$$\bar{r}_i = \alpha z_i a_i^{\alpha-1}$$

Plugging it into equation (21), it yields:

$$a_i = \frac{\alpha z_i}{r_f + p_i d}^{\frac{1}{1-\alpha}}$$

Consider a FDB policy that reallocates $dp > 0$ flood risk from protected county $dp_p = -dp$ to sacrificed county $dp_s = dp$. Assume that $\frac{z_p}{(p_p d + r_f)^{2-\alpha}} > \frac{z_s}{(p_s d + r_f)^{2-\alpha}}$. The impacts on aggregate capital investments and investment gap can be described by:

$$\frac{d(a_p + a_s)}{dp} = \frac{d}{1-\alpha}\left[\frac{\alpha z_p}{(r_f + p_p d)^{2-\alpha}}^{\frac{1}{1-\alpha}} - \frac{\alpha z_s}{(r_f + p_s d)^{2-\alpha}}^{\frac{1}{1-\alpha}}\right] > 0$$

$$\frac{d|a_p - a_s|}{dp} = \frac{d}{1-\alpha}\left[\frac{\alpha z_p}{(r_f + p_p d)^{2-\alpha}}^{\frac{1}{1-\alpha}} + \frac{\alpha z_s}{(r_f + p_s d)^{2-\alpha}}^{\frac{1}{1-\alpha}}\right] > 0$$

B.1.3 Direct Protection Effect

In Table A1, we first estimate the direct protection effect by running the regression

$$\text{lnLight}_{icpt} = \alpha + \beta_1 Flooded_{icpt} + \beta_2 Flooded \times FDB_{icpt} + \beta_3 Flooded \times Protected_{icpt} + X_{icpt} + \gamma_{pt} + \lambda_c + \varepsilon_c$$

where $lnLight_{icpt}$ is the ln(nighttime light intensity) of county i in city c, province p, at time t. $Flooded_{icpt}$ is a dummy variable that equals 1 if the county is flooded in year t, and 0 if not. FDB_{icpt} is a dummy variable that equals 1 if the county is an FDB county, and 0 if not. $Protected_{icpt}$ is a dummy variable that equals 1 if the county is an FDB-protected county, and 0 if not. X_{icpt} are controls. γ_{pt} is province-year fixed effect, η_t is time fixed effect, and λ_c is city fixed effect. ε_c is the standard error, which is clustered at the city level.

Following this specification, β_2 measures the impact of a county being designated as FDB county, while β_3 measures the impact of a county being protected by FDB counties. As shown in Table A1, we find that a protected county tends to suffer around 10% less when being hit by floods. However, an FDB county tends to suffer around 18% more when being hit by floods. This results indicates that FDB-protected counties are *direct*ly protected in flood events.

Table A1: Reduced Form: Direct Protection Effect

	ln(Nighttime Light Intensity)			
Flooded	−0.053**	−0.048*	−0.055*	−0.059*
	(0.027)	(0.027)	(0.031)	(0.032)
Flooded × FDB			−0.177*	−0.180*
			(0.092)	(0.067)
Flooded × Protected			0.105*	0.104*
			(0.061)	(0.067)
N(obs)	5,242	5,242	5,242	5,242
R^2	0.888	0.887	0.887	0.888
Fixed Effects				
Province-Year	Y	Y	Y	Y
City	Y	Y	Y	Y
Controls				
Demographic	Y	Y	Y	Y
Geographical	Y	N	Y	N

Note: (1) FDB is a dummy that equals 1 if the county i has once labeled as a Flood Detention Basin county, and equals 0 if not; (2) All regressions control for city fixed effects, province-by-year fixed effects, and a set of county-level controls (land area, population, and precipitation); (3) Standard errors are clustered at the county level.

B.1.4 Elasticity between Flood and Productivity

Table A2: Flood Impact on Productivity

(ln)	Total Productivity	Manufacturing Productivity
Size-adjusted Flooded Days	-0.043^{**}	-0.059^{*}
	(0.021)	(0.032)
N(obs)	1,283	1,283
Fixed Effects		
Year	Y	Y
City	Y	Y

Note: (1) FDB is a dummy that equals 1 if the county i has once labeled as a Flood Detention Basin county, and equals 0 if not; (2) All regressions control for city fixed effects, province-by-year fixed effects, and a set of county-level controls (land area, population, and precipitation); (3) Standard errors are clustered at the county level.

Appendix C

Appendix of Chapter 3

C.1 Supplementary Materials of Motivation

Figure A1: Centroid Change of Patent (Pink: 2008, Blue: 2013, Red: 2018)

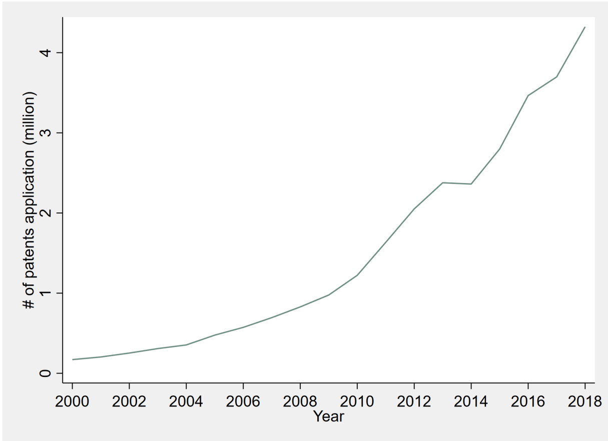


Figure A2: Patents Application by Year

C.2 Supplementary Materials of Data

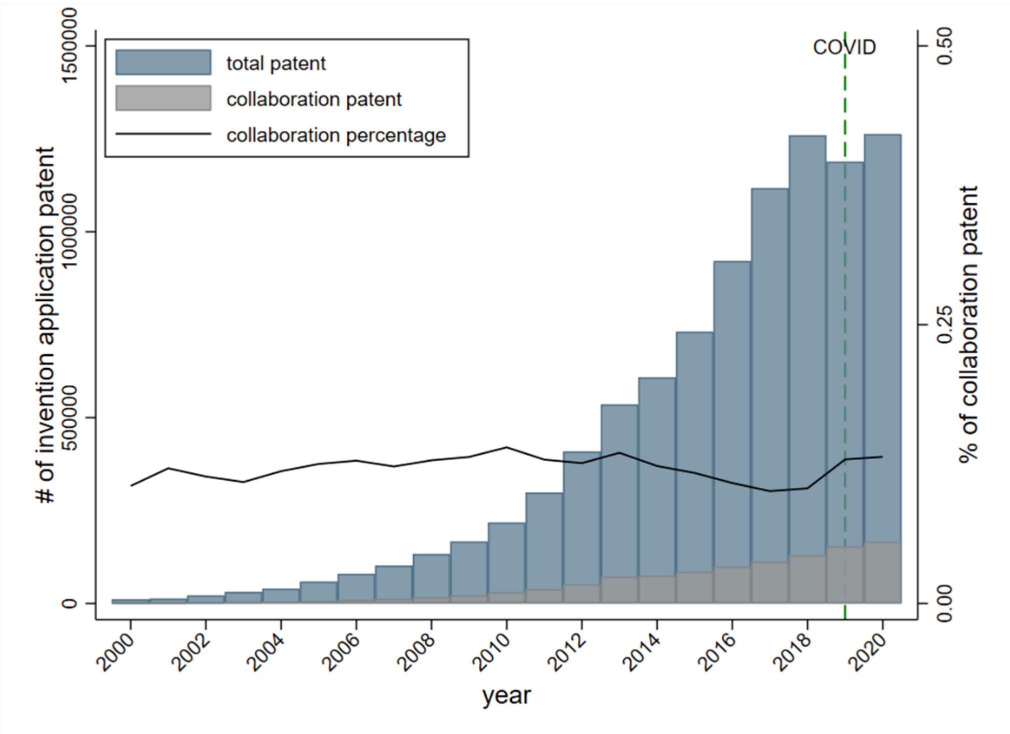


Figure B3: Total Patent Applications and Patent Collaborations (2000-2020)

C.3 Supplementary Materials of Empirical Results

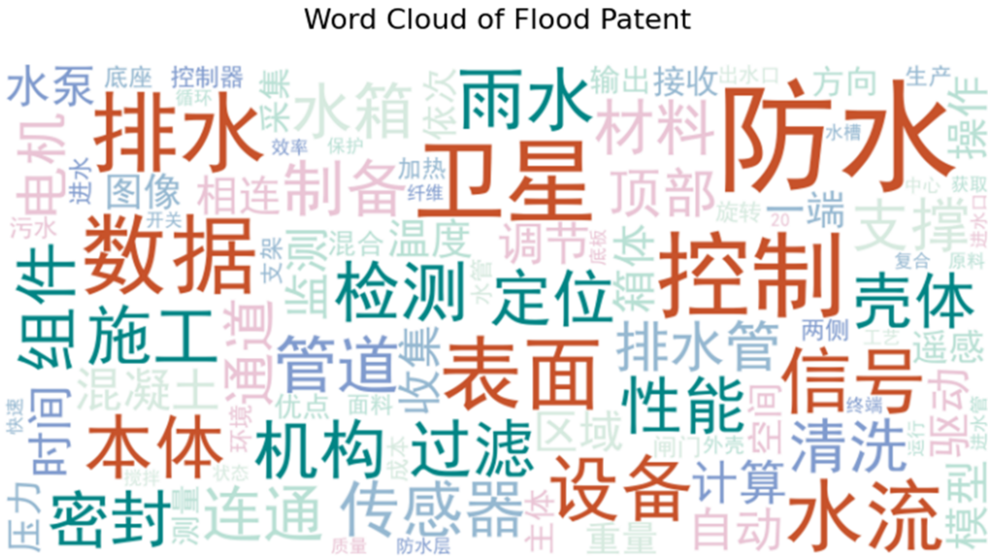


Figure C4: Keywords of Flood Patents
Note: The five most popular keywords of disaster related patents are: Water Proof, Drainage, Surface, Water Flow, and Control.

Table C1: Robustness Check: Flood Expectation and Patent Applications
- County Level Analysis

	ln(Number of Patent Applications)					
Type of Patents:	All		Single Applicant		Disaster Mitigation	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(ECFD)	−0.029***		−0.030***		0.003***	
	(0.004)		(0.004)		(0.001)	
ln(Weighted ECFD)		−0.032***		−0.034***		0.004***
		(0.005)		(0.005)		(0.001)
R-squared	0.864	0.864	0.856	0.856	0.582	0.582
N(obs)	28,867	28,867	28,867	28,867	28,867	28,867
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of Cumulative Flood Duration and Pixel-Adjusted Flood Duration can be found in Section 3.4.1; (2) The regression specification is: $DCollab_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1$(Cumulative Flood Duration)$)$, the logarithm of lagged CFD (lagged by one year), or L_1(Pixel-Adjusted Flood Duration), the lagged PFD (lagged by one year). (3) Standard errors are clustered at the county level.

Table C2: Robustness Check: Impacts of Floods on Patent Collaborations
- Patent Level Analysis (Changing Fixed Effects)

Dependent Variable:	Collaborative Patent Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_1(\text{CFD}))$	0.119***					
	(0.002)					
$L_1(\text{PFD})$		0.084**				
		(0.037)				
$\ln(\text{ECFD})$			0.106***			
			(0.002)			
EPFD				0.456***		
				(0.084)		
$\ln(L_1(\text{UCFD}))$					−0.067***	
					(0.001)	
$L_1(\text{UPFD})$						0.026
						(0.041)
R-squared	0.501	0.480	0.518	0.480	0.500	0.480
N(obs)	452,734	452,734	452,734	452,734	452,734	452,734
Fixed Effects						
County 1 × Year	Y	Y	Y	Y	Y	Y
County 2 × Year	Y	Y	Y	Y	Y	Y
Patent Type	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of Cumulative Flood Duration and Pixel-Adjusted Flood Duration can be found in Section 3.4.1; (2) The regression specification is: $DCollab_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{Cumulative Flood Duration}))$, the logarithm of lagged CFD (lagged by one year), or $L_1(\text{Pixel-Adjusted Flood Duration})$, the lagged PFD (lagged by one year). (3) Standard errors are clustered at the county level.

Table C3: Heterogeneous Impacts of Flood Expectation on Patent Collaborations - County Pair Level Analysis

	ln (Number of County-Pair Collaborative Patents)					
Sample:	Type A+B+C		Type A+C		Type B+C	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(ECFD)	0.008***		0.008**		0.006**	
	(0.003)		(0.003)		(0.003)	
ln(Weighted ECFD)		0.010***		0.010**		0.008**
		(0.003)		(0.004)		(0.036)
R-squared	0.434	0.434	0.438	0.438	0.438	0.438
N(obs)	65,828	65,828	58,760	58,760	62,410	62,410
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y	Y	Y
County 1	Y	Y	Y	Y	Y	Y
County 2	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of CFD (Cumulative Flood Duration) and PFD (Pixel-Adjusted Flood Duration) can be found in Section 3.4.1; (2) The regression specification is: $Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{CFD}))$, the logarithm of lagged CFD (lagged by one year), or L_1(PFD), the lagged PFD (lagged by one year); (3) 'Type A' refers to collaborations between flooded counties, 'Type B' refers to collaborations between flooded and non-flooded counties, while 'Type C' refers to collaborations between non-flooded counties; (4) Through Column (1) - (6), the control group is C Type cross-county collaborations; (5) Standard errors are clustered at the county-pair level.

Table C4: Heterogeneous Impacts of Unexpected Floods on Patent Collaborations - County Pair Level Analysis

	ln (Number of County-Pair Collaborative Patents)					
Sample:	Type A+B+C		Type A+C		Type B+C	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(UCFD)	−0.003**		−0.003**		0.002	
	(0.002)		(0.002)		(0.002)	
ln(Weighted UCFD)		−0.004**		−0.004**		−0.003
		(0.002)		(0.002)		(0.002)
R-squared	0.434	0.434	0.438	0.438	0.438	0.438
N(obs)	65,828	65,828	58,760	58,760	62,410	62,410
Fixed Effects						
Year	Y	Y	Y	Y	Y	Y
County-Pair	Y	Y	Y	Y	Y	Y
County 1	Y	Y	Y	Y	Y	Y
County 2	Y	Y	Y	Y	Y	Y

Note: (1) Detailed definition of CFD (Cumulative Flood Duration) and PFD (Pixel-Adjusted Flood Duration) can be found in Section 3.4.1; (2) The regression specification is: $Y_{ijt} = \alpha + \beta_1 FloodProxy_{ij,t} + \lambda_t + \gamma_{ij} + \eta_i + \pi_j + \varepsilon_{ijt}$ where $FloodProxy_{i,t-1}$ is either $\ln(L_1(\text{CFD})))$, the logarithm of lagged CFD (lagged by one year), or L_1(PFD), the lagged PFD (lagged by one year); (3) 'Type A' refers to collaborations between flooded counties, 'Type B' refers to collaborations between flooded and non-flooded counties, while 'Type C' refers to collaborations between non-flooded counties; (4) Through Column (1) - (6), the control group is C Type cross-county collaborations; (5) Standard errors are clustered at the county-pair level.

Bibliography

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *J. Am. Stat. Assoc.*, 105(490):493–505.

Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., and Wei, X. (2024). Land Security and Mobility Frictions*. *The Quarterly Journal of Economics*, 139(3):1941–1987.

Akcigit, U., Grigsby, J., Nicholas, T., and Stantcheva, S. (2022). Taxation and innovation in the twentieth century. *The Quarterly Journal of Economics*, 137(1):329–385.

Allen, T. and Arkolakis, C. (2022). The welfare effects of transportation infrastructure improvements. *The Review of Economic Studies*, 89(6):2911–2957.

Anderson, K. A. and Richards-Shubik, S. (2022). Collaborative production in science: An empirical analysis of coauthorships in economics. *Review of Economics and Statistics*, 104(6):1241–1255.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., and others (2021). Synthetic difference-in-differences. *American Economic Review*.

Arnell, N. W. and Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, 134:387–401.

Au, C.-C. and Henderson, J. V. (2006). Are chinese cities too small? *The Review of Economic Studies*, 73(3):549–576.

Balboni, C., Boehm, J., and Waseem, M. (2023). Firm adaptation in production networks: Evidence from extreme weather events in pakistan. Technical report, Working Paper.

Balboni, C. A. (2019). *In harm's way? Infrastructure investments and the persistence of coastal cities*. PhD thesis, London School of Economics and Political Science.

Basberg, B. L. (1987). Patents and the measurement of technological change: a survey of the literature. *Research policy*, 16(2-4):131–141.

Bennett, M. S. and Mays, L. W. (1985). Optimal design of detention and drainage channel systems. *Journal of Water Resources Planning and Management*, 111(1):99–112.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, page rdae007.

Boustan, L. P., Kahn, M. E., and Rhode, P. W. (2012). Moving to higher ground: Migration response to natural disasters in the early twentieth century. *American Economic Review*, 102(3):238–244.

Boustan, L. P., Kahn, M. E., Rhode, P. W., and Yanguas, M. L. (2020). The effect of natural disasters on economic activity in us counties: A century of data. *Journal of Urban Economics*, 118:103257.

Bradt, J. T. and Aldy, J. E. (2023). Private benefits from public investment in climate adaptation and resilience. *Unpublished Manuscript, July*.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946.

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., and Liu, Y. (2022). Global 1 km$\times$ 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *Scientific Data*, 9(1):202.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, 110(9):2964–2996.

Desmet, K., Kopp, R. E., Kulp, S. A., Nagy, D. K., Oppenheimer, M., Rossi-Hansberg, E., and Strauss, B. H. (2021). Evaluating the economic cost of coastal flooding. *American Economic Journal: Macroeconomics*, 13(2):444–86.

Duflo, E. and Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2):601–646.

Elliott, R. J., Strobl, E., and Sun, P. (2015). The local impact of typhoons on economic activity in china: A view from outer space. *Journal of Urban Economics*, 88:50–66.

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C., and Zidar, O. (2019). State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376.

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.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Georgic, W. and Klaiber, H. A. (2022). Stocks, flows, and flood insurance: A nationwide analysis of the capitalized impact of annual premium discounts on housing values. *Journal of Environmental Economics and Management*, 111:102567.

Gröger, A. and Zylberberg, Y. (2016). Internal labor migration as a shock coping strategy: Evidence from a typhoon. *American Economic Journal: Applied Economics*, 8(2):123–53.

He, G., Wang, S., and Zhang, B. (2020). Watering down environmental regulation in china. *The Quarterly Journal of Economics*, 135(4):2135–2185.

He, Z.-L., Tong, T. W., Zhang, Y., and He, W. (2018). A database linking chinese patents to china's census firms. *Scientific data*, 5(1):1–16.

Head, K., Mayer, T., and Thoenig, M. (2014). Welfare and trade without pareto. *American Economic Review*, 104(5):310–316.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American economic review*, 102(2):994–1028.

Hornbeck, R. and Naidu, S. (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review*, 104(3):963–90.

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

Hsiao, A. (2023). Sea level rise and urban adaptation in jakarta. *Technical Report*.

Hsiao, A. (2024). Sea level rise and urban inequality. In *AEA Papers and Proceedings*, volume 114, pages 47–51. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Hu, A. G., Zhang, P., and Zhao, L. (2017). China as number one? evidence from china's most recent patenting surge. *Journal of Development Economics*, 124:107–119.

Imbens, G. and Wager, S. (2019). Optimized regression discontinuity designs. *Review of Economics and Statistics*, 101(2):264–278.

Jia, R., Ma, X., and Xie, V. W. (2022). Expecting floods: Firm entry, employment, and aggregate implications. *working paper*.

Kocornik-Mina, A., McDermott, T. K., Michaels, G., and Rauch, F. (2020). Flooded cities. *American Economic Journal: Applied Economics*, 12(2):35–66.

Koh, Y., Li, J., and Xu, J. (2024). Subway, collaborative matching, and innovation. *Review of Economics and Statistics*, pages 1–18.

König, M., Storesletten, K., Song, Z., and Zilibotti, F. (2022). From imitation to innovation: Where is all that chinese r&d going? *Econometrica*, 90(4):1615–1654.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives*, 34(3):174–92.

Martinez, L. R. (2022). How much should we trust the dictator's gdp growth estimates? *Journal of Political Economy*, 130(10):2731–2769.

Mays, L. W. and Bedient, P. B. (1982). Model for optimal size and location of detention. *Journal of the Water Resources Planning and Management Division*, 108(3):270–285.

Mingkai, Q. and Kai, W. (2017). Flood management in china: The huaihe river basin as a case study. *Flood Risk Management*, pages 129–152.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10):3328–3375.

Mulder, P. (2021). Mismeasuring risk: The welfare effects of flood risk information. *URL https://faculty. wharton. upenn. edu/wp-content/uploads/2017/07/MismeasuringRisk_Mulder2021. pdf*.

Park, M. J., Ha, R., Kim, N. W., Lim, K. J., and Kim, S. J. (2014). Assessment of future climate and vegetation canopy change impacts on hydrological behavior of chungju dam watershed using swat model. *KSCE Journal of Civil Engineering*, 18:1185–1196.

Patel, D. (2023). Floods. *Available at SSRN 4636828*.

Perez-Pedini, C., Limbrunner, J. F., and Vogel, R. M. (2005). Optimal location of infiltration-based best management practices for storm water management. *Journal of water resources planning and management*, 131(6):441–448.

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. *Handbook of regional and urban economics*, 5:1339–1398.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244.

Strobl, E. (2011). The economic growth impact of hurricanes: Evidence from us coastal counties. *Review of Economics and Statistics*, 93(2):575–589.

Taur, C.-K., Toth, G., Oswald, G. E., and Mays, L. W. (1987). Austin detention basin optimization model. *Journal of Hydraulic Engineering*, 113(7):860–878.

Taylor, C. A. and Druckenmiller, H. (2022). Wetlands, flooding, and the clean water act. *American Economic Review*, 112(4):1334–1363.

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, 596(7870):80–86.

Wang, K., Wang, Z., Liu, K., Cheng, L., Bai, Y., and Jin, G. (2021). Optimizing flood diversion siting and its control strategy of detention basins: A case study of the yangtze river, china. *Journal of Hydrology*, 597:126201.

Wu, H., Lin, J., and Wu, H.-M. (2022). Investigating the real effect of china's patent surge: New evidence from firm-level patent quality data. *Journal of Economic Behavior and Organization*, 204:422–442.

Zeng, J. and Zhou, Q. (2024). Mayors' promotion incentives and subnational-level gdp manipulation. *Journal of Urban Economics*, 143:103679.

Zhang, L., Ren, Z., Chen, B., Gong, P., Xu, B., and Fu, H. (2024). A prolonged artificial nighttime-light dataset of china (1984-2020). *Scientific Data*, 11(1):414.

Zhang, X. and Song, Y. (2014). Optimization of wetland restoration siting and zoning in flood retention areas of river basins in china: A case study in mengwa, huaihe river basin. *Journal of Hydrology.*, 519:80–93.

Zhao, Y. (1999). Leaving the countryside: Rural-to-urban migration decisions in china. *American Economic Review*, 89(2):281–286.