Flooding and Firms in Indonesia*

Azhar Hussain†

March 13, 2025

Latest version here

Low- and middle-income countries, with their economic centers often located in vulnerable areas, are expected to bear the brunt of climate change impacts. Indonesia faces an elevated risk of disasters, with floods posing the most significant threat. In this paper, I first estimate the immediate effects of flooding on key economic variables, finding that more severe floods result in a larger reduction in aggregate economic indicators, reduced entry of new businesses, and a substantial depletion of firm-level capital stock. Some of these effects are likely driven by firms' evolving perceptions of flood risk in flood-prone areas. To examine these anticipatory effects of flooding, I develop a model of firms that incorporates flood risk and endogenous entry decisions. The analysis reveals that perceived flood risk, rather than actual flood events, has a more significant impact on firm behavior. While installing flood defenses in flood-prone regions could help mitigate these impacts, the resulting gains are diminished due to equilibrium adjustments and reduced firm selection effects on market entry.

JEL codes: D22, D84, E22, O13, Q54, R11

Keywords: Climate change, flood, flood risk, production, regional output, industry heterogeneity, low- and middle-income countries, Indonesia.

*I am grateful to Tim Besley, Gharad Bryan, Robin Burgess, and Daniel Sturm for their invaluable guidance and support on this project. I thank Clare Balboni, James Fenske, and Allan Hsiao for detailed discussions and comments on current and previous versions of this paper. The paper has also benefited from conversations with Sam Asher, Arun Chandrasekhar, Don Davis, Matthias Doepke, Dave Donaldson, Arnaud Dyèvre, Paolo Falco, Vernon Henderson, Namrata Kala, Rocco Macchiavello, Isabela Manelici, Guy Michaels, Esteban Rossi-Hansberg, Yannick Schindler, Javad Shamsi, and Silvana Tenreyro. Feedback and comments received from the participants at Essex PhD Conference, IFS-UCL-LSE Work-in-Progess Seminar, LSE Environment Day, Oxford Development Economics Workshop, and Transport, Energy and Climate Economics Working Group at Paris Dauphine further improved this paper. This work was supported by the Private Enterprise Development in Low-Income Countries under the Exploratory Research Grants program. I am responsible for the remaining errors.

†Department of Economics and STICERD, London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK. Email: a.hussain21@lse.ac.uk

1 Introduction

This paper examines the effects of flooding on firms in a low- and middle-income country, Indonesia. Given that the Global South is expected to experience heightened direct and indirect consequences of climate change (Cruz and Rossi-Hansberg 2021), the manifestation of such changes through extreme weather events is increasingly evident (IPCC 2023). Therefore, it is crucial to investigate how such events affect firms in different regions to better understand the response of production activities amid climate change. The insights gained through this investigation would be instrumental in designing appropriate adaptation strategies, such as reforming industrial zoning policies, for the changing world.

There is a negative contemporaneous relationship between flooding and measures of regional economic activity. When a region is hit by a flood, characterized by its spatial coverage and temporal extent, it shows an immediate reduction in aggregate value-added—a measure of economic output and capital stock and labor employment. This negative relationship is partly driven by the reduced entry of firms in flood-affected regions. At the firm level, flooding has a more pronounced negative effect on capital stock and positive effect on temporary labor hiring. Although this analysis sheds light on the costs associated with floods, it is unclear how floods interact with firm-level decision-making. In particular, due to the persistent nature of flood shocks in areas with high economic activity within Indonesia, the flood measure partially captures the evolving flood risk that firms perceive.

To address this, I introduce a quantitative framework where firms' input choices are influenced by different aspects of flooding. Specifically, firms choose their capital levels before flood shocks, taking flood risk into account, whereas labor is chosen afterward, once the actual flood shock has realized. The framework offers a novel microfoundation for understanding how perceived flood risk and actual flood shocks interact with firm behavior. The main findings indicate that perceived flood risk, rather than the occurrence of actual flooding, plays a more significant role in influencing firms' input decisions. I also conduct a counterfactual analysis in the spirit of building flood defenses to secure flood-prone areas. Building flood defenses has a direct positive impact on aggregate output, as protected areas become less vulnerable to flooding. However, these gains are partially offset by the entry of less productive firms into the now-safer areas, along with an upward pressure on equilibrium wages due to increased competition for scarce labor inputs.

The case in point is Indonesia, the world's tenth largest economy, which has maintained a high disaster risk profile mainly due to catastrophic flooding and accelerated sea-level rise affecting its major economic centers (World Risk Report 2023). As Figure 1 shows, flood events frequently affect areas in the southern islands of Java and

Sumatra, where a disproportionately large share ($> 90\%$) of manufacturing firms are located. This persistence of flood shocks potentially drives firms to update their perceived flood risk over time. Furthermore, as Figure 2 indicates, flooding is not just a recurring challenge but an escalating threat over time in Indonesia. Floods are the single most catastrophic natural disaster in terms of economic damage and human loss that Indonesia faces today (Government of Republic of Indonesia 2007). This confluence of factors makes Indonesia a prime case for examining the impacts of flooding.

In the first part of the paper, I provide reduced-form evidence on the contemporaneous effects of flooding on economic variables at both aggregate and firm levels using a static difference-in-differences research design. I propose and develop a regional flood index based on the spatial and temporal expanse of flooding. Since the index is continuous, I can analyze the effects by varying the intensity of flooding. I find that a 90th percentile flood leads to 20%, 25%, and 15% declines in aggregate value-added, capital stock, and labor employment, respectively, at the sector-region level. Estimating the same relationship at the firm level suggests that a 90th percentile flood is associated with a 5.7% reduction in the value of capital stock for a typical firm, with significant heterogeneity across sectors—more capital-intensive sectors, such as Iron and Steel, show more significant decline. Notably, the association of floods with firm exit is limited, while a firm's decision to enter a sector within a region is affected by floods in that region. Specifically, a 90th percentile flood is associated with 20% less firm entry at the sector-region level in the year of flooding. Given the spatial concentration of firms in areas prone to persistent flood shocks, these effects are partly driven by firms' evolving perceptions of flood risk.

The reduced-form findings yield estimates that capture both the actual damages from flooding and the adjustments that firms make in response to their evolving perception of flood risk over time. However, these findings do not provide a framework for understanding the mechanisms through which flooding influences firm behavior. In particular, capital installation decisions are typically made well in advance of the realization of flood shocks, and firms would anticipate these shocks and choose capital accordingly. In this context, the second part of the paper introduces a quantitative framework with flood risk and endogenous entry decisions to study the anticipatory effects of flooding. The time-varying parameters governing the regional flood exposure are estimated using the empirical distribution of firm-level production capacity utilization, which is negatively affected by flooding. The model builds on the seminal work by Lucas (1978) on understanding the impact of managerial talent on the distribution of firms. The foundational elements of the model find roots in more recent *misallocation* research, such as Hsieh and Klenow (2009) and Besley, Roland, and Reenen (2020), particularly using a general equilibrium framework to analyze firm behavior. The model transcends further by integrating firm entry and exit along the lines of Hopenhayn (1992).

Firms use a production technology that combines capital and labor inputs with firms' idiosyncratic productivity to produce output. Decisions regarding the amount of capital to install are made before the realization of flood shocks and take into account the uncertainty surrounding its utilization based on the flood risks associated with the firms' locations. Labor adjusts flexibly after the realization of flood shocks; however, it is indirectly influenced by flood risk through prior capital investment decisions. Risk-neutral firms maximize expected profits, where the expectations are based on the share of capital that can be utilized in a given year. This results in time-varying flood risk that differs across regions and sectors, acting as aggregate misallocation force that impacts capital allocation and ultimately output in equilibrium. Additionally, firms exhibit variations in their idiosyncratic productivity levels, which remain constant over time and are drawn from a common regional distribution. To enter a market, firms must incur a one-time fixed cost, making their entry decision contingent upon expected profits net of this fixed cost. This creates a productivity cutoff below which firms opt not to enter certain sectors within a region. This cutoff productivity, combined with labor market clearing, determines the mass of firms, their allocations, and the equilibrium wages in the market.

I estimate key model parameters to conduct quantitative analysis of the equilibrium. The novel regional shape parameters of the distribution of the share of capital utilized in a given year are estimated using the empirical distribution of production capacity utilization across firms located in a region. Production capacity utilization is the percentage of actual production over the planned production by a firm in a given year. Firms in flood-affected regions report lower production capacity utilization. Based on this finding, the share of capital that can be utilized in a given year is proxied by its production capacity utilization. Using such an objective measure that remains unaffected by prices and other short-run equilibrium adjustments ensures a more accurate assessment of the impact of flooding on firms. Additionally, the parameters governing the regional component of flood risk are strongly correlated with the empirical flood index used in the reduced-form analysis. I further estimate sector-specific production function parameters and the parameters that govern the distribution of firm productivity using standard methods from the literature.

In the analysis section, I start by disentangling the effects of flood risk and flood shocks on firm behavior. I employ firm-level equilibrium conditions for optimal capital and labor allocation to create a linear specification that can be estimated using ordinary least squares. The results indicate that firms reduce their capital investment and increase labor hiring in response to flood risk. In contrast, the impact of flood shocks—which affects equilibrium input allocations directly—is found to be limited. As an experimental counterfactual exercise in the spirit of flood defense systems used across the world, I compare observed outcomes with those generated after bringing the top 20th percentile of flood-prone regencies to the median value of the distribution by constructing flood defenses there. This intervention benefits all sectors and regions, but the benefits are larger for more capital-intensive sectors. The direct impact of the intervention increases annual aggregate output across sectors (regions) by 7% (16%). However, allowing for the entry of new firms reduces the gains by almost half, as less productive firms are now able to enter into these safer areas. The influx of new firms consumes scarce production resources, intensifying competition and ultimately driving wages upward. This underscores the potential downsides of such costly protective investments that are prevalent worldwide, particularly in low- and middle-income countries.

The paper relates to a strand of literature that studies the impact of climate change and natural disasters on the distribution of economic activity within and across regions (see, for example, Castro-Vincenzi 2024; Balboni 2024; Hsiao 2024; Nath 2024; Bilal and Rossi-Hansberg 2023; Desmet et al. 2021; Jia, Ma, and Xie 2022; Kocornik-Mina et al. 2020; Cruz and Rossi-Hansberg 2021; Balboni, Boehm, and Waseem 2023). I contribute to this literature in various ways. First, I employ a continuous measure of regional flooding that allows me to establish the relationship between flooding and economic variables at different flood intensities. Second, I provide a microfoundation for understanding how flooding affects firm decision-making by integrating the former within the firms' production function. I demonstrate how firms in low- and middle-income countries adjust their production inputs in response to threats posed by flooding. Specifically, they reduce investment in capital stock and increase temporary labor hiring, a more readily available resource in these settings, thus highlighting both production resilience and a form of adaptation to deal with such disruptions. Third, I develop a time-varying measure of flood risk, which is *perceived* by the firms. The literature has mostly employed time-invariant measures of flood risk, which is informed through atmospheric and hydrology models.[1] However, firms' decision-making is responsive to the actual flood events, and thus the perceived flood risk should evolve over time. Fourth, I examine the sectoral heterogeneity in the impact of flooding on both the intensive and extensive margins. On the methodological side, the paper is related to the literature on firm dynamics, misallocation, and their aggregate productivity effects (see, for example, Hsieh and Klenow 2009; Besley, Roland, and Reenen 2020; Midrigan and Xu 2014; Bento and Restuccia 2017; Bartelsman, Haltiwanger, and Scarpetta 2013; Hopenhayn 2014; Gopinath et al. 2017; Restuccia and Rogerson 2008; Banerjee and Duflo 2005)). I contribute to this literature by applying the methodology in the context of flooding, which generates distortions for

1. See, for example, Fathom Global Flood Map

firm-level capital decisions. I also integrate firm entry and exit dynamics along the lines of Hopenhayn (1992) and solve the model equilibrium analytically.

The remainder of the paper is organized as follows. The next section provides details on the data used. Section 3 presents the reduced-form findings on the contemporaneous effects of flooding. Section 4 develops the theoretical model for studying the effects of flood shocks and flood risk. Section 5 discusses the estimation of key model parameters, and Section 6 presents the analysis. Section 7 contains some concluding remarks.

2 Data

There are two main datasets used in the paper. The first is an extract of historical floods with various pieces of information to facilitate the construction of a flood index. The data on outcomes is derived from the census of medium and large manufacturing establishments located in Indonesia.

2.1 Large Flood Events

The data on floods is obtained from the Dartmouth Flood Observatory (DFO), which is a global, dynamic archive of large flood events starting in the year 1985 (Kocornik-Mina et al. 2020). The data provides start and end dates along with the extent of *affected area* for each flood event. Polygons representing the areas affected by flooding are drawn in a GIS program based upon information acquired from governmental, instrumental, news, and remote-sensing sources. Considering a longer time frame and the reliance on media reports, there could be concerns around plausible spatial and temporal bias in the reporting of flood events. For example, media reporting has improved over the years due to the development of technology and transportation infrastructure, and floods are more likely to be reported in areas with large population settlements and economic activity. In an extreme case, Figure 1 could merely reflect population settlement patterns across regencies in the Indonesian archipelago, rather than the actual number of flood events hitting them. To rule out this possibility, I redraw the map using only flood events confirmed through satellite observations, which are not subject to such biases. Specifically, I use inundation maps from 41 individual flood events across Indonesia during the period 2002–2018, as identified by Tellman et al. (2021). These maps are based on satellite imagery captured by the Moderate-Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra and Aqua satellites, which image the globe daily at a spatial resolution of 250 meters.[2] The spatial pattern

2. MODIS is well-suited for detecting large, slow-moving flood events but has limited capacity to resolve urban floods.

of flooding derived from these objective satellite-based measures aligns closely with the pattern depicted in Figure 1. This consistency is illustrated in Figure 3, which confirms the robustness of the observed spatial distribution of floods. Although 30% of the reported flood events in Indonesia last three days or less, media reporting could still miss some small flood events, particularly due to the low- and middle-income country setting (see, for example, Patel 2024, for estimates of such biases in Bangladesh). Following variables are then constructed from this data:

- *FloodAreaShare*$_{rt}$: It is the share of the flood-affected area of regency[3] r in year t. For a few cases where a regency witnessed multiple flooding episodes in a year, this is the average of all those flood-affected area shares. As discussed earlier, this variable captures the extent of geographic regions affected by a flood event, rather than just the extent of inundation. Flood-affected area is usually larger than the inundated area, and is a more relevant measure for studying the effect of flooding on economic activities. Using inundation maps of individual flood events from Tellman et al. (2021), I study the relationship between inundated area share and flood-affected area share at the regency level for these specific events using both non-parametric and parametric methods.[4] Table B.1 in the Appendix report measures of association between the two variables using two non-parametric methods and a regression analysis. Column 1 reports the Kendall's Tau-b coefficient and Column 2 reports the Somers' D coefficient, both of which range from -1 (perfect inversion) to +1 (perfect agreement), with 0 indicating no association. Clearly, the two variables are positively related. Using estimates of regression analysis reported in Column 3, a unit increase in flooded area increases the flood-affected area by 1.42 units.

- *FloodDaysShare*$_{rt}$: It is the share of days in year t that a regency r remains affected by flooding. It is calculated using the total duration (end date - start date + 1) of all flood events in a year. In most cases, these dates are derived from news reports. In a few cases for which beginning dates could not be determined, the

3. A regency is an administrative level-2 unit located within a province in Indonesia. As of 2020, there were 34 provinces containing a total of 522 regencies. However, many of these provinces and regencies were born out of administrative divisions among the existing ones through the years 1990-2010. Since the analysis starts in 1990, I merge some of these divisions to be representative of the 1990 administrative boundaries. Moreover, I drop four provinces on the Eastern islands namely, Papua, Papua Barat, Maluku, and Maluku Utara, as these provinces are sparsely populated by mainly indigenous tribes, and are predominantly engaged in activities such as forestry and fishing. After implementing all these changes, the analysis is representative of approximately 270 regencies that cover the entire economic map of the Indonesian archipelago.

4. Several areas within a regency could not be observed due to cloud cover on some days. In addition, to prevent misclassification of terrain shadows as water, areas with gradient larger than 5° were masked out in the maps. This means that the flooded area share is calculated out of the observed pixels, which do not necessarily cover the whole regency. Flooded areas would also be missed if they happen to lie within the masked areas.

starting dates are assumed to be the 15th day of the respective months.[5] Ending dates can either be exact—based on dates on which flood water starts to recede as per the news reports or estimated—based on a qualitative judgment concerning the flood event. Najibi and Devineni (2018) analyzed the issue of misreporting of duration using all flood events in the DFO catalogue for 1985–2015 time period. Their comparative analysis using the *in situ* streamflow observations obtained from the gauge stations suggests that the flood duration data from DFO is reliable and does not suffer from misreporting issues.

Figure B.1 in the Appendix shows the distribution of above variables by pooling all regency-year observations used in the analysis. Flood index is then generated by taking a simple product of the above two variables and rescaling the product by its maximum value so that it lies in the interval [0,1]. Thus, the index provides a measure of the intensity of floods by capturing both the spatial and temporal extents of each flood episode. Considering the spatial and temporal extent of flooding helps capture aspects of flood risk that both affected and unaffected firms may internalize in their decision-making. This approach acknowledges that firms, regardless of direct impact, could adjust their decisions based on the potential threat posed by flooding in their locations over time. As shown in Figure B.2 in the Appendix that includes all regency-year pairs (even those without floods), the flood index has a Pareto-like distribution with a long tail of extreme values. This is intuitive since extreme flood events are rare. Table B.2 reports the summary statistics on the flood index by utilizing only those regency-year pairs for which the index takes non-zero values. In the reduced-form analysis, the 25th, 50th, 75th, and 90th percentiles are used to capture and report the differential effects of flooding across various intensities. For example, at the 50th percentile, the estimated coefficients would represent the relationship between an outcome variable and a flood event at the median intensity level—an event more severe than 50% of all observed flood events in Indonesia from 1990 to 2012.

2.2 Information on Manufacturing Establishments

Data on the manufacturing establishments is obtained from the Annual Census of Medium and Large Manufacturing Establishments in Indonesia, also known as *Statistik Industri*. This data collection exercise was initiated by the Government of Indonesia in 1975 to survey all the manufacturing establishments with twenty or more workers annually. The central statistical agency, Statistics Indonesia, manages the collection and distribution of this data across different public departments and research organizations. Statistics Indonesia sends a questionnaire (asking details about previous

5. Only around 2.5% of the flood events that occurred in Indonesia during the 1985-2012 period have a starting date as the 15th day of a month.

year's operations), containing 150+ questions in a typical year, annually to all registered manufacturing establishments. In case of no response, field agents attempt to visit these establishments to either encourage compliance or confirm that the establishment has ceased operations (Blalock and Gertler 2008).

The establishment-level data includes information on industrial classification (5-digit ISIC), first year of commercial production, ownership structure, assets, income, output, value-added, expenses, capital stock, and other specialized information specific to a year for each establishment.[6] The main variables used in the analysis include measures on value-added, capital stock, labor employment, age (based on reported birth year), and location (regency where plant is located). All monetary variables are reported in nominal terms and are deflated using the wholesale price index at the 5-digit ISIC level to obtain their real values. Establishments are expected to report both market and book values of their capital stock, broken down by categories such as land, buildings, and equipment. However, book values are missing for the majority of observations, and not all categories of capital are consistently reported across time. Therefore, I use the market value of capital stock in all cases, unless it is unavailable and the book value is reported for those observations. All variables are winsorized at the 1% level on both the lower and upper tails each year to help mitigate potential measurement error concerns. To prevent compositional changes across variables from influencing the estimates, I include only plant-year observations with complete data on all three variables—value-added, total capital stock, and labor employment—following the final data cleaning step. Finally, I exclude all state-owned establishments, which represent less than 3% of total establishments in any given year, to avoid potential biases related to the implicit government insurance available to this group.

In a typical year, around 21,000 establishments are surveyed, with locations identified up to the regency where they are located. This establishment-level data is representative of firm-level analysis, as over 95% of surveyed establishments are single-branch entities. Therefore, hereafter, I use the term "firm" instead of establishment to refer to these manufacturing enterprises. Although these firms represent only about 2% of the total number of manufacturing firms operating in Indonesia in any given year, they contribute approximately 80% of the total value-added in the country's manufacturing sector. In terms of spatial distribution, these firms are primarily concentrated on the southern islands of Java and Sumatra, where flooding episodes are

6. Statistics Indonesia checks the reported values for inconsistencies and missing values and tries to make in-house corrections and imputations using the previous rounds of data before releasing it to the users. I do some additional data cleaning to match location identifiers consistently across years, impute some variables to correct for non-reporting in just one or two years, fix outliers identified at the firm and industry levels by interpolating between years, and fix a few obvious mistakes made in the data entry process.

also more frequent and intense (Asian Development Bank 2019).

3 Reduced-form Evidence

This section presents reduced-form results on the immediate impact of floods on aggregate and firm-level value-added, capital stock, and labor employment. It then examines the extensive margin by estimating the impact of flooding on firm entry and exit. Given the unique setting in this study where manufacturing hubs, primarily located in the flood-prone regions, contend with recurrent large flood events—it becomes challenging to isolate and interpret the long-run impact of individual flood occurrences without introducing potential biases.[7] Therefore, to examine the long-run consequences of flooding in such a context, one could use some kind of cumulative measure of flooding that *integrates* floods over an extended period of time. However, the focus of this paper is on estimating the contemporaneous effects of flooding and learning about firms' reaction towards regional flood risk. Therefore, one such analysis is included in the Appendix Section A where I estimate the effects of cumulative flood innovations on economic variables.

3.1 Effect of Flooding on Economic Variables

3.1.1 Econometric Model

I estimate the contemporaneous effects of flooding on aggregate (sector-regency) outcomes i.e., logarithm of total (labor share-weighted) value-added, capital stock, and labor employed at sector-regency level using the following specification:[8]

$$y_{srt} = \upsilon + \beta^J Flood^J_{rt} + \zeta_r + \nu_{st} + \varepsilon_{srt} \tag{III.1}$$

7. Figure B.3 in the Appendix shows that most of the regencies located on the islands of Java and Sumatra, which account for more than 95% of the manufacturing value-added, are affected by a large flood every alternate year. Figure B.4 in the Appendix shows the effect of the *first* flood on the logarithm of aggregate value-added, capital stock, and labor employed using the imputation-based difference-in-differences estimator proposed by Borusyak, Jaravel, and Spiess (2024). The estimated coefficients are not only less representative and far from what I would *want* to estimate but are also difficult to interpret. This is because the focal manufacturing regencies are not represented in the estimation of most of the dynamic effects coefficients as they get dropped too early because of the immediate second flooding episode.

8. It is implausible to check for the validity of canonical parallel trends assumption as those floods that struck some of these regencies before 1985 cannot be tracked. However, one can check if the growth rates across more and less flood-prone areas are significantly different. Figures B.5, B.6, and B.7 plot the logarithms of aggregate value-added, capital stock, and labor employment across low and high flood-prone regions. The plots suggest that high flood-prone regions did start with larger firms, but their year-on-year growth rates of value-added, capital stock, and labor employment measures are not significantly different from those of the less flood-prone ones.

where, y_{srt} is the logarithm of total (labor-share weighted) value-added, capital stock, or labor employment in sector s located in regency r in year t. $Flood^J_{rt}$ is a dummy variable that is assigned value 1 when the flood index in a regency-year exceeds the J^{th} percentile value for each $J \in \{25, 50, 75, 90\}$. Therefore, β^J captures the effect of J^{th} percentile flood on y_{srt}. ζ_r controls for time-invariant regency level characteristics, such as pre-existing differences in flood exposure and industrial settlements across regencies. ν_{st} controls for sectoral growth over time, where s denotes 2-digit ISIC sector. Where applicable, the outcome variables have been deflated by using 5-digit ISIC industry wholesale price index, and trimmed by 1% on both the tails for each year before being collapsed at the aggregate level.

Next, I conduct the firm-level estimation using the following specification:

$$y_{isrt} = \upsilon + \beta^J Flood^J_{rt} + \iota X_{isrt} + \zeta_i + \nu_{st} + \psi_{pt} + \varepsilon_{isrt} \tag{III.2}$$

where, y_{isrt} is the logarithm of value-added, capital stock, and (permanent and temporary) labor employed for firm i, belonging to sector s, located in regency r, in year t. $Flood^J_{rt}$ has the same definition as earlier. X_{isrt} includes time-varying, firm-level controls. Given the extensive impact of floods on firms, there are few suitable candidates for valid control variables. Consequently, only the logarithm of firm age and its squared term are included as firm-level controls. Firm fixed effects are controlled for as represented by parameter vector ζ_i. Similar to Equation (III.1), sector $\times$ year fixed effects are included. Since spatial margin is a key component in the flood index, with more statistical power available, province $\times$ year fixed effects denoted by ψ_{pt} are also included to control for removing differential geographic trends in flooding and outcome variables. In both estimations, control observations are defined by regency-year pairs that are not affected by flooding, meaning that the flood index is zero for control observations.

Different manufacturing sectors within the economy may be affected differently by flooding, depending on their production characteristics, such as input mix. To examine sectoral heterogeneity in the impact of flooding, I estimate an interaction version of the firm-level specification. This involves interacting the 2-digit ISIC sector dummies with flood dummies for different percentiles. The resulting specification is as follows:

$$y_{isrt} = \upsilon + \beta^J Flood^J_{rt} + \gamma^J_s \mathbb{1}\{Sector = s\} \times Flood^J_{rt} + \iota X_{isrt} + \zeta_i + \psi_{pt} + \varepsilon_{isrt} \tag{III.3}$$

Here, γ^J_s captures the estimated effect of the flood dummy J on firm-level economic variables, while all other symbols retain the same definitions as in Equation (III.2). When presenting the results, I report the combined main and interaction effects, that is, $\beta^J + \gamma^J_s$.

3.1.2 Results and Discussion

Figure 4 reports the results from estimating Equation (III.1). The contemporaneous effects of flooding on aggregate value-added (left), capital stock (centre), and labor employed (right) are negative, and the effects become stronger as the flood intensity increases. A 90th percentile flood is associated with 20%, 25%, and 15% decline in the sector-regency value-added, capital stock, and labor employment respectively.[9]

Figure 5 reports the results of estimating the firm-level specification outlined in Equation (III.2). Keeping everything else constant, a 90th percentile flood leads to a 5.7% decrease in the firm-level capital stock. While permanent labor employment does not respond to flood shocks, firms tend to increase their hiring of temporary workers during flood events. In addition, the effects on capital are driven by a few fixed capital categories, as reported in Figure 6.[10] The estimates indicate that the negative effects on the total capital stock are primarily driven by structures and land, which are more susceptible to lose their values due to the anticipatory effects of flooding, as firms preemptively reduce investments in response to the expected flooding. In contrast, machinery and other equipment, which are more vulnerable to the direct destruction effects of flooding, do not respond significantly to floods. This suggests that the perceived risk of flooding instead of actual flood shocks could be more important for capital investment decisions.[11] The temporal resolution of the analysis may also influence these results. Firms may be able to quickly *rebuild* some of their destroyed capital within a year, meaning such effects may not be captured in the annual analysis. There is evidence suggesting that firms often return to their initial production levels relatively quickly after floods (see, for example, Balboni, Boehm, and Waseem 2023, which examines firm responses to flooding in Pakistan). However, the impact on the capital stock in these cases remains unclear. Even if firms can quickly replace some of their damaged capital, the remaining capital that remains affected over a longer time horizon would influence the long-term development trajectory.[12]

Figure 7 reports the results of the heterogeneity analysis across sectors by estimat-

9. Though most of the regencies are observed for the entire 23 years of study period, to avoid compositional changes driving the estimates, I conduct a robustness check by using only those regencies for which at least 20 years of data are available. Results are robust to using this more balanced sample, and is reported in Figure B.8 in the Appendix.

10. As mentioned in the data section, the reporting on different capital categories is not consistent over time, so the number of observations used in the estimation of coefficients are different across the four categories.

11. Similar to the aggregate results, firm-level results are robust to using a more balanced sample. Results using only those firms for which data is available for at least 20 years are reported in Figure B.9 in the Appendix.

12. In 2023, 2,955 firms across the manufacturing, services, and retail sectors were surveyed in Indonesia between December 2022 and September 2023. Approximately 16% of these firms reported being impacted by a natural disaster during that period, but only about one-fourth of the affected firms experienced damage to their physical assets. *Source:* World Bank Enterprise Surveys, www.enterprisesurveys.org.

ing Equation (III.3) for the 90th percentile flood dummy.[13] The findings reveal two key patterns. First, the sector-specific results are qualitatively consistent with the overall trends shown in Figure 5, indicating that a typical firm experiences a decline in capital stock and increases temporary labor hiring after floods. Second, capital-intensive sectors—such as food processing, iron and steel, and ceramics, glass, and clay products—exhibit a more significant reduction in capital stock. This suggests that the impact of flooding on the capital margin is particularly pronounced in sectors that depend heavily on capital for their production processes.

3.2 Effect of Flooding on Firm Exit and Entry

Changes in the number of operational firms across years would contribute to the effects of flooding on aggregate economic variables. For example, reduced entries and/or increased exits could explain the negative effects obtained. Similarly, sample attrition owing to exiting firms could potentially bias the firm-level results. Firms that survive negative shocks, such as floods, are not only more adapted to deal with these shocks but are also more adept in their operations.[14] Therefore, if some small and less efficient firms are shutting down after floods, then the firm-level estimates are the lower bounds of total effects.

3.2.1 Econometric Model

To estimate the contemporaneous effect of flooding on firm exit, I employ the following econometric specification on the firm-level data:

$$y_{isrt} = \upsilon + \beta^J Flood^J_{rt} + \iota X_{isrt} + \zeta_r + \nu_{st} + \psi_{pt} + \varepsilon_{isrt} \tag{III.4}$$

where y_{isrt} is an exit dummy for firm i, belonging to 2-digit ISIC sector s, located in regency r, in year t and other terms have the same interpretation as the previous specifications.[15]

3.2.2 Results and Discussion

The left plot in Figure 8 reports the results of estimating Equation (III.4) on all the firms in the data. The results suggest that floods are not associated with firm exits. This might not be a surprising result considering that the sample comprises of medium

13. Results for flood dummies at other percentiles are provided in the Appendix.
14. Figure B.13 in the Appendix suggests that exiting firms, on average, have lower value-added, capital stock, and labor employment, compared to both new entrants and survivors.
15. The exit dummy is an "implied" variable in the sense that they are backed out from the longitudinal observation of a firm in the data. A firm's last year in the data is taken as its exit year.

and large manufacturing firms that are likely more adept at dealing with such shocks. Null results on the exit margin also indicate that the biases introduced in the previous analysis on firm-level variables are small.

The right plot in Figure 8 reports the results from estimating Equation (III.1) on the logarithm of number of new firms entering into a sector-regency in a given year. A 90th percentile flood leads to a 20% reduction in the number of new firms entering in a sector within a regency. This evidence suggests that firms avoid flood-affected regencies when setting up their operations.[16] Entry is typically a costly decision requiring investments, so firms would typically enter if they expect to make some profits net of entry costs. However, the results on entry indicate lack of good foresight on floods, as firm entry is reduced in the year of flooding. This very fact motivates the modeling assumptions in the firm entry part of the theoretical framework, in particular, an entrant's decision to enter depends on its expected profits in the current period only. Additionally, similar to the findings related to capital, these extensive margin results point towards the role of perceived flood risk at the firm level due to the evolving distribution of flood shocks.

3.3 Scope of the Reduced-form Approach and Way Forward

The reduced-form findings capture both the elements of destruction by flooding and anticipation due to the evolving flood risk across regencies. Firms located in regencies where floods are more persistent will take more anticipatory actions aimed at mitigating the effects of flooding, thereby reducing the actual destruction resulting from floods. Therefore, the reduced-form approach fails to offer a framework for understanding the mechanisms leading to the impact, particularly in linking to the literature on misallocation. The capital installation decisions are usually taken well in advance of such shocks, and firms can use some available signals on these shocks to inform their decision on the current level of capital to install. In addition, the approach becomes unreliable for investigating some sources of heterogeneity across firms that could drive the impact of flooding. Since floods are evolving shocks that are heterogeneous across space, any policy analysis would be considered incomplete without the evaluation of meaningful counterfactual scenarios beyond the existing flood experiences. A model-based approach could potentially address some of these issues.

To get to the anticipatory element of flood shocks, I exploit the uncertainty that

16. I validate this result using a more direct measure of flood risk made available for year 2013 by *Indeks Risiko Bencana Indonesia* (IRBI). I use regency-level average flood index (over the period 1990-2012) and the flood risk data available for the year 2013 to estimate their impact on average firm entry rate and firm entry rate in 2012 respectively by employing the specification: $y_r = \upsilon + \beta Flood_r + \varepsilon_r$. Entry rate is defined as the count of new firms in the current year over the total count of firms in the previous year for a given regency. Table B.3 in the Appendix suggests that regencies with high flood index and high flood risk profile witness lower firm entry rates.

firms face due to regional flooding when choosing the optimal capital to install in a given period. Capital installation decisions are taken in anticipation of floods experienced by the regencies where firms are located. This generates a time-varying flood risk variable, which vary across regencies and also across industries located within a regency. Modeling floods in this manner links them closely with the firm's production enterprise, thereby offering a microfoundation for understanding how perceived flood risk interacts with firm behavior.

The share of installed capital that can be utilized in a regency in a given year is proxied by the firm-level production capacity utilization (PCU) for that year. PCU is the percentage of available production capacity that a firm is able to utilize in a given year.[17] Figure 9 reports the results of estimating Equation (III.2) with PCU as dependent variable; a 90th percentile flood is associated with a 3.6% decline in the firm-level PCU.[18] Therefore, the empirical distribution of firm-level PCU could be used to calibrate a measure of flood risk, which captures the anticipatory effects of flooding on firms. Unlike other measures of flood risk that are generated using climate and atmospheric models and might not capture local conditions, this measure is informed from the firm-level decisions itself, and is therefore more relevant for studying the effects of flooding on firms. Given that capital installation decisions are made in anticipation of current flooding, firms account for flood risk when choosing the optimal level of capital. Firms located in flood-affected regencies are able to utilize a lower share of their installed capital, which in turn impacts the level of capital that they install to start with. The potential entrants in these markets face similar constraints due to the stochastic nature of flooding and their entry decisions are driven by expected profits. Below, I propose a model with heterogeneous firms to study the effects of flood risk and flood shocks on the production side of the economy.

4 Theoretical Framework

The multi-sector, multi-region general equilibrium model captures the interaction between regency-level flooding and firm-level economic variables. Central to the model are parameters that account for flood exposure at the regency level and firm selection based on the idiosyncratic productivity that each firm is endowed with. The investigation of the impact of flooding on firms employs a *within* production function

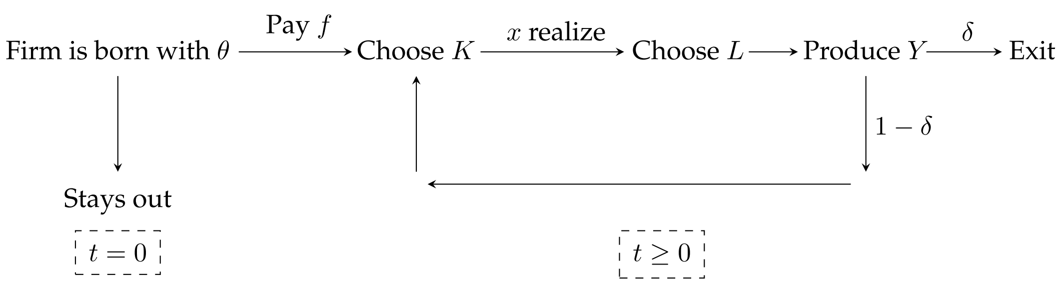
17. Figure B.14 shows that the share of firms reporting lower PCU levels tend to be higher in the flood-prone regencies, thereby providing a suggestive evidence for the impact of flooding on PCU.

18. There are some firm-year observations that report zero capacity utilization even when the output is non-zero. Since the data primarily comprises of manufacturing firms, it is less likely that they engage in other businesses to generate output, therefore, these observations could be excluded from the estimation. The reported estimates use all available observations, but the results remain qualitatively unchanged if those observations are excluded.

approach in which anticipation of floods affect firms' optimal capital installation decision. Specifically, incumbent firms internalize the constraint that flooding at their location will render only a fraction of their installed capital usable in each period. Firms are modeled as risk-neutral agents, so only the expected values of these shares matter for their decision-making. The flood risk arising from the anticipatory element of flooding also varies by industries, in particular, the capital intensity of their respective production technologies. Firms are price-takers, taking wages and rental price of capital as given. The fixed differences in productivity across firms, combined with decreasing returns to scale, generate rents for individual firms. Consequently, even though all firms within a sector are price-takers and produce a homogeneous product, the equilibrium features a distribution of firms differentiated by their idiosyncratic productivity levels.

On the extensive margin, firm entry is affected by flood risk. Each firm is born with an idiosyncratic productivity that remains constant over time. To enter a market and start production, each firm pays a one-time fixed cost. Entrants use their expected profits for the current year to decide if they should enter the market. This structure on entry decision generates firm selection, where only firms with large enough productivity decide to enter into a sector within a region.

Below is a schematic life-cycle diagram of a typical firm in the model that is born at $t = 0$ and has options to either stay out of market or pay a one-time fixed cost f to be able to choose the capital stock for period $t = 0$. The flood shock is then realized, and the firm chooses the flexible input, labor, to produce output in period $t = 0$. This three-step cycle of capital installation, flood shock realization, and flexible labor choice continues in this order with probability $(1 - \delta)$, as the firm can exit the market with probability δ in any period $t \geq 0$ for some exogenous reasons.



4.1 Technology

The production technology of a risk-neutral firm i located in regency r at time t is Cobb-Douglas in labor L and capital K, with the sector-specific output elasticity of capital and returns to scale parameters denoted by α_s and η_s respectively. The firm can only utilize a stochastic share $x_{it} \in [1, \infty)$ of the installed capital because of floods at time t, so it forms expectations on this random variable. Idiosyncratic productivity θ remains constant over time and has an ex-ante regency-specific distribution with c.d.f. (p.d.f.) denoted by H_r (h_r), which has full support in the domain $[1, \infty)$. The production function is defined as below:

$$Y_{it}(\theta, x, K, L) = \theta_i \left\{ \left(\frac{K_{it}}{x_{it}} \right)^{\alpha_s} L_{it}^{1-\alpha_s} \right\}^{\eta_s} \tag{IV.1}$$

Labor hiring decision is made after the realization of x_{it}, so the labor demand adjusts flexibly based on the capital input demand and the realized value of x_{it}. Firms take prices as given, so with wage rate w_t, the optimal choice of labor after maximizing the profit function can be written as follows:

$$L_{it}(\theta, x, K, w) = \left\{ \frac{w_t}{(1-\alpha_s)\eta_s \theta_i \left(K_{it}/x_{it}\right)^{\alpha_s \eta_s}} \right\}^{-\frac{1}{1-(1-\alpha_s)\eta_s}} \tag{IV.2}$$

Putting the optimal labor choice above back into the production function, the firm-level conditional (on capital) profit function can be written as follows:

$$\pi_{it}(\theta, x, K, z) = \Gamma_{it} \left(\frac{K_{it}}{x_{it}} \right)^{\frac{\alpha_s \eta_s}{1-(1-\alpha_s)\eta_s}} \tag{IV.3}$$

where $\Gamma_{it}(\theta, w) \equiv \left[1-(1-\alpha_s)\eta_s\right] \theta_i^{\frac{1}{1-(1-\alpha_s)\eta_s}} \left\{ \frac{w_t}{(1-\alpha_s)\eta_s} \right\}^{-\frac{(1-\alpha_s)\eta_s}{1-(1-\alpha_s)\eta_s}}$ is the product of aggregate, sector-, and firm-level parameters. z_t denotes the input price vector (w_t, ρ). Firms make their decision on the optimal capital to install before x_{it} is realized, so being risk-neutral, they maximize expected profits when making this choice under uncertainty.

4.2 Flood Shocks and Flood Risk

Firms anticipate flood shocks in a given year and choose capital accordingly. Firms internalize the fact that flooding can reduce the share of installed capital that they could utilize in a given year. Firms do not know this share precisely but have knowledge of the distribution from which the share is drawn. By maximizing the expected profits, firms choose the optimal capital to be installed in each period.

The share variables, $x \in [1, \infty)$, follow Pareto distributions with time-varying, regency-specific shape parameters ϕ_{rt}. The general form of the distribution is as follows:

$$G_{rt}(x) = \begin{cases} 1 - \left(\frac{1}{x}\right)^{\phi_{rt}} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

The above assumption on the distribution of flood impact on firms is natural and is motivated by the fact that some regencies are more flood-prone than others. Firm's expected share of capital that they can utilize in a given year would be lower for those regencies that experience more extreme flood events—regencies with heavier Pareto tails.

Capital is perfectly mobile, so the rental price of capital ρ, does not vary across regencies. This assumption is made because the central bank of Indonesia, *Bank Indonesia*, sets a national interest rate that influences lending rates across the country. Integrated financial markets within the country ensure that the no-arbitrage condition holds, leading to convergence of interest rates across regions.[19] Going back to the profit function defined in Equation (IV.3), the net expected profit function can be written as follows:

$$\Pi_{it}(\theta, x, K, z) = \Gamma_{it}\mathbb{E}\left[\left(\frac{K_{it}}{x_{it}}\right)^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\right] - \rho K_{it} \tag{IV.4}$$

Using the distribution function of share variable defined above, the objective function for the optimal capital can then be written as follows:[20]

$$K_{it} = \text{argmax}\left\{\Gamma_{it}\tau_{srt}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} - \rho K_{it}\right\} \tag{IV.5}$$

where $\tau_{srt}(\phi) \equiv \frac{\phi_{rt}}{\phi_{rt}+\alpha_s\eta_s/(1-(1-\alpha_s)\eta_s)}$ is a measure of flood risk, which captures distortions introduced in capital decisions due to flooding. This measure at the sector-regency level account for both the spatial differences in flood exposure and the differences arising due to sectoral characteristics, in particular, heterogeneity in capital intensity across industries. Note that τ_{srt} is an increasing function of ϕ_{rt}, but decreasing in η_s and also α_s in case of a decreasing return to scale technology.[21] Also, by the properties of Pareto distribution, increasing ϕ_{rt} decreases the probability of realization of larger values of x_{it}, that is, increases the probability of observing higher shares. Therefore, regencies experiencing more extreme floods on average, would have lower

19. There could still be some regional differences in the price of capital due to a multitude of factors, including regional flood risk. Ridhwan et al. (2012) show the map of these differences for rural and regional bank interests rates for the period 2000-08. First, the total assets held by these two categories of non-central banks combined is less than 10% of the total assets in the Indonesian banking industry (Financial Services Authority of Indonesia 2024). Second, the spatial pattern of deviations in interest rates reported in Figure 1 of Ridhwan et al. (2012) does not correspond to the pattern of flooding shown in Figure 1.

20. Detailed derivations are shown in the Appendix Section D

21. The setup is similar to Lucas (1978) in which the source of decreasing returns is managerial limits on the production side. Alternatively, the origin of decreasing returns could be on the demand side as shown in Hopenhayn (2014). The equivalence of results obtained from these two different approaches can be ensured by calibrating the demand elasticity, ϵ, in the latter to be equal to $(1/(1-\eta))$.

values of respective tail parameters. Finally, solving for the optimal capital choice using Equation (IV.5) delivers the equilibrium value of capital installed as outlined below:

$$K_{it}(\theta, \phi, z) = \frac{\alpha_s \eta_s}{\rho} \tau_{srt}^{\frac{1-\eta_s+\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st} \theta_i^{\frac{1}{1-\eta_s}} \tag{IV.6}$$

where $\Lambda_{st}(z) \equiv \left\{\frac{w_t}{(1-\alpha_s)\eta_s}\right\}^{-\frac{(1-\alpha_s)\eta_s}{1-\eta_s}} \left\{\frac{\rho}{\alpha_s\eta_s}\right\}^{-\frac{\alpha_s\eta_s}{1-\eta_s}}$ is the product of aggregate and sector-level parameters. To be precise, this is the stock of capital *installed* by a firm, but the amount capital that is eventually *utilized* will depend on the realization of x_{it}. More precisely, the part of capital stock that is utilized in production will be (K_{it}/x_{it}), as the remaining capital gets *destroyed* in flood. Using Equation (IV.2), the equilibrium value of labor demanded can then be written as follows:

$$L_{it}(\theta, \phi, z, x) = \frac{(1-\alpha_s)\eta_s}{w_t} \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st} \theta_i^{\frac{1}{1-\eta_s}} x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} \tag{IV.7}$$

Due to the multi-stage setup that firms use for choosing inputs within a period, the *ex-ante* (before realization of flood shocks) values of output and profit would differ from the respective *ex-post* (after realization of flood shocks) values. However, the entry decision of new firms will be based on the expected value of profit. Below are the expected values of output and profit in equilibrium:

$$Y_{it}(\theta, \phi, z) = \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st} \theta_i^{\frac{1}{1-\eta_s}} \tag{IV.8}$$

$$\Pi_{it}(\theta, \phi, z) = [1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt}]\, \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st} \theta_i^{\frac{1}{1-\eta_s}} = [1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt}]\, Y_{it} \tag{IV.9}$$

4.3 Firm Entry and Exit

Firm exit from the market is simple in that it is assumed to be exogenous and is parameterized by a constant probability of exit δ, which is independent of sector-, regency-, and firm-level variables.[22] The entry decision of firms has more subtleties involved. The pool of potential (identical) entrants is unbounded. Each entrant is born with a (constant) idiosyncratic productivity θ, which is drawn from a common regency-specific, time-invariant distribution $H_r(\theta)$. A potential entrant decides to stay out of the market in any given year if its productivity is too low. It is because each potential entrant needs to pays a one-time fixed cost, f to enter the market. Therefore, an entrant in period t would have its expected profit in period t exceed the fixed cost of entry. This generates a time-varying sector-region cutoff productivity θ^*_{srt}, below

22. Once a firm has entered, it cannot endogenously exit the market. Equation (IV.9) supports this assumption as follows. A firm can exit if its profit becomes negative in any period i.e., $1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt} < 0$, but this is not plausible in a decreasing returns to scale technology, where $\eta < 1$.

which potential entrants do not enter into sector s within regency r in year t. Combining all of the above, the expected net profit function of a potential entrant i at time t can be written as follows:

$$\sigma_{it}(\theta) = \max\{0, \Pi_{it}(\theta) - f\} \tag{IV.10}$$

where $\Pi_{it}(\theta)$ is the expected profit function defined in Equation (IV.9).

4.3.1 Aggregation

Owing to the above entry and exit dynamics, the equilibrium productivity distribution, $\mu_{srt}(\theta)$ could differ from the ex-ante distribution, $h_r(\theta)$. However, the exogeneity of exit decisions ensures that it does not affect the equilibrium distribution. In addition, all entrants with $\theta < \theta^*_{srt}$ stay out of the market, so the equilibrium distribution depends only on the productivity of entrants. That is, the equilibrium productivity distribution is a truncated version of the ex-ante distribution, as outlined below:

$$\mu_{srt}(\theta) = h_r(\theta | \theta \geq \theta^*_{srt}) = \begin{cases} \dfrac{h_r(\theta)}{1 - H_r(\theta^*_{srt})} & \theta \geq \theta^*_{srt} \\ 0 & \theta < \theta^*_{srt} \end{cases} \tag{IV.11}$$

where the probability of successful entry into sector s in regency r is $p^e_{srt} \equiv 1 - H_r(\theta^*_{srt})$.[23]

Studies have shown that the Zipf's law seems to be an empirical regularity for the firm size distribution; the results get even tighter at the upper tail, which tends to be well-approximated by a Pareto distribution (Geerolf 2017).[24] Using this information, the distribution of productivity is assumed to have the following form:

$$H_r(\theta) = \begin{cases} 1 - (\overline{\theta}_r/\theta)^{\xi} & \theta \geq \overline{\theta}_r \\ 0 & \theta < \overline{\theta}_r \end{cases}$$

where $\overline{\theta}_r$ is a regency-specific scale parameter that reflects the spatial differences in firm productivity. For example, a firm born in Jakarta might be more or less productive due to competing agglomeration and congestion externalities existing there. Without firm entry and exit, the equilibrium productivity distribution is same as the initial distribution, $h_r(\theta)$. But when firms are allowed to enter and exit the markets, the equilibrium distribution takes the form defined in Equation (IV.11). I now define the

23. Hopenhayn (1992) discusses the assumptions under which law of large numbers could be applied to determine the equilibrium distribution, $\mu_{srt}(\theta)$ from the initial distribution, $h_r(\theta)$.

24. Since, the data comprises only of medium and large manufacturing firms, the Pareto assumption on firm productivity distribution is even more innocuous.

expected aggregate output for sector s in regency r at time t, $\overline{Y}_{srt} \equiv \int_{\theta^*_{srt}}^{\infty} Y_{it}(\theta)\mu_{srt}(\theta)d\theta$, which is a total weighted output of all surviving firms, where the weights are mass of firms at each productivity level in the equilibrium distribution. Combining the definition of equilibrium distribution from Equation (IV.11) with the firm-level expected output defined in Equation (IV.8), the expected aggregate output can be written as follows:[25]

$$\overline{Y}_{srt} = \tau_{srt}^{\frac{\alpha_s \eta_s}{1-\eta_s}} \frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1} \Lambda_{st} \left(\theta^*_{srt}\right)^{\frac{1}{1-\eta_s}} \tag{IV.12}$$

The only parameter restriction that needs to be imposed here is that $(\xi(1-\eta_s) > 1)$, which, as would be seen later in the estimation part, is true for all the 3-digit ISIC manufacturing sectors in the economy. Also, since $(\theta^*_{srt} > 1)$ and $(\xi > 0)$, given constant wages, the expected aggregate output and thereby expected aggregate profit increase after considering the firm selection above. Similar to the above, expressions for all other expected aggregate variables could be derived using Equations (IV.8) and (IV.9)

One key object is the cutoff productivity θ^*_{srt}, which is the productivity of least productive firm deciding to enter into sector s in regency r at time t. This can be pinned down using the firm-level expected profit function as described next.

4.3.2 Zero Cutoff Net Profit Condition

Each entrant needs to pay a one-time fixed cost, f to enter into any market. This fixed cost can be thought of as a permit or license fee that each new firm needs to pay to a centralized authority. An entrant would be willing to pay this fee if it can make a non-negative expected net profit, that is, the cutoff productivity, $\theta^*_{srt} \equiv \inf\{\theta : \sigma_{it}(\theta) > 0\}$, where the net expected profit of entrant is defined in Equation (IV.10). Using the firm profit function defined in Equation (IV.9) delivers the expression for cutoff productivity below:

$$\theta^*_{srt} = \left\{ \frac{f}{\left[1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt}\right] \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st}} \right\}^{1-\eta_s} \tag{IV.13}$$

Keeping wages fixed, θ^*_{srt} is a decreasing function of τ_{srt},[26] so, the cutoff productivity level is larger for more vulnerable sectors and flood-prone regencies, that is, the idiosyncratic productivity needs to be high enough for firms belonging to these sectors if they intend to establish operations in the flood-prone regencies, thereby suggesting that the firm selection effects would be stronger in these cases.

25. The intermediate steps for getting to this final expression are outlined in the Appendix Section D.
26. For this to hold with certainty, η_s needs to be smaller than unity, that is, the production technology should have decreasing returns to scale.

4.4 Equilibrium

The definition of equilibrium is standard, with the labor market clearing at the aggregate level each year. Like capital, labor is mobile across regencies, so wages are constant in space, but they do adjust over time in response to flooding. Therefore, the equilibrium is pinned down by the wages w_t and the cutoff productivity levels θ^*_{srt}. The aggregate labor endowment, $\overline{L}$ is assumed to be exogenous and constant over years. The labor market clearing condition delivers the equilibrium wage equation as follows:[27]

$$w_t = \frac{f}{\overline{L}} \sum_{r=1}^{R} \sum_{s=1}^{S} \frac{(1-\alpha_s)\eta_s \tau_{srt}}{1-(1-\alpha_s)\eta_s - \alpha_s \eta_s \tau_{srt}} \frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1} \tag{IV.14}$$

It is easy to see that wages increase when the fixed cost of entry rises. It is because with increase in the fixed cost, the firm selection also gets stronger, that is, fewer and more productive firms are able to enter the markets. Since, the marginal product of labor depends on firm's productivity, when labor is reallocated to higher productivity firms, the marginal product of labor increases. Thus, wages also rise to match this increased marginal product in equilibrium.

Also, the sign of first derivative of wage equation w.r.t τ_{srt}, which quantifies the overall effect of flooding for sector s located in regency r at time t, is positive. This means that equilibrium wages go down as the impact of flooding increases. To get to the intuition behind this result, first, remember that flood shocks reduce the utilization of capital and that reduces the returns on capital in areas affected by flooding. In the model, firms rely on both capital and labor to maximize output net of input costs. When capital is less productive, the marginal productivity of labor also declines, since firms cannot use labor as effectively without undistorted capital. This lower productivity reduces firms' demand for labor at any given wage, as they need to scale down operations in response to flooding. With a reduced demand for labor, the equilibrium wages undergo downward adjustment to clear the labor markets. Wages also decrease due to an increase in flood risk that is captured by a decline in τ_{srt}. It is because with increased risk, less firms will enter into these markets, decreasing the competition for labor, and eventually driving the wages downwards to clear the labor market.

With the equilibrium wages in hand, using Equation (IV.13), the equilibrium values of cutoff productivity levels can be derived to follow the general analytical expression

27. Detailed derivation in the Appendix Section D

below:

$$\begin{aligned}\theta^*_{srt} = & \frac{f^{1-\eta_s}}{((1-\alpha_s)\eta_s)^{(1-\alpha_s)\eta_s}} \left(\frac{\rho}{\alpha_s\eta_s\tau_{srt}}\right)^{\alpha_s\eta_s} \\ & \times \left\{\frac{1}{1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt}}\right\}^{1-\eta_s} \\ & \times \frac{f}{\overline{L}} \left\{\sum_{r=1}^{R}\sum_{s=1}^{S} \frac{(1-\alpha_s)\eta_s\tau_{srt}}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}} \frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}\right\}^{(1-\alpha_s)\eta_s}\end{aligned} \tag{IV.15}$$

The derivative of the above expression w.r.t. τ_{srt} has a negative sign, thereby suggesting that even after accounting for equilibrium wage adjustments, the cutoff productivity level for entering into more vulnerable sectors located in flood-prone regencies needs to be higher.

The set of parameters: $\left(\{\{\phi_{rt}\}_{r=1}^{R}\}_{t=1}^{T}, \{\alpha_s\}_{s=1}^{S}, \{\eta_s\}_{s=1}^{S}, \xi, \{\overline{\theta}_r\}_{r=1}^{R}, \rho, f, \delta, \overline{L}\right)$ would need to be estimated or calibrated to compute each equilibrium object in levels. However, the subsequent analyses will require estimates for most—but not all—of these parameters. Specifically, estimates for the last four parameters, which represent the rental price of capital, fixed cost of entry, exit probability, and aggregate labor supply, will not be needed.

5 Estimation

This section describes the estimation of some of the parameters in the model. Table 1 reports the summary on these parameters along with the estimation and calibration techniques employed.

5.1 Flooding Shape Parameters (ϕ_{rt})

To estimate the shape parameters of the Pareto distributions of the share of installed capital that can be utilized in a regency in a given year, I use the empirical distribution of PCU for firms located in the regency in that year. One key advantage of using an objective measure, such as PCU, is that it is immune to price changes and market adjustments due to new firm entries. Other similar measures, such as output and value of capital stock, would potentially be impacted by these equilibrium adjustments. Additionally, PCU is a relative measure based on what firms were able to produce relative to what they had *planned* in a given year, so it can be used as a proxy for flood shocks at the firm level. With this, the maximum likelihood estimator of the shape parameter can be derived and that has the following form:[28]

$$\widehat{\phi}_{rt} = \frac{N_{rt}}{\sum_{i=1}^{N_{rt}} \ln(x_{it})}$$

where x_{it} is the reciprocal of PCU for firm i in year t and N_{rt} is the number of firms located in regency r in year t. Under the assumption that firms' expectations are informed *only* by past events, one could alternatively use the past realizations of PCU to estimate this parameter. However, the current realizations would have elements from the past floods due to the auto-correlated nature of flood shocks. In the model, due to the multi-stage decision process within a period, entrants would not know the realized flooding before they enter nor the incumbents can choose capital under perfect information about flooding.

Figure 10 shows the distribution of average (over years) estimated flood exposure, ϕ_{rt} for each regency in the data. Due to the property of Pareto distribution, lower the value of shape parameter, higher is the exposure of regency to extreme flooding. Clearly, regencies in the West Java province, including Jakarta, are some of the most exposed regencies in Indonesia.

Figure 11 shows the correlation between the average (over years) shape parameters with the empirical flood index across regencies. The relationship is statistically significant and suggests that flood-prone regencies experience more severe flooding (thicker Pareto tails) on average. Therefore, regency-level flood index and model-based flood exposure are related, and the relationship is in the expected direction.

5.2 Firm Productivity Parameters ($\overline{\theta}_r, \xi$)

I calibrate the scale parameter, $\overline{\theta}_r$ to match the logarithm of aggregate regency value-added (akin to regional GDP) that is computed using the data I have on medium and large manufacturing firms in Indonesia. The shape parameter ξ, is then estimated using the following maximum likelihood estimator:[29]

$$\widehat{\xi} = \frac{N_r}{\sum_{i=1}^{N_r} \ln(\theta_i/\overline{\theta}_r)}$$

where N_r is the number of firms in regency r and θ_i is the logarithm of average value-added for firm i located in regency r. The point estimate is 4.514 with a robust standard

28. Detailed derivation of the estimator is in the Appendix Section D.
29. I use the logarithm of average value-added over the entire period for which a firm has been operational in the data. This averaging exercise fixes each firm's productivity to a constant, but alleviates some concerns around measurement error in the reported figures.

error of 0.0237.[30]

Given that the production function parameters, α_s and η_s are already estimated, one could compute the Solow residuals as a measure of productivity. However, this approach could be problematic for two reasons. First, flood risk can affect productivity. I follow an approach that closely mirrors Besley, Roland, and Reenen (2020) to show that this case does not arise here, owing to the aggregate nature of shocks. Second, measurement error in firm-level data, especially with respect to capital, could severely bias the productivity estimates (Collard-Wexler and De Loecker 2016). However, the impact of measurement error on the estimates of production function parameters should be minimal due to their aggregate nature.

The output share of firm i, belonging to sector s and located in regency r, $\kappa_{it} \equiv Y_{it}/\overline{Y}_{srt}$, will be equal to the *relative* productivity of firm, ω_{it} in a world without friction. However, in the presence of distortions due to flooding, the relative productivity could also change. Using Equation (IV.8), the productivity terms can be written as follows:

$$\theta_i^{\frac{1}{1-\eta_s}} = \frac{Y_{it}}{\overline{Y}_{srt}} \frac{\overline{Y}_{srt}}{\tau_{srt}^{\frac{\alpha_s \eta_s}{1-\eta_s}} \Lambda_{st}}$$

Further, including the general expression for aggregate output in the above, one gets ω_{it} to be equal to κ_{it}. Therefore, productivity is not affected by flood risk.

5.3 Production Function Parameters (α_s, η_s)

I employ the production function estimation approach using the technique proposed in Levinsohn and Petrin (2003) for each 3-digit ISIC sector. The method uses similar identification ideas as in a control-function setup to address the issue of endogeneity of input choices. Unlike the seminal work Olley and Pakes (1996), which uses investment as a proxy to control for unobserved productivity shocks, Levinsohn and Petrin (2003) uses intermediate inputs, such as materials or energy, as proxies for unobserved productivity shocks. Intermediate inputs are often more flexible and can adjust more quickly to productivity changes than investment, which also suffers from issues such as lumpiness in choice and a lot of zero entries in the data. In the estimation, logarithms of real value-added, real capital stock, quantity of labor employed, and real material costs are used as left-hand side, state, static input, and productivity proxy variables respectively.

Table 2 reports the computed production function parameters for each 3-digit ISIC

30. One potential concern here could be that the equilibrium distribution, $\mu_{srt}(\theta)$, is *related* to the ex-ante distribution, $h_r(\theta)$, but it is not the same distribution. This issue becomes important when we are dealing with small samples within sector-region pairs because the non-applicability of the law of large numbers will not let us approximate the equilibrium distribution from the ex-ante distribution.

sector; some of the capital-intensive sectors are Industrial Chemical Products, Basic Iron and Steel, and Machines and Repairs.[31] With the production function parameters in hand, I can compute the flood risk, represented by τ_{srt} in Equation (IV.5), for each sector within a given regency. Figure 12 shows the distribution of average (over years and sectors) flood risk across regencies. Lower value of flood risk point to larger capital distortions caused by flooding. Additionally, the spatial distribution of flood risk has noticeable differences from the flood exposure map shown in Figure 10, thereby highlighting significant sectoral variations in industrial settlement patterns across regencies.

With the parameter estimates in hand, I show some analysis on flood risk and flood shocks using the equilibrium conditions derived from the model. Due to the exogenous nature of flood shocks, one could look at the effects of flooding on output both with and without equilibrium adjustments. The next section delves into the details.

6 Analysis

This section delineates the results of the analysis conducted using the structure of model and parameter estimates from the previous two sections. In the first part, I use the equilibrium conditions from the model to disentangle the effects of flood and flood risk. Next, I conduct a counterfactual analysis on flood defenses and disentangle effects due to different margins.

6.1 Flood Shocks and Flood Risk

In the model, flood risk directly affect the optimal capital installation decision, while labor hiring decision is affected by flood shocks. However, the installed capital could suffer destruction ex-post due to realized flood shocks. So, I can use the expressions for firm-level equilibrium capital stock (accounting for destruction) and labor to understand how each gets impacted by different elements of flooding. Taking logs of Equations (IV.6) and (IV.7) delivers the following two equations:

$$\ln K_{it} = \underbrace{\ln\left(\frac{\alpha_s\eta_s}{\rho}\right) + \ln\Lambda_{st}}_{\text{sector} \times \text{year fixed effect}} + \frac{1-\eta_s+\alpha_s\eta_s}{1-\eta_s}\ln\tau_{srt} + \underbrace{\frac{1}{1-\eta_s}\ln\theta_i}_{\text{firm fixed effect}}$$

$$\ln L_{it} = \underbrace{\ln\left(\frac{(1-\alpha_s)\eta_s}{w_t}\right) + \ln\Lambda_{st}}_{\text{sector} \times \text{year fixed effect}} + \frac{\alpha_s\eta_s}{1-\eta_s}\ln\tau_{srt} + \underbrace{\frac{1}{1-\eta_s}\ln\theta_i}_{\text{firm fixed effect}} - \frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}\ln x_{it}$$

31. Table B.4 reports the estimated output elasticities of capital and labor for each 3-digit ISIC sector.

One could compute the coefficients on τ_{srt} and x_{it} directly using the estimated parameters but it would be subject of endogeneity concerns. Therefore, to estimate the elasticities of capital and labor with respect to flood risk and actual flood shock, I employ the following econometric specifications based on the above two equations, accounting for the destruction effect of flood shocks on capital:

$$\ln K_{it} = \nu_{st} + \beta^K \ln \tau_{srt} + \zeta_i + \gamma^K \ln x_{it} + \varepsilon_{it}^K \quad \text{(VI.1)}$$

$$\ln L_{it} = \nu_{st} + \beta^L \ln \tau_{srt} + \zeta_i + \gamma^L \ln x_{it} + \varepsilon_{it}^L \quad \text{(VI.2)}$$

In the above specifications, τ_{srt} is a measure of flood risk. The reason why this variable captures only flood risk and not the combined effects of flood risk and actual flooding is that it reflects average disruptions in production capacity under expected flooding. Therefore, the effect of flood risk is identified by the cross-regency differences in the mean probability of flooding, while the effects of actual flood shocks uses the variation induced by annual flood events. There is a plausible endogeneity concern with this setup if PCU is itself influenced by labor and capital levels. For example, firms with less labor or capital might experience less or more production disruptions during floods, which would impact capacity utilization. One way to address this concern is to instrument the flood risk and actual flooding with some relevant objective measures of flooding. For example, flood index could be a potential IV in this case. However, as discussed in the reduced-form section, flood index potentially captures both the effects, so it will not be a valid IV for either of the two variables. It is also a coarser variable as it only varies across regencies over time, so the analysis could potentially suffer from weak instrument problem (Angrist and Pischke (2009)). To partially alleviate the concerns around endogeneity, I look at whether there is heterogeneous effect of flooding on PCU due to the firm size, where the firm size is calculated based on both the labor employment and capital stock measure. The results reported in Table B.5 in the Appendix suggests that such a heterogeneity in the impact of flooding on PCU is non-existent.

The flood risk variable above uses the estimated parameter values of regency-level flood exposure based on the distribution of firm-level PCU from past years only. This is essential for the analysis to have enough statistical power to identify both β and γ parameters capturing effects of flood risk and flood shocks respectively. Table 3 report the results of estimating Equations (VI.1) and (VI.2) in Columns 5 and 6 respectively. Result in Column 5 suggests that an increase in τ_{srt}, that is, a decrease in flood risk leads to an increase in the capital stock at firm level. In terms of magnitude, a 1% decrease in τ_{srt} leads to a 0.25% decrease in the value of capital stock. Coefficient estimates on flood risk variable in Column 6 suggests that the labor demand increase in response to an increase in flood risk. More precisely, a 1% decrease in τ_{srt} leads to a 0.13% increase in labor employment. Firms tend to reduce investment in capital, which is less flexible and more vulnerable to the effects of flooding, and increase hiring of labor, which is a relatively more flexible input in production. The coefficients on flood shock variable suggests that its impact on both capital and labor inputs is small relative to flood risk. Flood shock affect labor and capital directly, while flood risk enters in the labor decision only through the capital margin. The results highlight that controlling for the indirect effects of flooding operating through flood risk, the direct impact of flooding on firm-level capital stock and labor employment is limited.

6.2 Flood Defenses

One counterfactual exercise that is natural in this setting is the installation of flood defense systems, such as flood barriers and fences in flood-prone areas. This exercise is in the spirit of various mitigation efforts undertaken by both local and central governments in Indonesia through both in-house and international support (Islam et al. 2019). For this *experimental* exercise, I assume installation of flood defenses in the most-affected regencies of Indonesia in terms of flooding. The metric used for classifying these regencies is the average flood exposure, which is estimated using the empirical distribution of firm-level PCU, across years. Flood defense systems reduce the flood exposure of regencies where they are installed by bringing the exposure level down to a lower level.

Flooding is an exogenous event, and the parameter governing it is also independent of equilibrium effects in the model viz. wage adjustments and endogenous entry decisions. Thus, one can quantify the *direct* effects of flood risk and then examine how these effects change once equilibrium forces viz. wage adjustments and endogenous firm entry, are taken into account. For this reason, the analyses are reported in two distinct scenarios outlined below:

1. *Flood Risk Only*: This scenario quantifies the change in direct effects of flood risk before and after the installation of flood defenses. Due to the exogeneity of parameter governing the direct effects of flood risk, it can be inspected separately from the equilibrium adjustments that occur as a result of it.

2. *Equilibrium With Entry*: This scenario quantifies the total change, including both direct and indirect effects of flood risk, before and after the installation of flood defenses. Flood risk affect firm entry decision and equilibrium wage, so accounting for these adjustments is essential for capturing the overall benefits of flood defenses.

To quantify the effects of flood defenses on the aggregate output, I compare the aggregate output after the installation of flood defenses to the observed aggregate output. In this thought experiment, flood defenses reduce the flood exposure of the top 20th percentile of regencies by bringing them down to the median level of flood exposure distribution. The flood exposure distribution is generated using average (over years) flood exposure of all the regencies in Indonesia. The counterfactual assigns same flood exposure level to the most-affected regencies in all the years, but different sectors within a regency would still be impacted differently due to their sectoral characteristics. The motivation for using the median benchmark is firstly to be realistic that flooding cannot be completely eliminated, so all regencies should experience some level of flooding in the *constrained* best-case scenario. Secondly, flooding, being a spatial shock by design, primarily creates differences across regencies, and switching off this channel by bringing all treated regencies to the same level could provide insights on the spatial misallocation effects of flooding.

To operationalize this experimental exercise, consider an exogenous change in the flood exposure of regency r from ϕ_{rt} to $\widetilde{\phi}_r$, where $\widetilde{\phi}_r$ is the median value of regency-level flood exposure. Only the top 20th percentile regencies undergo this change in flood exposure from the start of the period, while the remaining regencies remain unaffected in all the years. Figure 13 shows the spatial distribution of flood exposure following the installation of flood defenses. As the most-exposed regencies are now more secure from flooding, the minimum value of average ϕ increases for 53 treated regencies out of a total of 266 regencies in the sample. Similarly, Figure 14 depicts the spatial distribution of flood risk, which decreases for the top 20th percentile of the most-exposed regencies. For the observed outcomes, Equation (IV.12) delivers the expected equilibrium value of aggregate output. The same equation can then be used to write the counterfactual output after the installation of flood defenses as below:

$$\widetilde{\overline{Y}}_{srt} = \widetilde{\tau}_{srt}^{\frac{\alpha_s \eta_s}{1-\eta_s}} \frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1} \widetilde{\Lambda}_{st} \left(\widetilde{\theta}^*_{srt}\right)^{\frac{1}{1-\eta_s}}$$

There are various ways in which the above output could be compared to the observed output. In the most comprehensive analysis, one can ideally calibrate or estimate all the parameters involved in both the objects and compute the objects in *levels*. However, this exercise would require imposing additional assumptions on the model structure and would also be prone to measurement errors. Therefore, I take the ratio of the two objects, which cancels all the *fixed* terms that are assumed to not change in the counterfactual world.[32] The ratio can be written as follows:

$$\frac{\widetilde{\overline{Y}}_{srt}}{\overline{Y}_{srt}} = \left(\frac{\widetilde{\tau}_{srt}}{\tau_{srt}}\right)^{\frac{\alpha_s \eta_s}{1-\eta_s}} \frac{\widetilde{\Lambda}_{st}}{\Lambda_{st}} \left(\frac{\widetilde{\theta}^*_{srt}}{\theta^*_{srt}}\right)^{\frac{1}{1-\eta_s}}$$

32. The assumption imposed on the equilibrium is that the exogenous change in flood exposure does not affect the production function, entry fixed cost, and productivity distribution parameters.

Taking the log of the above ratio, I derive the (log) change in aggregate output as follows:

$$\widetilde{\Omega}_{srt} = \underbrace{\frac{\alpha_s \eta_s}{1-\eta_s} \ln\left(\frac{\widetilde{\tau}_{srt}}{\tau_{srt}}\right)}_{\text{flood risk}} + \underbrace{\ln\left(\frac{\widetilde{\Lambda}_{st}}{\Lambda_{st}}\right) + \frac{1}{1-\eta_s} \ln\left(\frac{\widetilde{\theta}^*_{srt}}{\theta^*_{srt}}\right)}_{\text{wage and entry adjustments}} \tag{VI.3}$$

$\widetilde{\Omega}$ captures the change in (log) aggregate output in the counterfactual with respect to the real world. So, a positive $\widetilde{\Omega}$ would mean that the aggregate counterfactual output is higher, and for small changes, the magnitude would represent the percentage increase in output relative to the observed output. The total change in aggregate output can be decomposed into two parts: (A) flood risk and (B) wage and entry adjustments. The first part captures the direct impact of flood risk, while the second part provides estimates of indirect effects due to the equilibrium forces in place. In the results that follow, *Flood Risk Only* scenario reports estimates of (A), while *Equilibrium With Entry* scenario reports the sum of (A) and (B).

Flood Risk Only. In this scenario, the estimates of the first term in Equation (VI.3) is reported. Given that there are two key margins of variation viz. sector and regency, I report results on both the margins.[33] Although flooding is inherently a regional shock, its impact can vary significantly across economic sectors, depending on each sector's specific vulnerabilities to flood risk and flood events. For example, an iron and steel firm and a furniture producer might both be located on the same floodplain in Jakarta, yet the impact of flood events—and how each firm perceives these events to inform their flood risk—could differ greatly owing to their sectoral characteristics. For the aggregation step, I take simple average across all qualifying observations. Figure 15 shows the distribution of change in aggregate output due to flood risk across sectors, where the sectors are ordered in increasing order of their capital intensities from left to right. First, on average, all sectors derive direct benefits in terms of aggregate output from the installation of flood defenses in the top 20th percentile of most flood-affected regencies in Indonesia. This confirms a well-known empirical fact that industries in Indonesia are primarily clustered in flood-prone areas. Such clustering in high-risk zones is due to various factors, including historical path dependence, agglomeration externalities, and higher demand due to richer population. Second, the benefits increase moving from left to right on the graph, thereby suggesting that the sectors using capital-intensive technology for production reap more direct rewards from such protective investments. This is because the sectors that rely on capital heavily face more distortions in their production decisions because of flood risk, and flood defenses help

33. There is also time variation but it is not key part of the analysis as all flood defense systems everywhere are installed at the start of the period.

alleviate those distortions. On average, the aggregate annual sector-level output increases by 7%, but with significant heterogeneity across sectors. Figure 16 illustrates the spatial distribution of changes in aggregate output resulting from reduced flood risk across regencies. On average, the aggregate annual output increases by 16% in the treated regencies, though the range is considerable, varying from 9% to 30%. The sectoral composition of each regency plays a significant role in shaping these outcomes; for instance, regencies with a higher concentration of capital-intensive sectors tend to experience greater gains in aggregate output compared to those with a higher proportion of less capital-intensive sectors. For example, regencies in the West Java province, including Jakarta, see disproportionately higher gains in their expected aggregate output.

Equilibrium With Entry. This scenario captures the total change in aggregate output—the sum of the two parts outlined in Equation (VI.3) after the flood defenses intervention. Understanding the direct effects of protective investments, such as flood defenses, is important to justify the monetary costs involved in their construction and maintenance. However, the indirect effects could potentially increase these benefits further or decrease them depending on the margin looked at in calculating these effects. One such margin that is important to consider before such interventions are commissioned is how potential firms, which are still out of the market, would respond to the changes resulting from interventions. Installation of flood defenses potentially increase the pool of new firms entering into these safer areas, but that would also increase competition among incumbent firms for the scarce resources employed in the production of final goods.

Figure 17 shows the distribution of total change in aggregate output across manufacturing sectors; results of the previous scenario are also included side-by-side for comparison. The sum of direct and indirect effects is positive for all the sectors. On average, aggregate output increases by 4% from the observed outcome after the installation of flood defenses. However, accounting for the equilibrium forces of wage adjustment and firm selection on entry decreases the aggregate gains for all the manufacturing sectors relative to the previous scenario, which captures only the direct effects of flood defenses. Flood defenses, by design, make risky regencies safer for economic operations. Since the potential market entrants make their entry decision on the expected profits, which depend on the anticipation of flooding, installation of flood defenses increase these expected profits. This means that less productive firms, which were unable to enter earlier due to high flood risk, would be able to enter into these markets now. Due to the larger mass of incumbent firms in equilibrium, the competition for the scarce labor inputs also increases, thereby exerting an upward pressure on the equilibrium wages. Therefore, the increased competition driving wages upwards combined with the reduction in firm selection on entry decreases the aggregate output in equilibrium relative to the direct impact on aggregate output due to flood risk after the installation of flood defenses.[34] Figure 18 shows the distribution of total change in aggregate output across treated regencies. Similar to the sectoral distribution above, total benefits of flood defenses decrease for all regencies when indirect equilibrium effects are accounted in the change calculation. Overall, the yearly aggregate output increases by 9%, which is about half of the gains from considering direct effects only.

7 Conclusion

This paper investigates the impact of flooding on the manufacturing sector of a low- and middle-income country. Using historical data on floods, I show that severe floods are associated with significant reductions in aggregate measures of production inputs and economic output. Though at the firm level, the value of capital stock declines and hiring of temporary labor increases, with the risk of floods also acting as a deterrent to firm entry. However, in regions with persistent flood shocks, both the actual damages from floods and the anticipatory adjustments in response to evolving perception of flood risk, could play a significant role in generating these results. To address this, I develop a model of firms with endogenous entry and flood risk affecting capital installation decisions, to assess the effects of different elements of flooding on firm behavior. I provide a microfoundation for understanding how flood risk and flood shocks interact with firm behavior by linking them to entry decision and input choice. The equilibrium analysis reveals that perceived flood risk, rather than actual flood shocks, has more significant impact on firm behavior. I conduct a counterfactual analysis in the spirit of building flood defenses to secure flood-prone regions and find that there are large gains in aggregate output from such an intervention, but equilibrium adjustments, in particular, upward pressure on wages and the entry of less productive firms, reduce these gains by half.

The theoretical framework developed in this paper could be adapted to examine other aggregate and firm-level distortions generated by anticipation, such as AI adoption, technological disruptions, and policy-induced market frictions. Incorporating firm entry and exit dynamics into the model, while maintaining its tractability for policy analysis, could be useful in various contexts, such as assessing the impact of global trade disruptions, regional economic integration, and industry regulations.

34. One concern might be that the results are driven by regency characteristics where the sectors are located. To address this issue, I plot the same figures keeping regencies constant across sectors. There are only six regencies where all 25 3-digit ISIC sectors are located, so the average is taken over these six regencies only. Figures B.15 and B.16 show the new graphs, which point to the same qualitative findings as reported in the main figures.

References

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Asian Development Bank. 2019. *Policies to Support the Development of Indonesia's Manufacturing Sector During 2020–2024.* Technical report. A Joint ADB–BAPPENAS Report. https://dx.doi.org/10.22617/TCS199910-2.

Balboni, Clare. 2024. "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities." *Working Paper.*

Balboni, Clare, Johannes Boehm, and Mazhar Waseem. 2023. "Firm adaptation and Production Networks: Structural Evidence from Extreme Weather Events in Pakistan." *Working Paper.*

Banerjee, Abhijit, and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." Chap. 07 in *Handbook of Economic Growth,* 1st ed., edited by Philippe Aghion and Steven Durlauf, vol. 1, Part A, 473–552.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103 (1): 305–34.

Bento, Pedro, and Diego Restuccia. 2017. "Misallocation, Establishment Size, and Productivity." *American Economic Journal: Macroeconomics* 9 (3): 267–303.

Besley, Timothy J., Isabelle A. Roland, and John Van Reenen. 2020. "The Aggregate Consequences of Default Risk: Evidence from Firm-level Data." *NBER Working Paper* 26686.

Bilal, Adrien, and Esteban Rossi-Hansberg. 2023. "Anticipating Climate Change Across the United States." *NBER Working Paper* 31323.

Blalock, Garrick, and Paul J. Gertler. 2008. "Welfare Gains from Foreign Direct Investment through Technology Transfer to Local Suppliers." *Journal of International Economics* 78 (2): 402–421.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *The Review of Economic Studies* rdae007.

Castro-Vincenzi, Juanma. 2024. "Climate Hazards and Resilience in the Global Car Industry." *Working Paper.*

Collard-Wexler, Allan, and Jan De Loecker. 2016. "Production Function Estimation and Capital Measurement Error." *NBER Working Paper* 22437.

Cruz, José-Luis, and Esteban Rossi-Hansberg. 2021. "The Economic Geography of Global Warming." *University of Chicago, Becker Friedman Institute for Economics Working Paper,* nos. 2021-130.

Desmet, Klaus, Robert E. Kopp, Scott A. Kulp, Dávid Krisztián Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg, and Benjamin H. Strauss. 2021. "Evaluating the Economic Cost of Coastal Flooding." *American Economic Journal: Macroeconomics* 13 (2): 444–86.

Financial Services Authority of Indonesia. 2024. *Indonesia Banking Statistic - January 2024.* Technical report. https://www.ojk.go.id/en/kanal/perbankan/data-dan-statistik/statistik-perbankan-indonesia/Pages/Indonesia-Banking-Statistic---January-2024.aspx..

Geerolf, François. 2017. "A Theory of Pareto Distributions." *Working Paper.*

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital Allocation and Productivity in South Europe*." *The Quarterly Journal of Economics* 132 (4): 1915–1967.

Government of Republic of Indonesia. 2007. *Indonesia Country Report: Climate Variability and Climate Changes, and Their Implication.* Technical report. Ministry of Environment, Jakarta.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–1150.

———. 2014. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics* 6:735–770.

Hsiao, Allan. 2024. "Sea Level Rise and Urban Adaptation in Jakarta." *Working Paper.*

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403–1448.

IPCC. 2023. *Climate Change 2023: Synthesis Report Summary for Policymakers.* Technical report. Intergovernmental Panel on Climate Change.

Islam, S., C. Chu, J. C. R. Smart, and L. Liew. 2019. "Integrating Disaster Risk Reduction and Climate Change Adaptation: A Systematic Literature Review." *Climate and Development* 12 (3): 255–267.

Jia, Ruixue, Xiao Ma, and Victoria Wenxin Xie. 2022. "Expecting Floods: Firm Entry, Employment, and Aggregate Implications." *NBER Working Paper* 30250.

Kocornik-Mina, Adriana, Thomas K. J. McDermott, Guy Michaels, and Ferdinand Rauch. 2020. "Flooded Cities." *American Economic Journal: Applied Economics* 12 (2): 35–66.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies* 70 (2): 317–341.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9 (2): 508–523.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2).

Najibi, Nasser, and Naresh Devineni. 2018. "Recent Trends in the Frequency and Duration of Global Floods." *Earth System Dynamics* 9 (2): 757–783.

Nath, Ishan. 2024. "Climate Change, the Food Problem, and the Challenge of Adaptation through Sectoral Reallocation." *Journal of Political Economy* forthcoming.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263–1297.

Patel, Dev. 2024. "Floods." *Working Paper.*

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* 11 (4): 707–720.

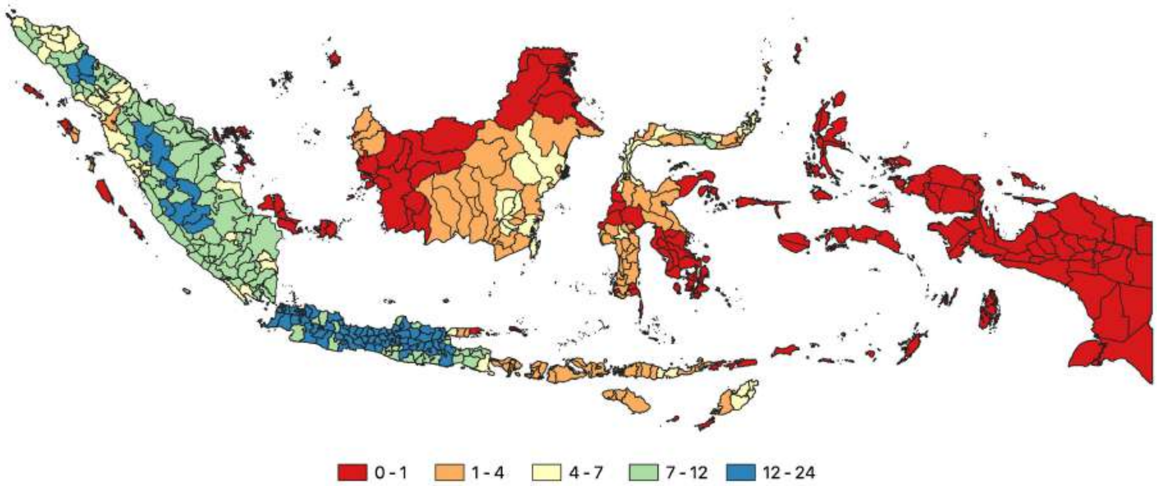
Ridhwan, Masagus M., Henri L.F. de Groot, Piet Rietveld, and Peter Nijkamp. 2012. "Regional Interest Rate Variations: Evidence from the Indonesian Credit Markets." *Tinbergen Institute Discussion Papers* 12-073/3.

Tellman, B., J. A. Sullivan, C. Kuhn, A. J. Kettner, C. S. Doyle, G. R. Brakenridge, T. A. Erickson, and D. A. Slayback. 2021. "Satellite Imaging Reveals Increased Proportion of Population Exposed to Floods." *Nature* 596:80–86.

World Risk Report. 2023. *World Risk Report 2023.* Technical report. Bündnis Entwicklung Hilft, Ruhr University Bochum – Institute for International Law of Peace and Conflict.

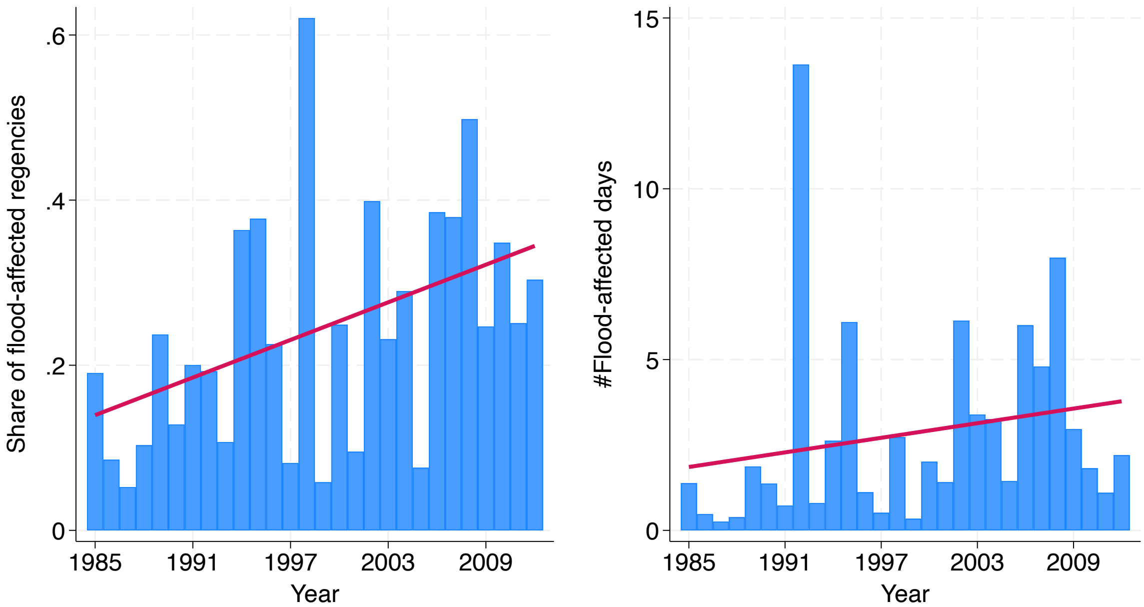
Main Tables and Figures

Figure 1: **Total count of large floods affecting regencies in 1985-2012 period**



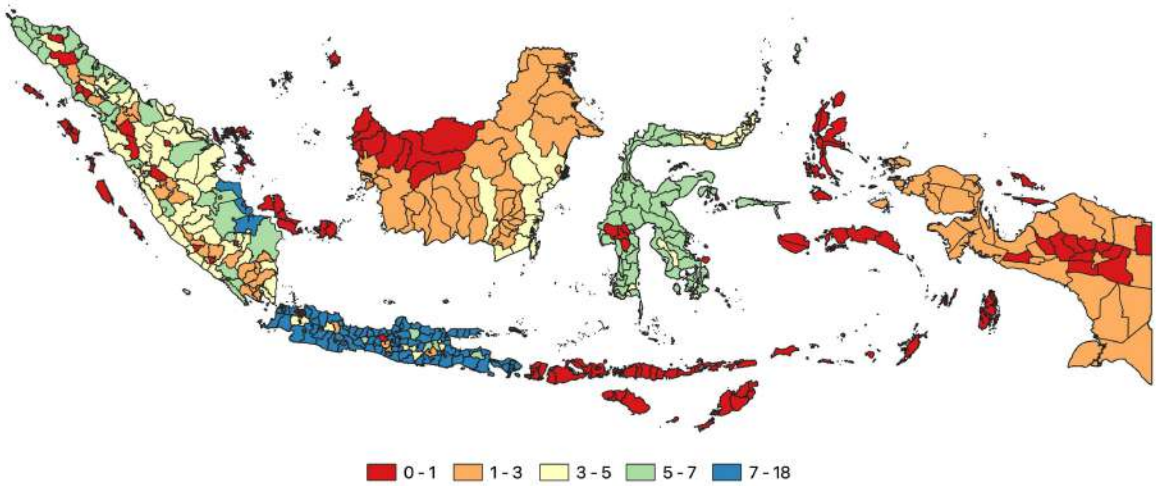
Notes: The map shows the total number of large flood events (as per the DFO archive of large flood events) that each Indonesian regency got affected by during the period 1985-2012. The internal boundaries are regency boundaries, and the legend entries represents number of large flood events experienced during the period 1985-2012. Regency boundaries correspond to the administrative divisions for the year 2020.

Figure 2: **Number of flood-affected regencies and average count of flood days**



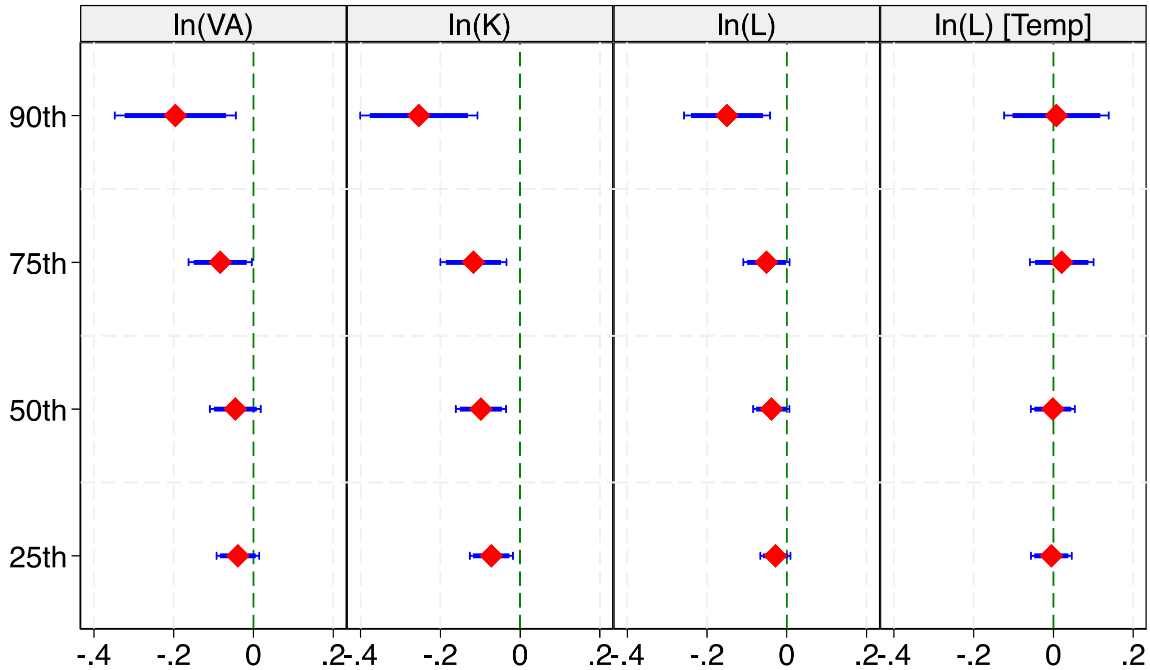
Notes: The graphs show the flooding trends within Indonesia using the information from the DFO archive of large flood events. The left (right) figure plots the count of flood-affected regencies (days) in each year for the period 1985-2012. Both the variables are trending positively over the years.

Figure 3: **Total count of large floods based on satellite observation in 2002-18 period**



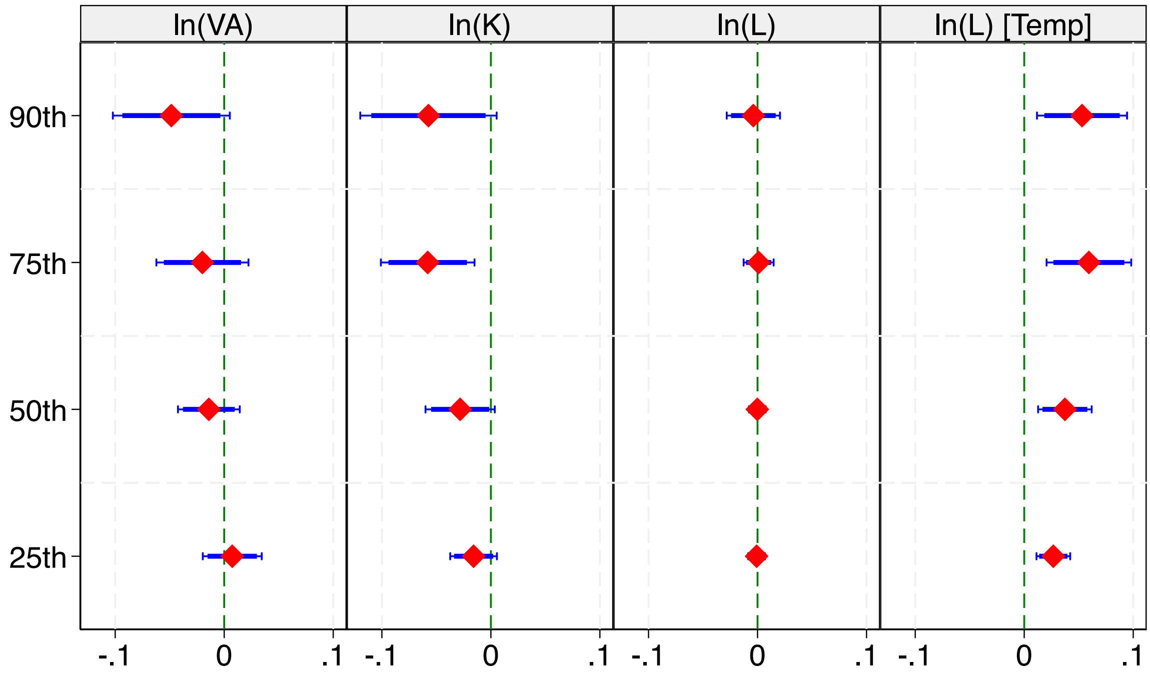
Notes: The map shows the total number of large flood events from the DFO archive of large flood events, which are confirmed using satellite observations in Tellman et al. (2021) for the period 2002-18. The internal boundaries are regency boundaries, and the legend entries represents number of large flood events experienced during the period 2002-18. Regency boundaries correspond to the administrative divisions for the year 2020.

Figure 4: **Effect of flooding on sector-regency-level variables**



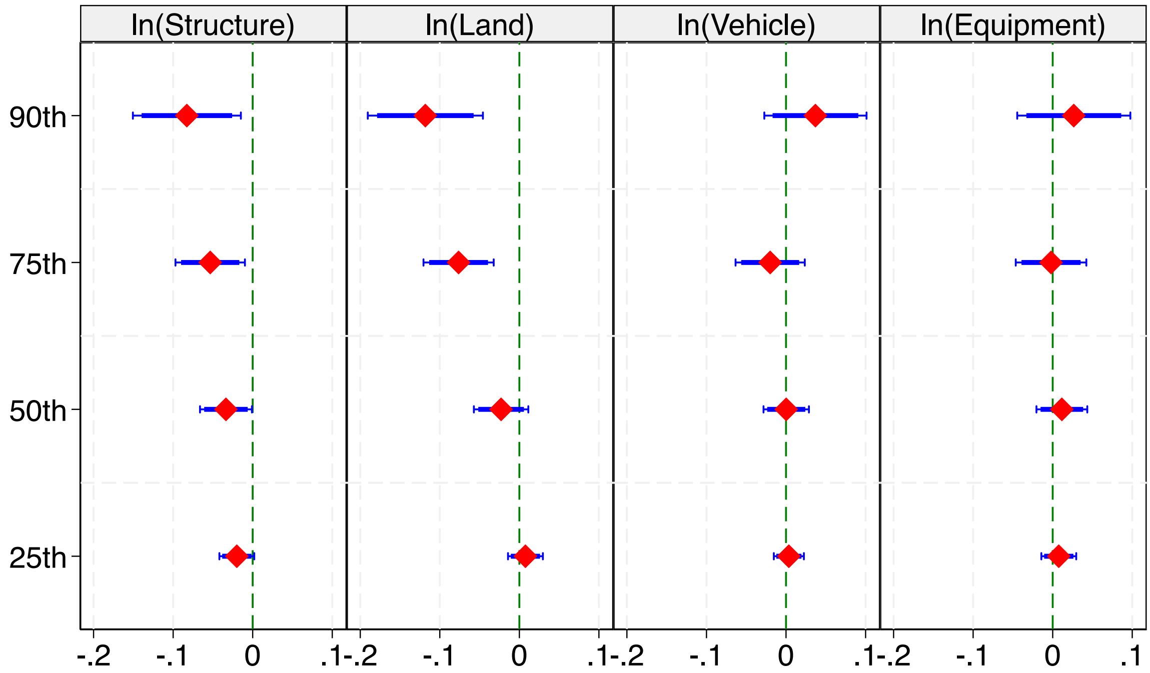
Notes: The graph presents the results of estimating Equation (III.1) for aggregate variables i.e., logarithms of total value-added (left), capital stock (centre), and labor employment (right) at the sector-regency level. To get to the aggregate variables from firm-level information, following steps are undertaken. First, The un-logged version of all the monetary variables are deflated by the whole-sale price index at the 5-digit ISIC level to reflect their real values. Second, the tails on both ends of the resulting variables are trimmed by 1% for each year to address measurement error issues. Third, the variables are then summed across sector-regency for each year using labor share weights. Finally, the variables are log-transformed and used in the regressions. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.1 in the Appendix.

Figure 5: **Effect of flooding on firm-level variables**



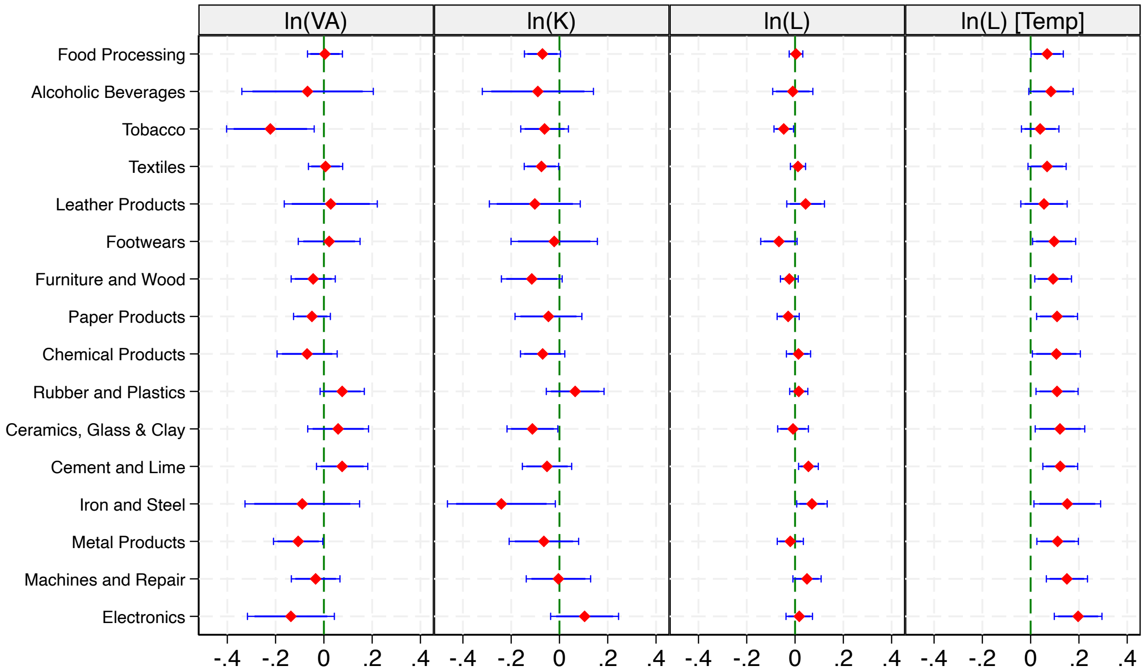
Notes: The graph presents the results of estimating Equation (III.2) for firm-level variables i.e., logarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right). The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.2 in the Appendix.

Figure 6: **Effect of flooding on firm-level capital categories**



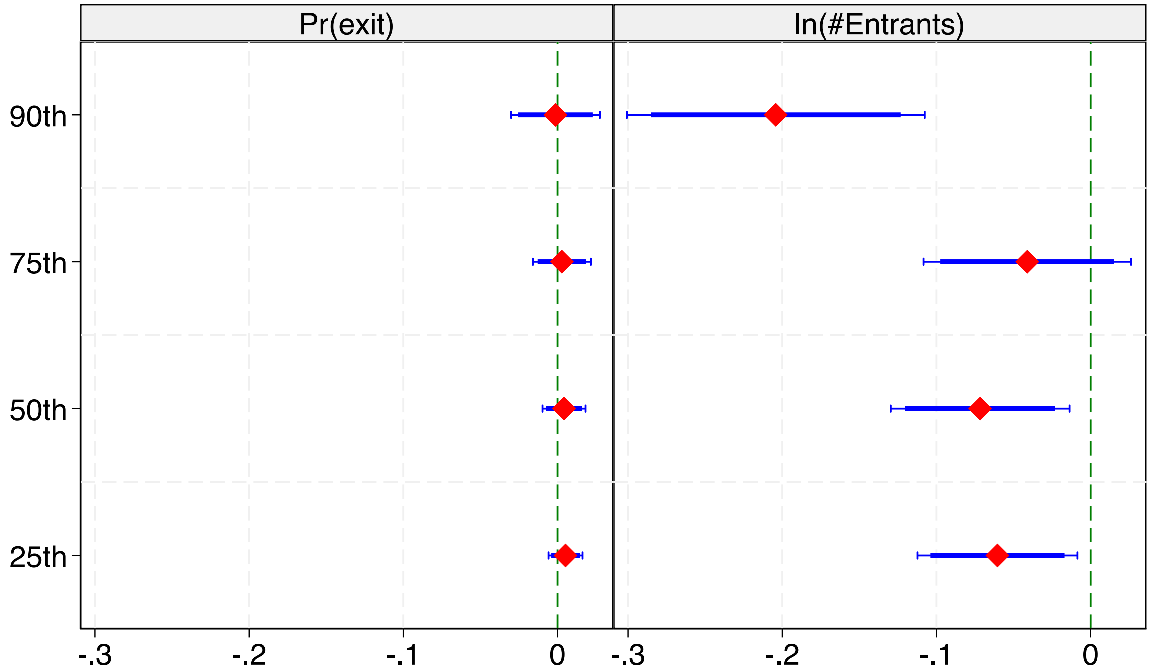
Notes: The graph presents the results of estimating Equation (III.2) for four different capital categories at the firm level. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Going left to right, first plot reports the results for value of structures, which include buildings and all man-made constructions to support the manufacturing activities within the firm. Second plot shows results on land, which is the total value of land occupied by the manufacturing firm. Third plot reports results on the value of vehicles and other transportation equipment owned by the firm. Last plot shows results for value of machinery and other production equipment employed in the firm. As mentioned in the data section, the reporting on different capital categories is not consistent over time, and that is why the number of observations are different across all four columns. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.3 in the Appendix.

Figure 7: **Effect of 90th percentile floods on firm-level variables by sectors**



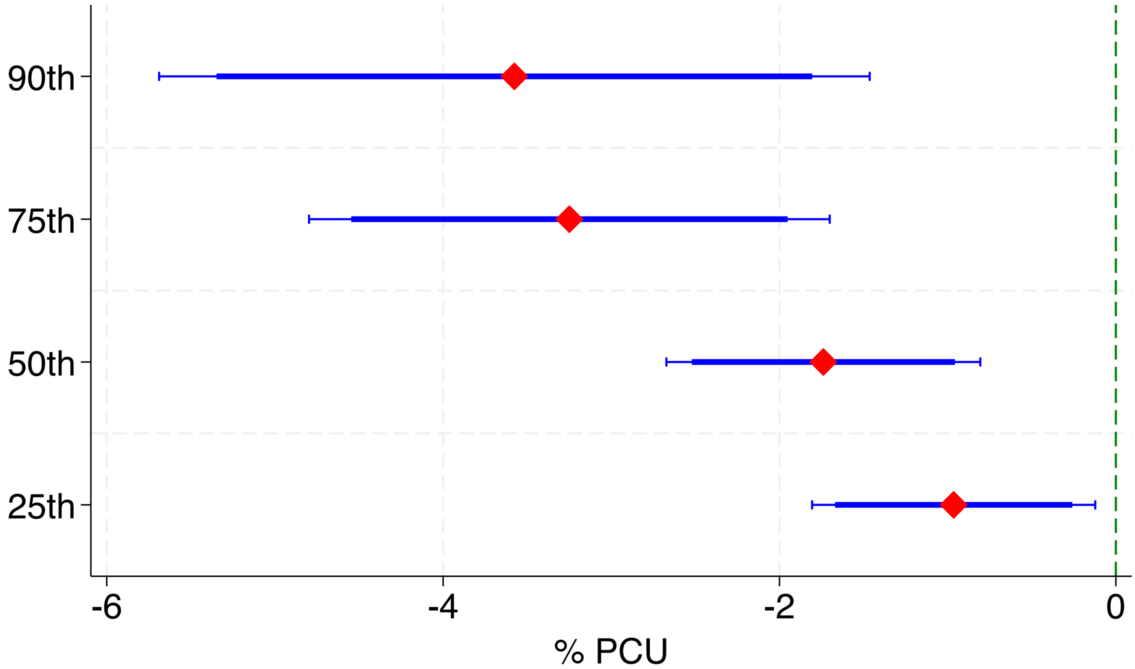
Notes: The graph presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right) using the 90th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the 2-digit ISIC manufacturing sectors. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.7 in the Appendix.

Figure 8: **Effect of flooding on firm exit and entry**



Notes: The graphs present results on firm exit and entry. Left graph presents the results of estimating Equation (III.4) for firm exit dummy, where the dummy variable takes a value of 1 in the last year of firm observation in the data. Right graph presents the results of estimating Equation (III.1) with the logarithm of number of new firms entering in a sector-regency in a given year as the dependent variable. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.8 in the Appendix.

Figure 9: **Effect of flooding on firm capacity utilization**



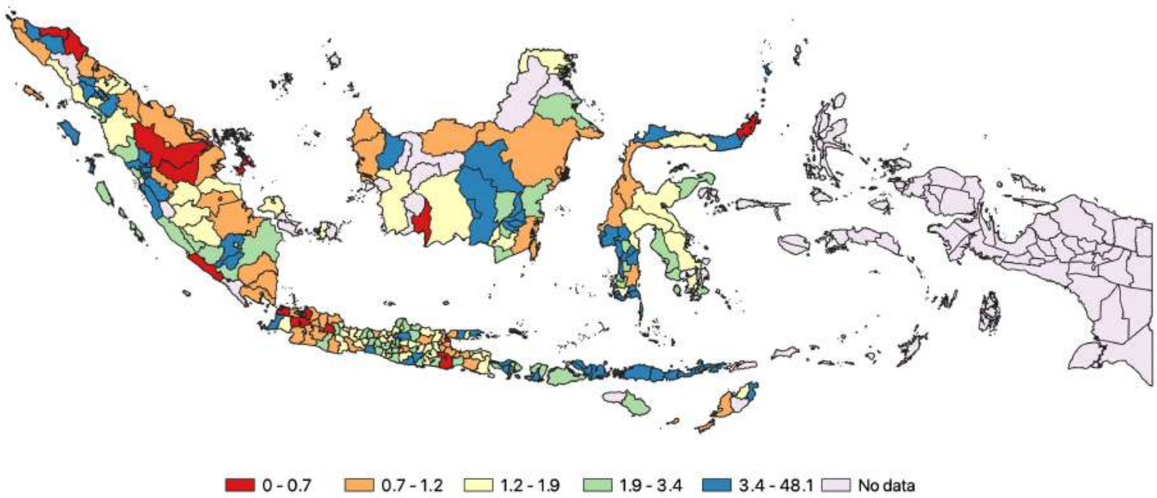
Notes: The graph presents the results of estimating Equation (III.2) for firm-level production capacity utilization (PCU). PCU measures the percentage of the potential firm capacity, in terms of production, that is realized in a given year. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.9 in the Appendix.

Table 1: **Summary of model parameters**

	(1) Parameter	(2) Level	(3) Value	(4) Method/Source
ϕ	Flooding shape	Regency-Year	-	MLE on firm-level PCU data
α	Output elasticity	Sector	-	PF estimation (Levinsohn and Petrin (2003))
η	Returns to scale	Sector	-	PF estimation (Levinsohn and Petrin (2003))
$\bar{\theta}$	Productivity scale	Regency	-	Aggregate regency manufacturing value-added
ξ	Productivity shape	Aggregate	4.514	MLE on firm-level value-added data

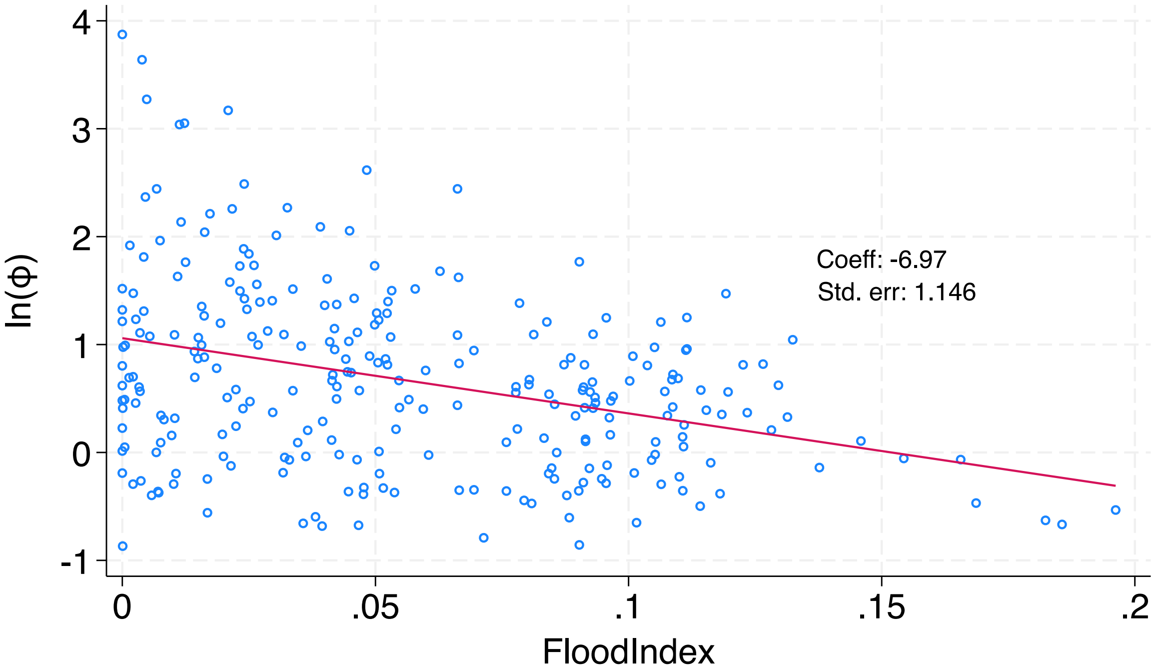
Notes: The table presents the summary of model parameters' estimation or calibration exercise. Values of some of the parameters are not included in the above table as they are too many in number to report. However, their estimation method/ calibration source is discussed in detailed in the Estimation section.

Figure 10: **Distribution of flood exposure across regencies**



Notes: The map shows the distribution of average (over years) regional flood exposure as captured by the regency-level shape parameters, ϕ_{rt}, estimated using the firm-level production capacity utilization data. In line with the properties of Pareto distribution, smaller value for a regency suggests that the regency, on average, faces more extreme flooding over time. The regencies for which flood exposure could not be estimated are shaded in pink and coded as "No data". Regency boundaries correspond to the administrative divisions for the year 1990.

Figure 11: **Pareto tail exponent versus flood index across regencies**



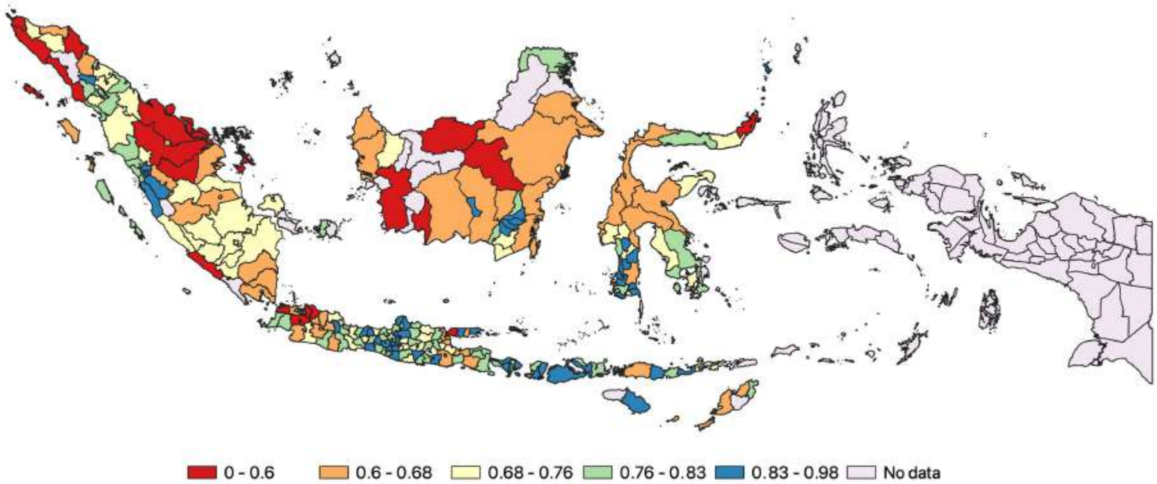
Notes: The graph plots the regency-level average regional flood exposure as captured by the regency-level shape parameters, ϕ_{rt} against the average regency-level flood index. The averages are taken across all the years in sample i.e., 1990-2012. Each dot represents one regency in Indonesia.

Table 2: **Sectoral production function parameters**

Industry name	3-digit ISIC	α_s	η_s
Food Processing	311	0.213	0.700
Food Processing 2	312	0.261	0.662
Cigarettes and Tobacco	314	0.253	0.487
Textiles	321	0.204	0.613
Leather Products	323	0.188	0.770
Manufacture of Footwear	324	0.150	0.668
Wood Products	331	0.230	0.710
Furniture	332	0.134	0.693
Paper Products	341	0.267	0.562
Paper Products, Finished	342	0.114	0.704
Chemical Products, Industrial	351	0.304	0.546
Chemical Products, Household	352	0.180	0.596
Rubber Products	355	0.104	0.625
Plastic Wares	356	0.229	0.652
Ceramics	361	0.341	0.593
Glass Products	362	0.295	0.705
Cement and Lime	363	0.223	0.687
Structural Clay Products	364	0.193	0.773
Other Non Metal Mineral Products	369	0.201	0.646
Basic Iron and Steel	371	0.265	0.742
Metal Products, Finished	381	0.194	0.744
Machines and Repair	382	0.289	0.707
Electronics	383	0.138	0.718
Motor Vehicles	384	0.268	0.601
Other Manufacturing	390	0.172	0.778

Notes: The table reports the computed values of production function parameters for each 3-digit ISIC sector. Using data from Table B.4, the computation uses the following formulae for scale parameter, $\eta_s = ln(L)coeff + ln(K)coeff$ and index on capital, $\alpha_s = \frac{ln(K)coeff}{\eta_s}$.

Figure 12: **Distribution of flood risk across regencies**



Notes: The graph plots the distribution of flood risk as captured by τ_{srt} variable, where $\tau_{srt}(\phi) \equiv \frac{\phi_{rt}}{\phi_{rt}+\alpha_s\eta_s/(1-(1-\alpha_s)\eta_s)}$ captures distortions introduced in the optimal capital installation decisions due to flooding. Both the regency-level flood exposure, ϕ_{rt} and production function parameters for each 3-digit ISIC sector (α_s, η_s) are estimated. Lower values of τ_{srt} suggest larger capital distortions due to flooding. The regencies for which flood exposure could not be estimated are shaded in pink and coded as "No data". Regency boundaries correspond to the administrative divisions for the year 1990.

Table 3: **Effect of flood shock and flood risk on capital and labor**

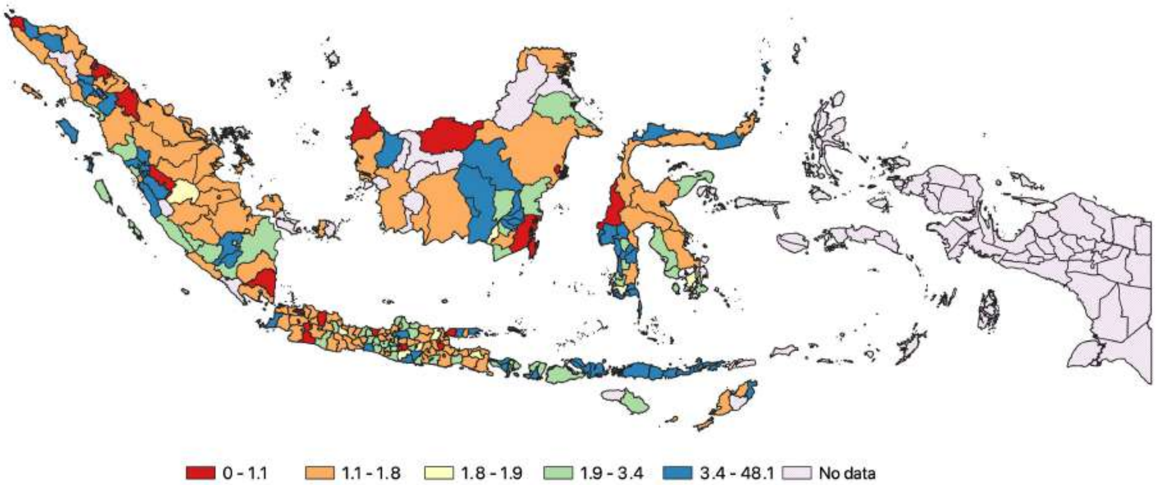
	(1) ln(K)	(2) ln(L)	(3) ln(K)	(4) ln(L)	(5) ln(K)	(6) ln(L)
ln(τ)	0.259***	-0.127***			0.249***	-0.129***
	(0.082)	(0.036)			(0.082)	(0.036)
ln(x)			-0.008***	-0.002***	-0.007***	-0.002***
			(0.001)	(0.000)	(0.001)	(0.000)
Observations	316,788	316,788	330,577	330,577	316,610	316,610
Dep. var mean	8.656	4.131	8.659	4.131	8.656	4.132
Firm FE	Y	Y	Y	Y	Y	Y
3-digit ISIC × year FE	Y	Y	Y	Y	Y	Y

Standard errors clustered at the sector-regency level are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

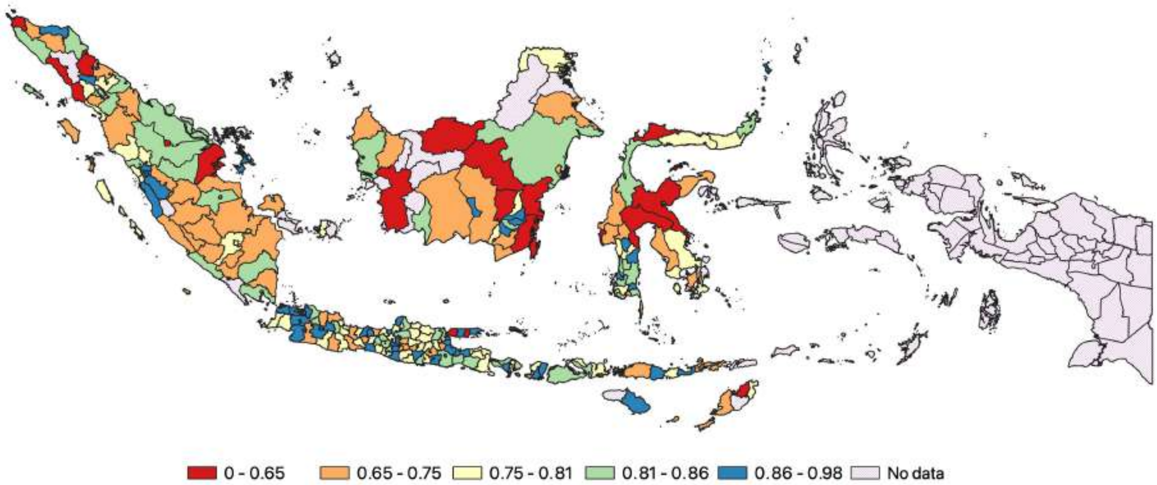
Notes: The table presents the results of estimating Equations (VI.1) and (VI.2) in Column 5 and 6 respectively. The regency-level flood exposure component of flood risk are estimated using the empirical distribution of PCU across firms within a regency for the past years. The un-logged version of value of capital stock has been deflated by the wholesale price index at the 5-digit ISIC level to reflect its real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Results reported both columns control for firm and sector × year fixed effects. Standard errors are clustered at the sector-regency level.

Figure 13: **Distribution of flood exposure across regencies after flood defenses**



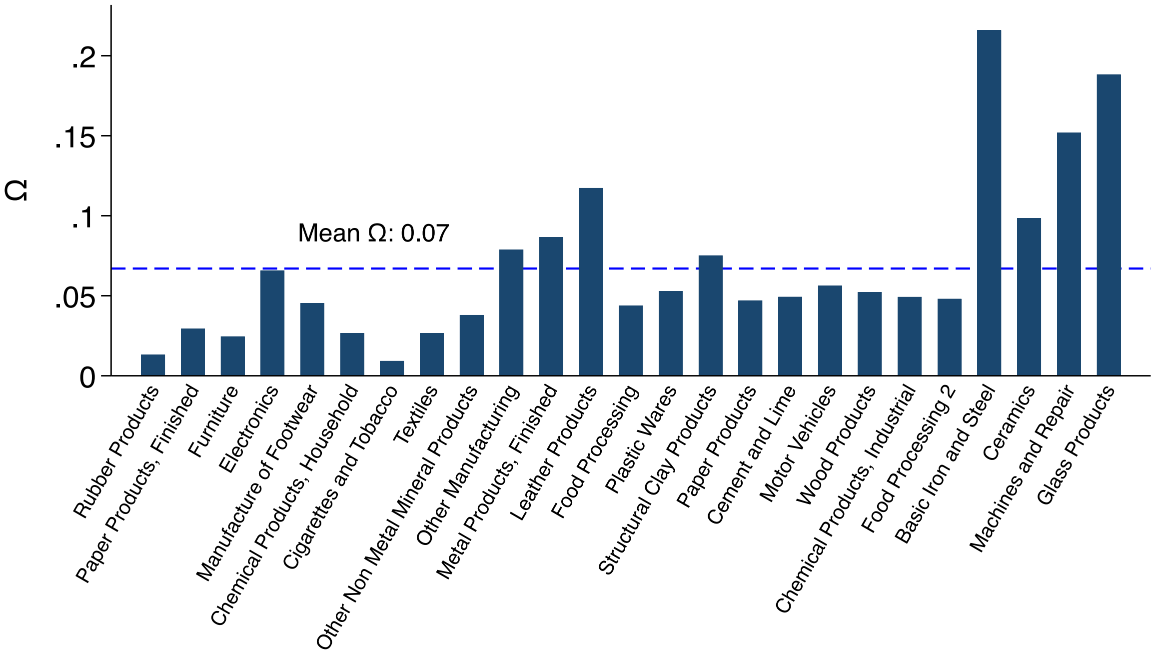
Notes: The map shows the distribution of average (over years) regional flood exposure as captured by the regency-level shape parameters, $\widetilde{\phi}_{rt}$, estimated using the firm-level production capacity utilization data after the installation of flood defenses. In line with the properties of Pareto distribution, smaller value for a regency suggests that the regency, on average, faces more extreme flooding over time. The regencies for which flood exposure could not be estimated are shaded in pink and coded as “No data”. Regency boundaries correspond to the administrative divisions for the year 1990.

Figure 14: **Distribution of flood risk across regencies after flood defenses**



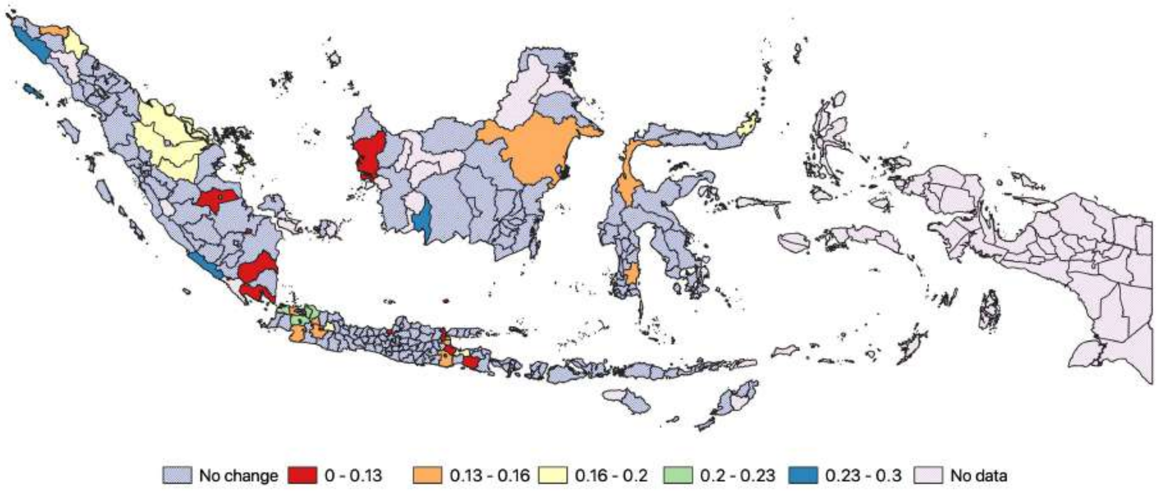
Notes: The graph plots the distribution of flood risk as captured by $\widetilde{\tau}_{srt}$ variable, where $\widetilde{\tau}_{srt}(\phi) \equiv \frac{\widetilde{\phi}_{rt}}{\widetilde{\phi}_{rt}+\alpha_s\eta_s/(1-(1-\alpha_s)\eta_s)}$ captures distortions introduced in the optimal capital installation decisions due to flooding after the installation of flood defenses. Both the regency-level flood exposure, $\widetilde{\phi}_{rt}$ and production function parameters for each 3-digit ISIC sector (α_s, η_s) are estimated. Lower values of $\widetilde{\tau}_{srt}$ suggest larger capital distortions due to flooding. The regencies for which flood exposure could not be estimated are shaded in pink and coded as "No data". Regency boundaries correspond to the administrative divisions for the year 1990.

Figure 15: **Change in output across sectors due to flood risk**



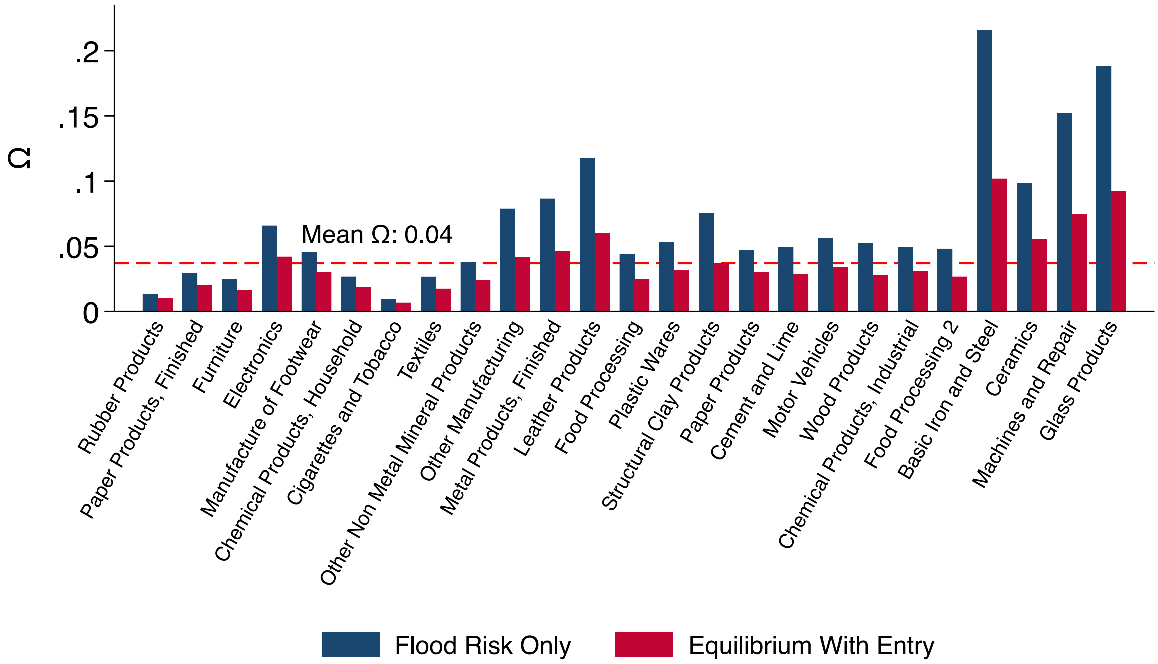
Notes: The graph plots the (log) change in aggregate output due to flood risk as outlined in Equation (VI.3) across 3-digit ISIC sectors. This represents the Flood Risk Only scenario where in the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The sectors are ranked from left to right in the increasing order of their respective capital intensities.

Figure 16: **Change in output across regencies due to flood risk**



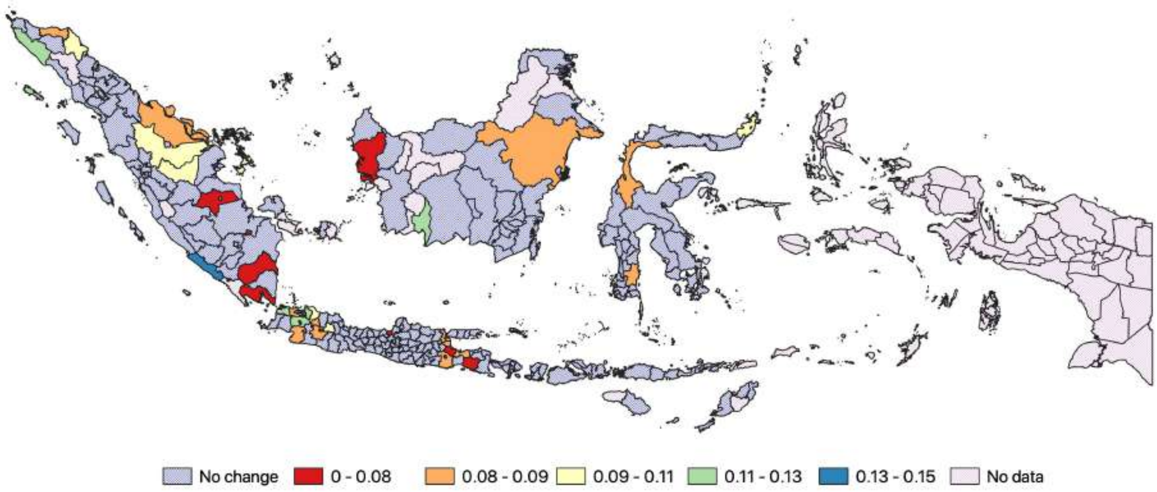
Notes: The map shows the (log) change in aggregate output due to flood risk as outlined in Equation (VI.3) across regencies. This represents the Flood Risk Only scenario where in the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The regencies for which flood exposure could not be estimated are shaded in pink and coded as "No data", while those regencies that remain unaffected by the flood defenses installation are shaded in blue and coded as "No change". The remaining legend values represent percentage change in aggregate output due to reduction in flood risk after the installation of flood defenses. Regency boundaries correspond to the administrative divisions for the year 1990.

Figure 17: **Change in output across sectors due to flood risk and equilibrium**



Notes: The graph plots the (log) change in aggregate output due to flood risk in blue and sum of (log) change in aggregate output due to flood risk and (log) change in aggregate output due to equilibrium wage adjustments and firm entry in red as outlined in Equation (VI.3) across 3-digit ISIC sectors. In the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The sectors are ranked from left to right in the increasing order of their respective capital intensities.

Figure 18: **Change in output across regencies due to flood risk and equilibrium**



Notes: The map shows the sum of (log) change in aggregate output due to flood risk and (log) change in aggregate output due to equilibrium wage adjustments and firm entry as outlined in Equation (VI.3) across regencies. This represents the Equilibrium With Entry scenario where in the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The regencies for which flood exposure could not be estimated are shaded in pink and coded as “No data”, while those regencies that remain unaffected by the flood defenses installation are shaded in blue and coded as “No change”. The remaining legend values represent percentage change in aggregate output due to reduction in flood risk and accounting for equilibrium adjustments after the installation of flood defenses. Regency boundaries correspond to the administrative divisions for the year 1990.

Appendix

A Reduced-form Evidence on Long-run Effects of Flood

A.1 Effect of Flooding on Production Variables

A.1.1 Econometric Model

As discussed in the main section of the paper, employing an event-study framework to estimate the long-term effects of flooding is not suitable in this context. Moreover, our interest lies not in assessing the impact of individual flood shocks, but rather in understanding the collective effects of flooding experienced by a region over an extended period of time. To achieve this objective, I look at the relationship between cumulative flooding and the long-term differences in aggregate variables and firm entry and exit. The long-run period is defined as starting from 1994 and ending in 2008, so it is representative of 15 years duration.[A.1]

First, cumulative flood shocks are obtained by aggregating innovations in flood index at the regency level over 1994-2008 period. Flood innovations are residuals obtained from the following linear regression on flood data for 1985-2012 period:[A.2]

$$FloodIndex_{rt} = \zeta_r + \chi_t + \varepsilon_{rt} \tag{A.1}$$

The residuals, $\widehat{\varepsilon}_{rt}$, are then summed across regencies over the period 1994-2008 to compute regency-level cumulative flood shocks as below:

$$CumulativeFlood_r = \sum_{t=1994}^{2008} \widehat{\varepsilon}_{rt}$$

I employ the first-difference (FD) estimator on the aggregate variables in 1994 and 2008, where the treatment variable is cumulative flood that takes zero as its initial value in 1994. I estimate the following econometric specification at the sector-regency

A.1. There is no particular reason for choosing this epoch, except that it coincides with some of the major flood events in Indonesia, including the 2000 Sumatra floods, which killed 120 and affected around 600,000 people and 2007 Jakarta Floods, which resulted in $850 million worth of monetary loss, 80 deaths, and over 500,000 human displacements. Results are robust to choosing alternate epochs.

A.2. This formulation is simply used to match the specifications employed for studying contemporaneous effect of flooding on economic variables. Results are robust to adding lags of flood index or regency-level linear time trends.

or regency level:

$$\Delta y_{sr} = \upsilon + \beta \Delta CumulativeFlood_r + \nu_s + \varepsilon_{sr} \tag{A.2}$$

where Δy_{sr} denotes the difference in the logarithm of regency or sector-regency level value-added, capital stock, or labor employed between the years 1994 and 2008. β is the coefficient capturing relationship between the change in the cumulative flood shock from 0 to its end-of-period value and change in the aggregate variable. υ captures the linear time trend.[A.3]

A.1.2 Results and Discussion

I use the value at the 90th percentile from Table B.6 for interpreting the strength of the relationship in all cases. Table B.7 reports the results from estimating Equation (A.2). Interpreting the magnitude using estimates from Columns 4-6, a 90th percentile cumulative flood shock at the regency level is associated with 18.9%, 20.9%, and 16.2% decrease in the aggregate value-added, capital stock, and labor employment respectively.

A.2 Effect of Flooding on Firm Exit and Entry

A.2.1 Econometric Model

To estimate the relationship between cumulative flood shocks and the exit decision of firms, I first define a cumulative flood shock at firm level by aggregating the flood index over the long-run analysis period for each firm. Obviously, a typical firm does not continue operating in all these years, so the last year of its observation in data is assumed to be the exit year. Under these assumptions, the firm-specific cumulative flood shock is defined as below:

$$CumulativeFlood_{irt} = \frac{1}{t - T^i_{start} + 1} \sum_{s=T^i_{start}}^{t \le T^i_{end}} FloodIndex_{rs}$$

A.3. The estimation leverages cross-sectional variation in cumulative flood exposure across regencies. Since, the flood innovations always sum to zero for each regency over the complete 1985-2012 period, the realizations over the block of period 1994-2008 randomly sort regencies based on the realized exposures within this period. Figure B.17 illustrates this point by showing the realized values of flood innovations in circles and their moving sums for different years.

where T^i_{start} and T^i_{end} are the entry and exit years for firm i.[A.4] I then estimate the following relationship:

$$y_{isrt} = \upsilon + \beta CumulativeFlood_{irt} + \iota X_{isrt} + \zeta_r + \nu_{st} + \psi_{pt} + \varepsilon_{isrt} \tag{A.3}$$

where y_{isrt} is an exit dummy for firm i, belonging to 2-digit ISIC sector s, located in regency r, in year t and other terms have the same interpretation as Equation (III.4). Below is the summary statistics table on the cumulative shock variable.

A.2.2 Results and Discussion

Column 1 of Table B.8 reports the results from estimating Equation (A.3) for all the firms that start and end their life in the period 1994-2008. Unlike temporary shocks, cumulative flood shocks do lead to firm exits in the long run. In particular, a 90th percentile cumulative flood shock increases the firm exit probability by 0.26%.

Columns 2 and 3 of Table B.8 report the results from estimating Equation (A.2) with difference in the logarithm of number of firms operating in a regency or sector-regency between 2008 and 1994 as the dependent variable. Similar to the contemporaneous analysis, the evidence points towards the firms avoiding flood-prone locations when setting up their operations, a 90th percentile cumulative flood shock decreases the number of firms at the sector-regency (regency) level by 6.6% (8.5%).

A.4. Only those firms, which have their start and end years fall in the period 1994-2008 are included in the analysis.

B Additional Tables and Figures

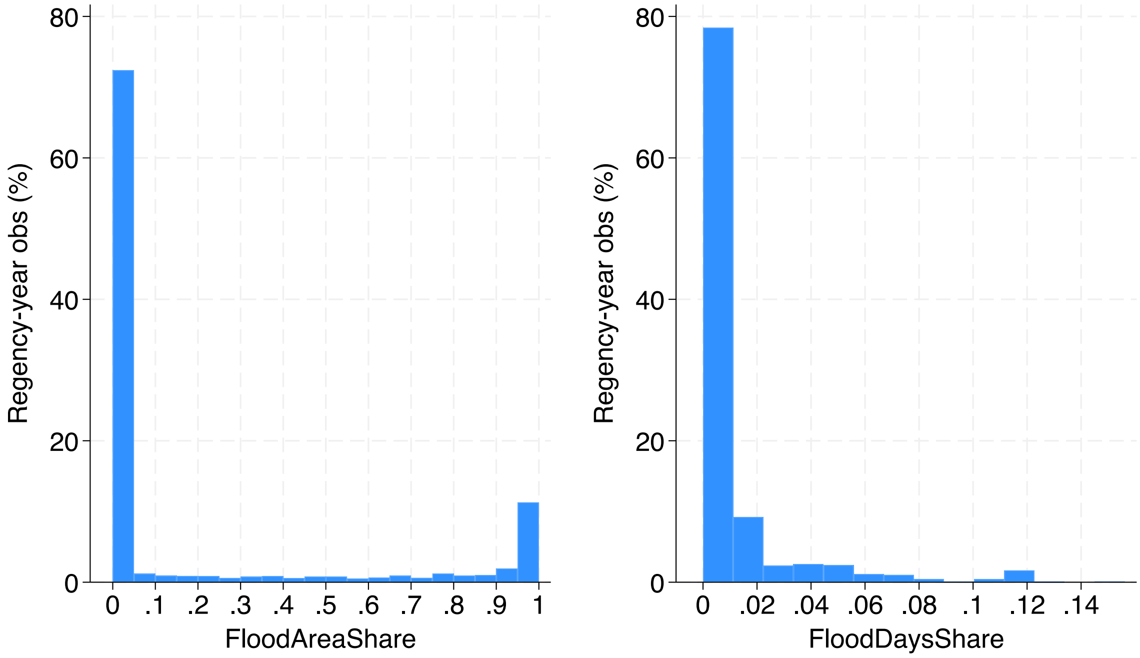
Table B.1: **Relationship between flooded-affected and flooded area share**

	(1) Affected area	(2) Affected area	(3) Affected area
Flooded area	0.300	0.513***	1.422***
		(0.026)	(0.240)
Observations	21,074	21,074	21,074
Regency FE	-	-	Y
Flood FE	-	-	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

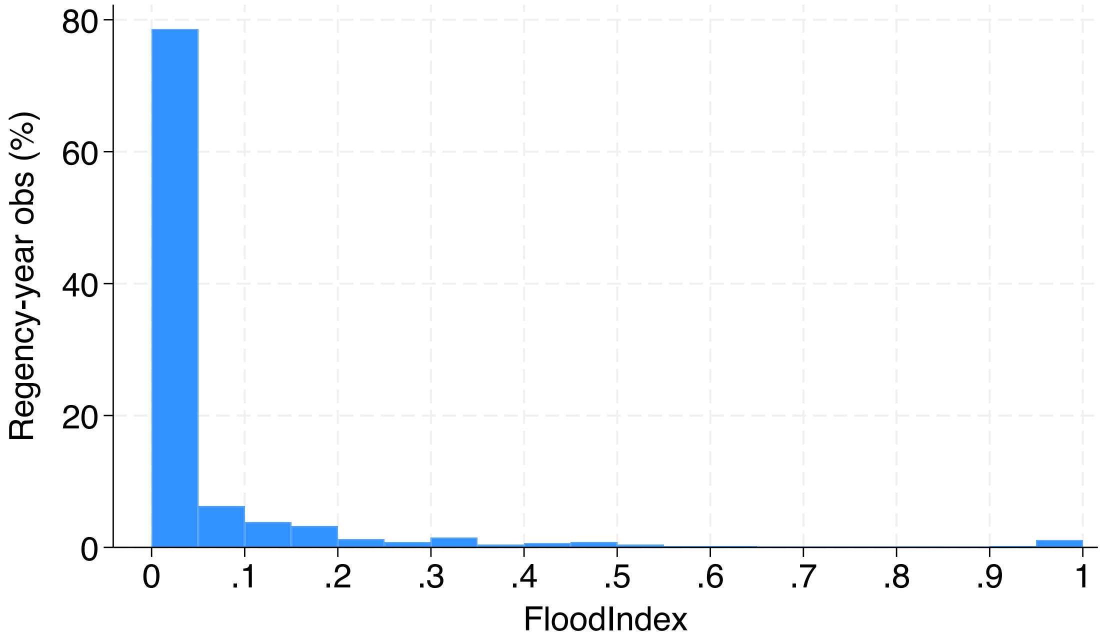
Notes: The table presents results on establishing relationship between flooded and flood-affected area in Indonesia. Flooded area metric at the regency level is constructed using detailed inundation maps available for 41 flood events within Indonesia. Flood-affected area within regencies for these 41 events are obtained from the polygons available in the DFO archive. Both these variables are then normalized by the total area of regency to represent area shares. Column 1 reports the Kendall's Tau-b coefficient of association between flooded and flood-affected area share. The coefficient ranges from -1 (perfect inversion) to +1 (perfect agreement) with 0 indicating no association. Column 2 reports the Somers' D coefficient, which also has the same range as Tau-b coefficient, but additionally it comes with standard errors on the coefficient of association that are generated using the jackknife variance calculation method. Column 3 reports the result of following a parametric approach by using fixed-effects regression analysis in which the flood-affected area is regressed on the flooded area share. Standard errors are clustered at the regency level.

Figure B.1: **Distribution of regencies on flood-affected area and days share**



Notes: The graphs show the distribution of components of flood index across regency-year pairs. The left graph presents the distribution of regency-level flood-affected area share that is computed using the polygons provided in the DFO archive of large flood events. The right graph presents the share of days in a year that a regency remains flooded, where the number of days are computed using the start and end dates for each flood event. In case of multiple flood events affecting a regency in a year, the flood-affected area is the average across all flooding episodes and the flood days are total count of days that any area in the regency remains flooded.

Figure B.2: **Distribution of flood index across regency-year pairs**



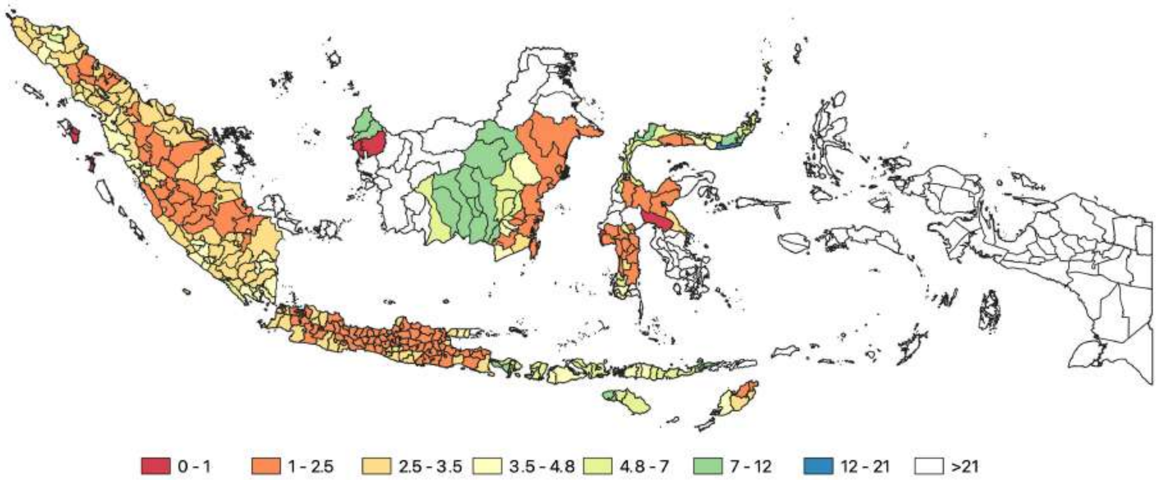
Notes: The graph shows the distribution of flood index across all regency-year pairs. Flood index is a rescaled product of flood-affected area share in a regency and flood days share in a year. The rescaling is done so that the index lies in the range 0 to 1. Most of the regency-year pairs remain unaffected by flooding with some extreme flood events affecting few of them located on the right tail of the distribution.

Table B.2: **Summary statistics on flood index**

(1)	(2)	(3)	(4)	(5)	(6)
Mean	Std. Dev.	25th Pctile	50th Pctile	75th Pctile	90th Pctile
0.186	0.235	0.04	0.098	0.219	0.455

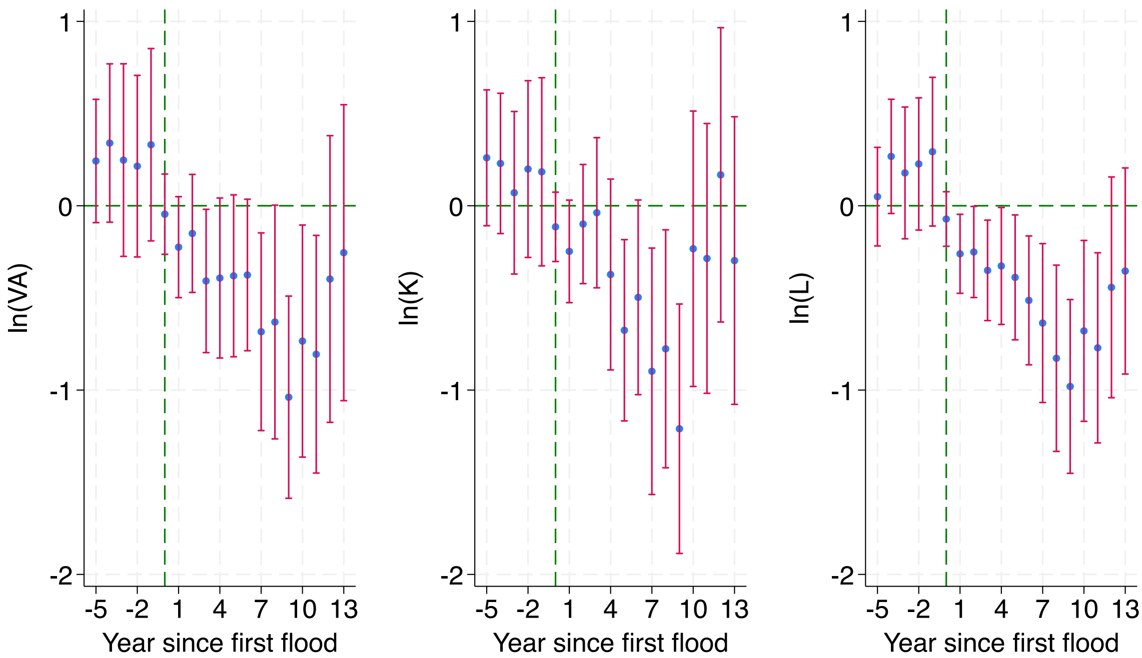
Notes: The table presents the summary statistics on flood index at the regency level that is defined by the rescaled product of flood-affected area share in a regency and flood days share in a year. Only those regency-year pairs that have non-zero flood index values are considered to derive these statistics.

Figure B.3: **Average time interval between two successive floods**



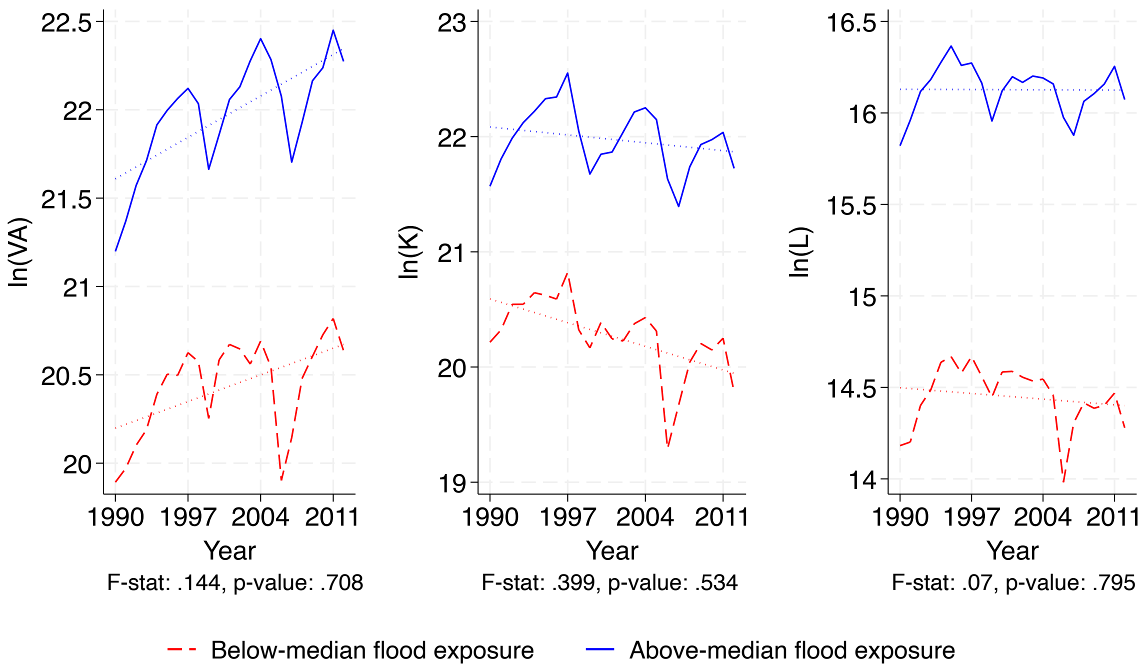
Notes: The map shows the average time interval in terms of years between two successive flood events in a regency during the period 1985-2012. The internal boundaries are regency boundaries, with the legends denoting the number of years between two successive flooding episodes. Most of the regencies located on the islands of Java and Sumatra witness a large flood almost every alternate year.

Figure B.4: **Effect of first flood on aggregate variables**



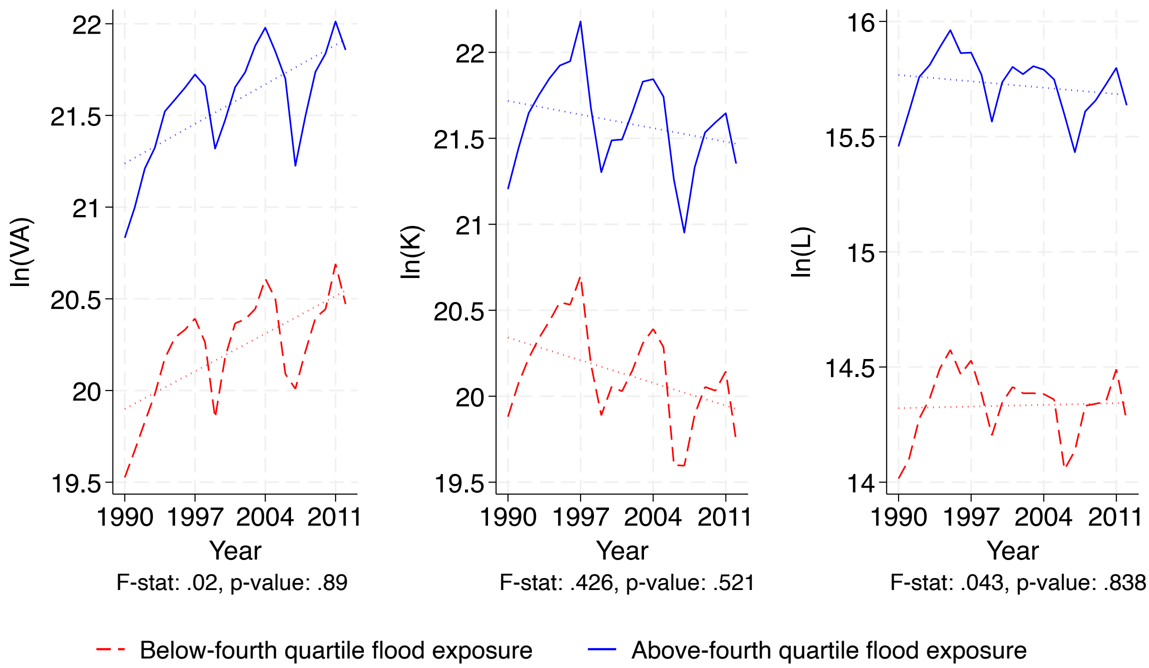
Notes: The graphs show the effect of "first" flood on the regency-level aggregate variables i.e., logarithms of total value-added (left), capital stock (middle), and labor employment (right) using the imputation-based difference-in-differences estimator proposed by Borusyak, Jaravel, and Spiess (2024). The first flood for a regency is defined as the first year in which the regency witnessed a flooding episode in the sample. Since the outcomes data starts in 1990 but the floods can be tracked since 1985, only those regencies that did not witness any flood event in the period 1985-89 are included in the analysis. To get to the aggregate variables from firm-level information, following steps are undertaken. First, The un-logged version of all the monetary variables are deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values. Second, the tails on both ends of the resulting variables are trimmed by 1% for each year to address measurement error issues. Third, the variables are then summed across regency for each year using labor share weights. Finally, the variables are log-transformed and used in the estimation. Standard errors are clustered at the regency level and whiskers on the point estimates show 90% confidence intervals.

Figure B.5: **Variable trends across high and low flood-prone regions**



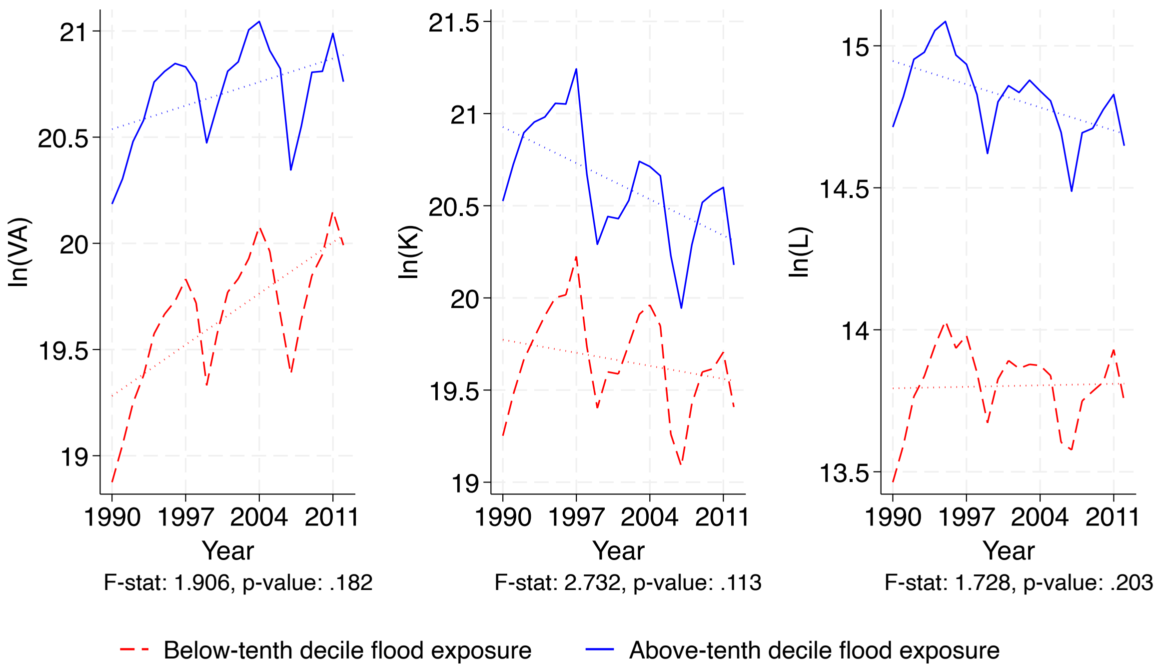
Notes: The graphs present a test of the parallel trends assumption for the aggregate variables i.e., logarithms of total value-added (left), capital stock (middle), and labor employment (right) across Indonesian regencies, which are classified as below-median or above-median flood exposure. Flood exposure of a regency is defined as the average flood index during the period 1990-2012. Two bins are then created using this measure by splitting the distribution of regencies at the median value of flood exposure. The results of F-test for the equality of growth rates of all variables across the two bins are also reported at the bottom of each graph.

Figure B.6: **Variable trends across high and low flood-prone regions**



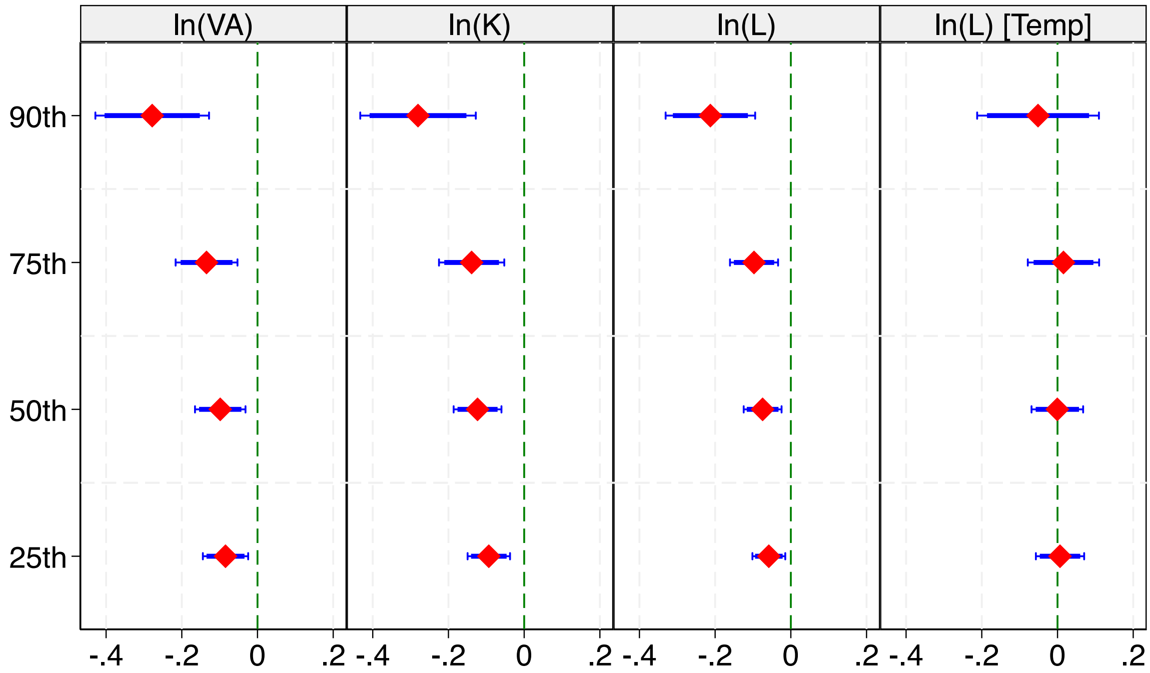
Notes: The graphs present a test of the parallel trends assumption for the aggregate variables i.e., logarithms of total value-added (left) and capital stock (right) across Indonesian regencies, which are classified as below-fourth quartile or above-fourth quartile flood exposure. Flood exposure of a regency is defined as the average flood index during the period 1990-2012. Two bins are then created using this measure by first splitting the distribution of regencies by quartiles of flood exposure, and then taking the average of first three quartiles to define below-fourth quartile flood exposure and the fourth quartile is defined as above-fourth quartile flood exposure. The results of F-test for the equality of growth rates of all variables across the two bins are also reported at the bottom of the graph.

Figure B.7: **Variable trends across high and low flood-prone regions**



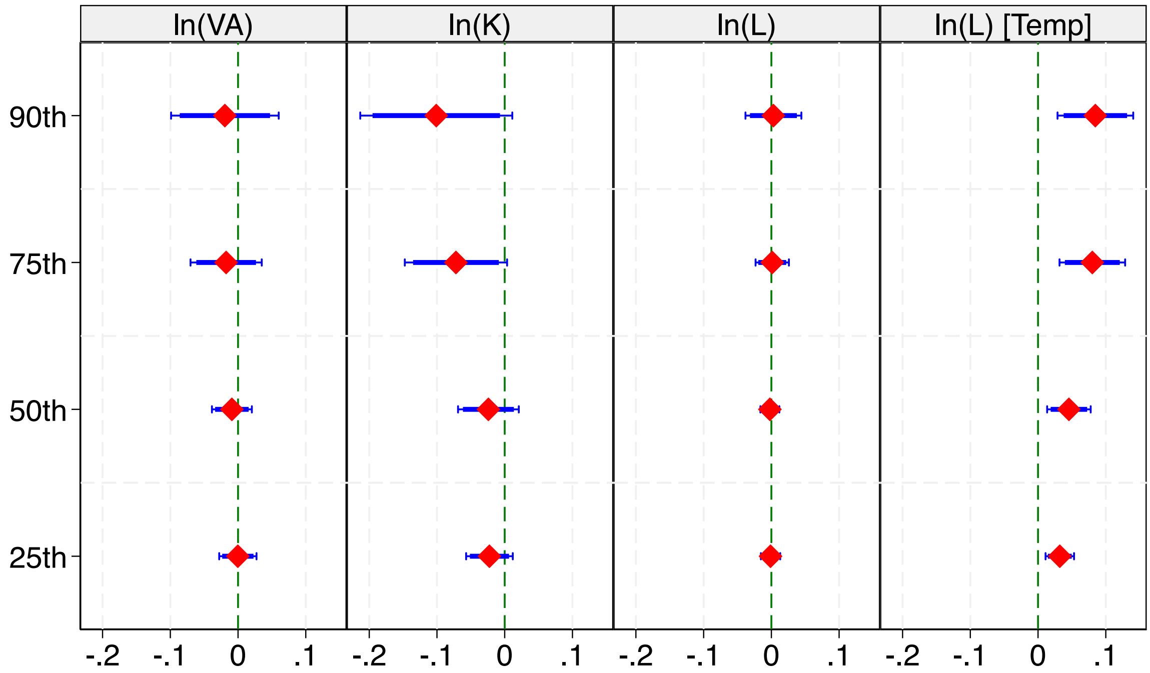
Notes: The graphs present a test of the parallel trends assumption for the aggregate variables i.e., logarithms of total value-added (left) and capital stock (right) across Indonesian regencies, which are classified as below-tenth decile or above-tenth decile flood exposure. Flood exposure of a regency is defined as the average flood index during the period 1990-2012. Two bins are then created using this measure by first splitting the distribution of regencies by deciles of flood exposure, and then taking the average of first nine deciles to define below-tenth decile flood exposure and the tenth decile is defined as above-tenth decile flood exposure. The results of F-test for the equality of growth rates of all variables across the two bins are also reported at the bottom of the graph.

Figure B.8: **Effect of flooding on sector-regency-level variables**



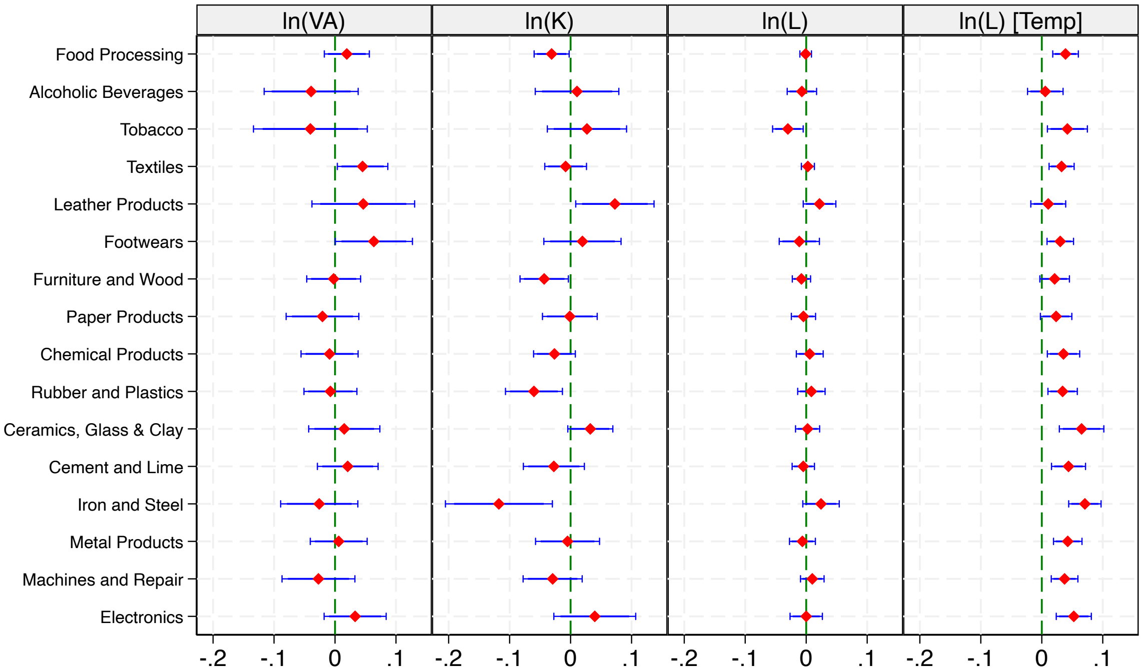
Notes: The graph presents the results of estimating Equation (III.1) for aggregate variables i.e., logarithms of total value-added (left), capital stock (centre), and labor employment (right) at the sector-regency level using only those regencies for which data is available for at least 20 years. To get to the aggregate variables from firm-level information, following steps are undertaken. First, The unlogged version of all the monetary variables are deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values. Second, the tails on both ends of the resulting variables are trimmed by 1% for each year to address measurement error issues. Third, the variables are then summed across sector-regency for each year using labor share weights. Finally, the variables are log-transformed and used in the regressions. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level.

Figure B.9: **Effect of flooding on firm-level variables**



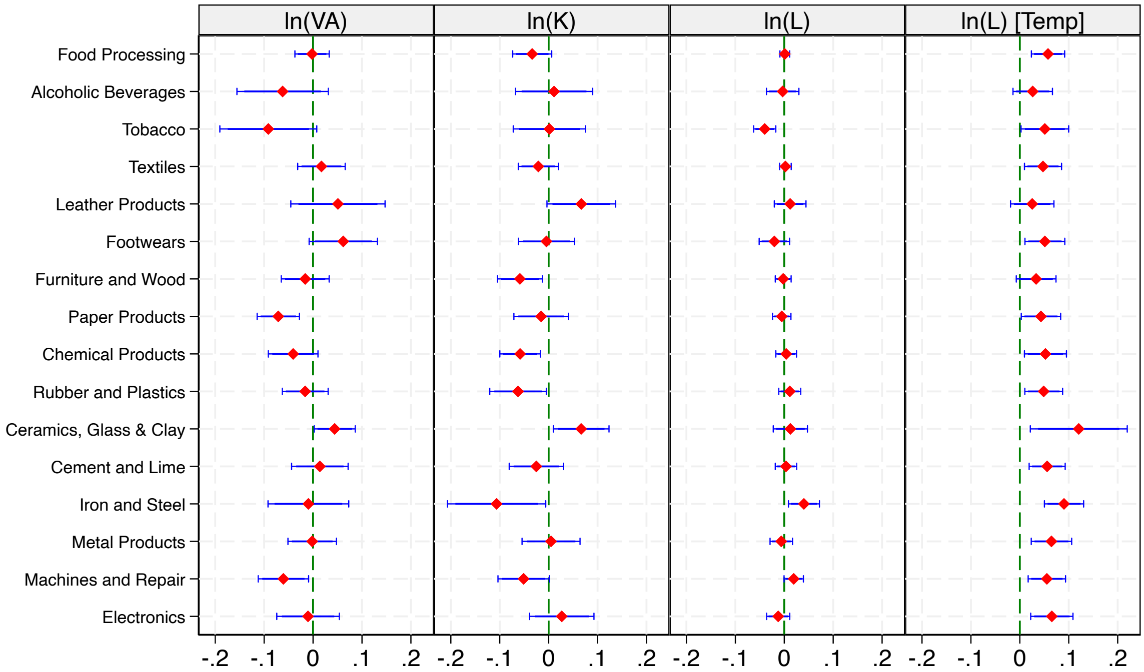
Notes: The graph presents the results of eslogarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right) using only those firm observations for which data is available for at least 20 years. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the percentiles of flood index for which dummy is used in the regression. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level.

Figure B.10: **Effect of 25th percentile floods on firm-level variables by sectors**



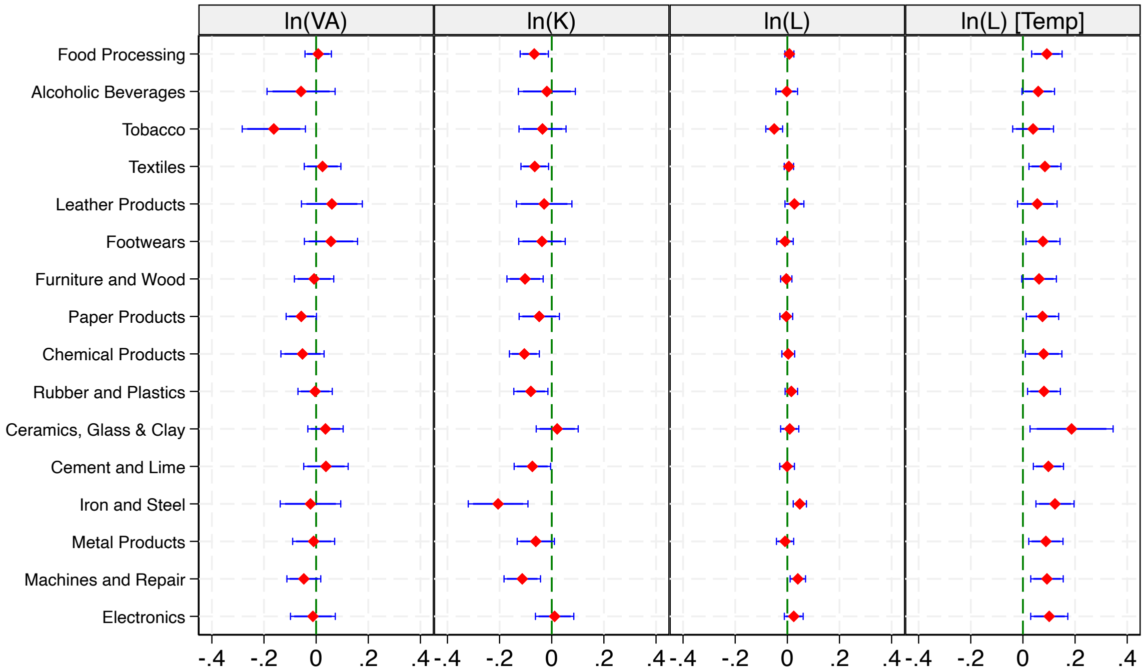
Notes: The graph presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right) using the 25th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the 2-digit ISIC manufacturing sectors. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.4 in the Appendix.

Figure B.11: **Effect of 50th percentile floods on firm-level variables by sectors**



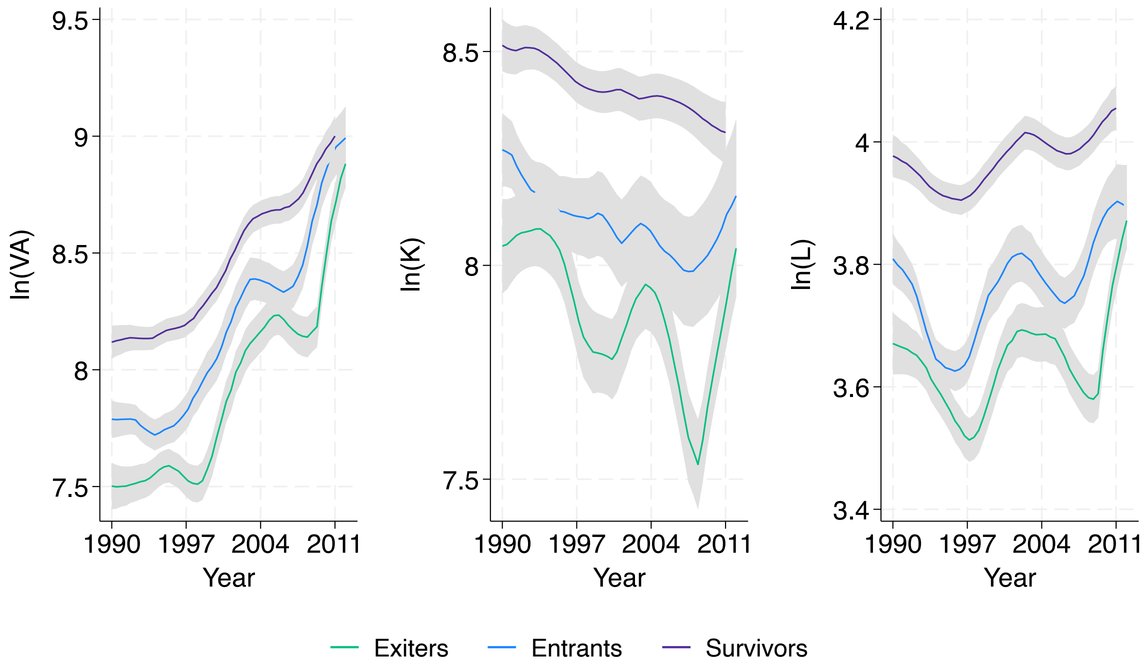
Notes: The graph presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right) using the 50th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the 2-digit ISIC manufacturing sectors. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.5 in the Appendix.

Figure B.12: **Effect of 75th percentile floods on firm-level variables by sectors**



Notes: The graph presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added (left), capital stock (second-left), permanent labor employment (second-right), temporary labor employment (right) using the 75th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. The labels on y-axis represent the 2-digit ISIC manufacturing sectors. The control observations in all cases are regency-year pairs that are not flooded. 90 and 95% confidence intervals are shown in thick and thin blue lines respectively over the point estimates. Standard errors are clustered at the regency level. Detailed regression results are reported in Table C.6 in the Appendix.

Figure B.13: **Aggregate variables for entering, exiting, and surviving firms**



Notes: The graphs plot the average of the three variables viz. logarithm of value-added (left), capital stock (middle), and labor employment (right) across regencies over time for three groups of firms: exiters, entrants, and survivors. A firm's year of exit is its last year of observation in the data, entry year is its first year of observation, and all the years in between are its years of survival.

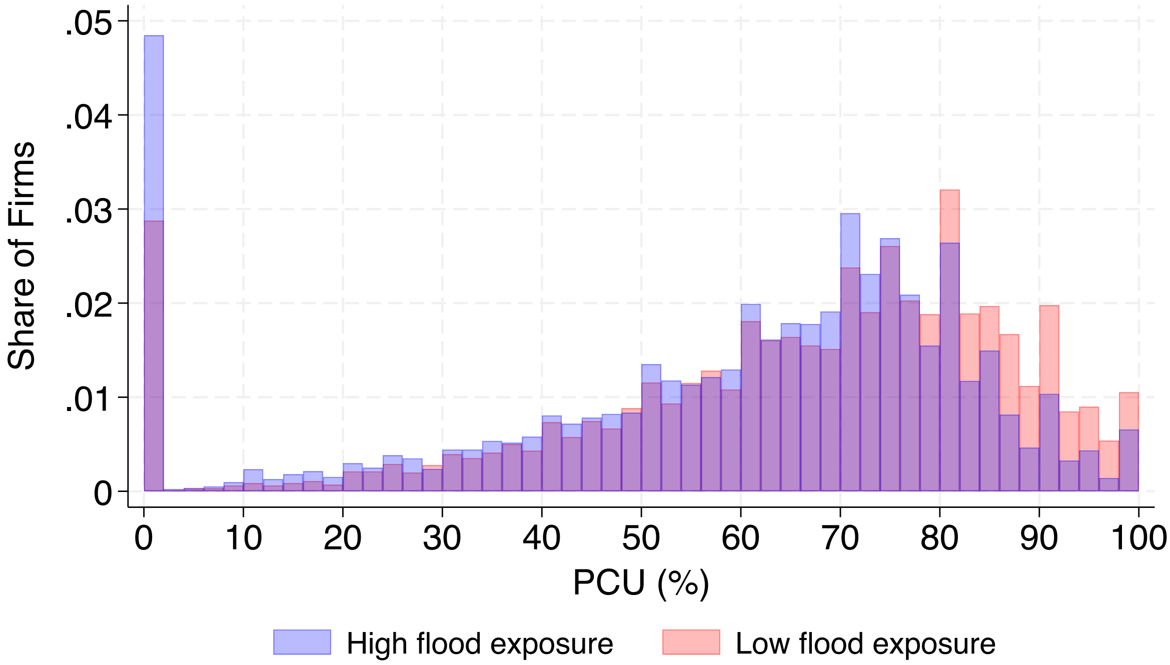
Table B.3: **Relationship between flooding and firm entry rate**

	(1) Plant entry	(2) Plant entry
FloodIndex	-0.299**	
	(0.144)	
FloodRisk		-0.247*
		(0.131)
Observations	273	211
R-squared	0.025	0.031

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the relationship between flood variables and regency-level firm entry rates. The entry rate is defined as the ratio of count of entrants in the current year over the count of survivors for the previous year. $FloodIndex$ is the average flood index at the regency level for the period 1990-2012 and has been rescaled to lie in the interval [0,1]. $FloodRisk$ variable reflects regency-level flood risk for the year 2013 as published by the IRBI in their annual report and also lie in the interval [0,1]. Each regency is given a risk score between 0 and 1 depending on its hazard profile, vulnerability index, and resilience to deal with destructive effects of flooding. The total number of regencies used in the flood risk analysis are smaller because the flood risk scores are unavailable for some regencies. However, the omission seems to be orthogonal to the flood risk, since some of the omitted regencies are also at high flood risk as outlined in the published report (IRBI 2013). Standard errors are clustered at the regency level.

Figure B.14: **Firm capacity utilization across low and high flood-prone regencies**



Notes: The graph plots the distribution of firm-level production capacity utilization (PCU) across low and high flood-prone regencies in Indonesia. PCU is defined as the percentage of available production capacity utilized by a firm in each year. Flood exposure of a regency is the average flood index over the period 1990-2012. The distribution of regencies on average flood index is split into 20 quantiles. The last quantile is defined as high flood exposure and the first 10 quantiles are defined as low flood exposure regencies. This sampling choice ensures that around 20% of regencies fall in each of the two bins. The core qualitative finding in the above graph is robust to changing this sampling criteria.

Table B.4: **Sectoral output elasticities of capital and labor**

Industry name	(1) 3-digit ISIC	(2) log(L) coeff	(3) log(L) se	(4) log(K) coeff	(5) log(K) se	(6) #Observations	(7) #Plants	(8) #Years (avg)
Food Processing	311	.551	.0002	.149	.0001	40914	5955	6.9
Food Processing 2	312	.489	.0003	.173	.0002	31821	4391	7.2
Cigarettes and Tobacco	314	.364	.0008	.123	.0011	16494	2638	6.3
Textiles	321	.488	.0003	.125	.0002	34911	4855	7.2
Leather Products	323	.625	.0011	.145	.0015	3094	469	6.6
Manufacture of Footwear	324	.568	.0007	.1	.0007	5491	886	6.2
Wood Products	331	.547	.0002	.163	.0001	22945	3877	5.9
Furniture	332	.6	.0003	.093	.0001	20661	3421	6
Paper Products	341	.412	.0018	.15	.001	5162	660	7.8
Paper Products, Finished	342	.624	.0011	.08	.0003	8522	1139	7.5
Chemical Products, Industrial	351	.38	.0016	.166	.0006	5706	781	7.3
Chemical Products, Household	352	.489	.0008	.107	.0004	8897	1006	8.8
Rubber Products	355	.56	.0008	.065	.0004	6494	734	8.8
Plastic Wares	356	.503	.0004	.149	.0002	14934	1979	7.5
Ceramics	361	.391	.0036	.202	.0026	1282	142	9
Glass Products	362	.497	.0026	.208	.0064	945	128	7.4
Cement and Lime	363	.534	.0014	.153	.0006	7965	1221	6.5
Structural Clay Products	364	.624	.0007	.149	.0002	15704	1983	7.9
Other Non Metal Mineral Products	369	.516	.0018	.13	.0007	4499	700	6.4
Basic Iron and Steel	371	.545	.003	.197	.002	2685	346	7.8
Metal Products, Finished	381	.6	.0006	.144	.0003	13941	1910	7.3
Machines and Repair	382	.503	.0025	.204	.0034	4475	593	7.5
Electronics	383	.619	.0011	.099	.0015	4389	736	6
Motor Vehicles	384	.44	.0012	.161	.0006	7959	1070	7.4
Other Manufacturing	390	.644	.0005	.134	.0004	8111	1272	6.4

Notes: The table presents the production function estimation results for each 3-digit ISIC sector by employing the Levinsohn and Petrin (2003) methodology in Stata through the `prodest` package. Columns 2 & 3 (4 & 5) report output elasticity of labor (capital) coefficient and standard errors respectively. Column 6 reports the total number of observations used in the estimation with Column 7 and 8 reporting statistics on number of firms used and average number of years observed for each firm in the estimation.

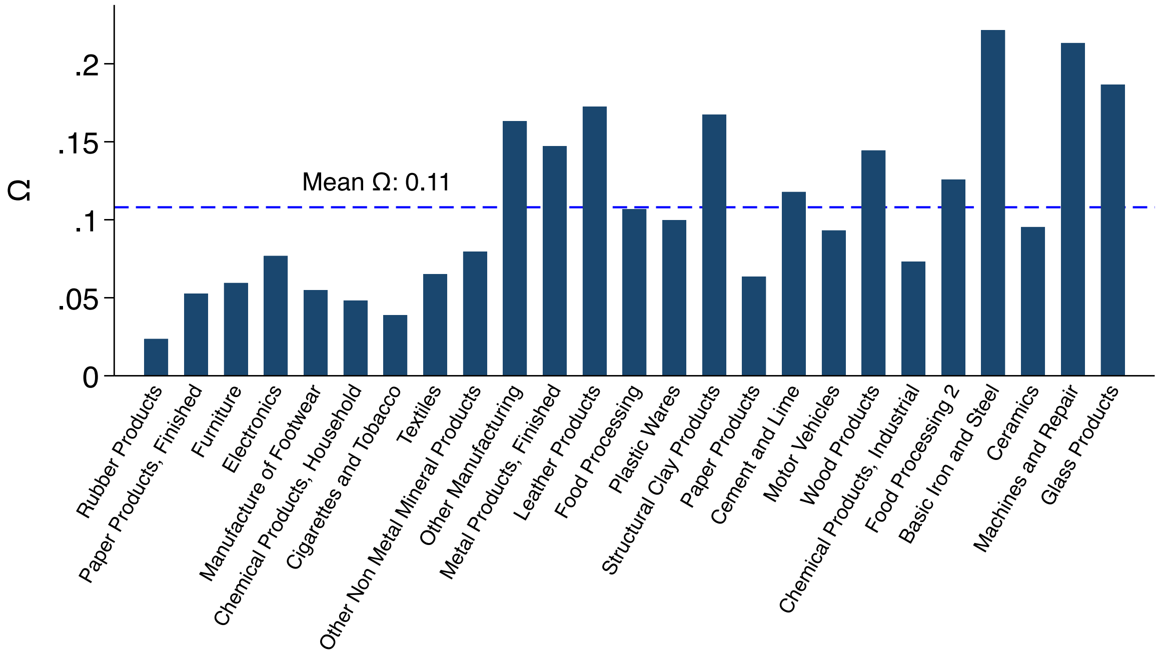
Table B.5: **Effect of flooding on firm capacity utilization by firm size**

	(1) % PCU	(2) % PCU
FloodIndex	-3.064**	-3.078**
	(1.268)	(1.301)
Large (L) Firm × FloodIndex	0.960	
	(0.648)	
Large (K) Firm × FloodIndex		1.080
		(0.777)
Observations	330,580	330,580
Adj R-squared	0.296	0.296
Dep. var mean	68.349	68.349
Firm FE	Y	Y
Province × year FE	Y	Y
2-digit ISIC × year FE	Y	Y
Plant-level controls	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

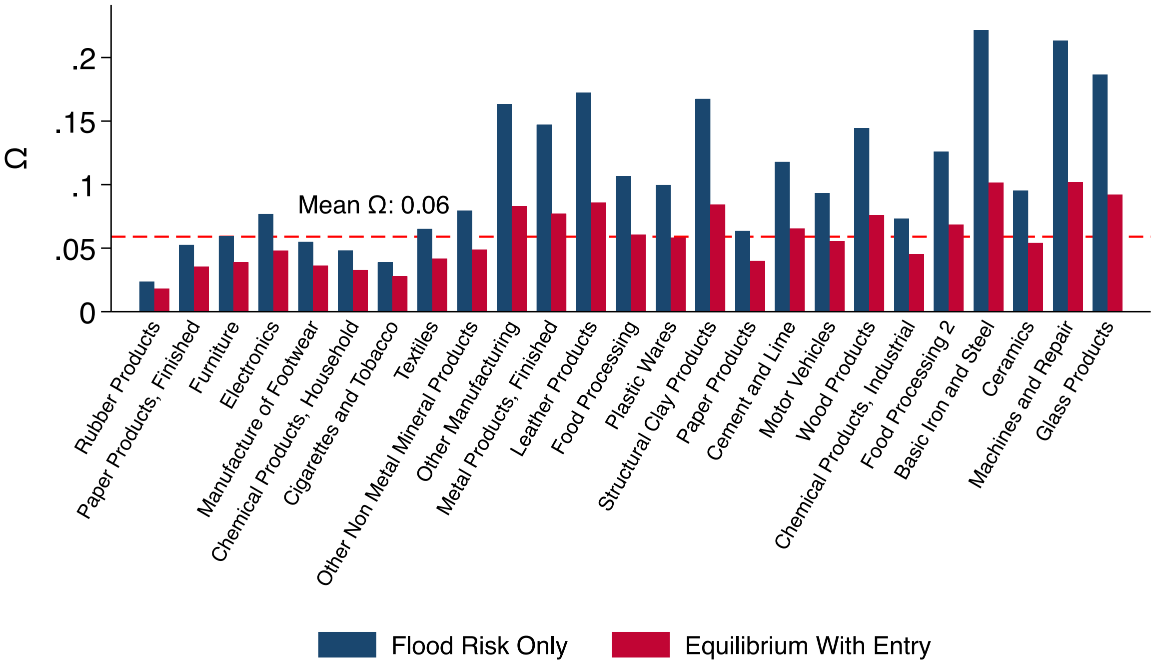
Notes: The table presents the results of estimating an interaction version of Equation (III.2) for firm-level production capacity utilization (PCU), where the flood index is interacted with the firm size dummy. The firm size dummy in Column 1 (2) uses average (over years) labor employment (capital stock) for each firm. PCU measures the percentage of the potential firm capacity, in terms of production, that is realized in a given year. Results reported in both the columns control for firm age controls, firm, province × year, and sector × year fixed effects. Standard errors are clustered at the regency level.

Figure B.15: **Change in aggregate output across sectors due to flood risk**



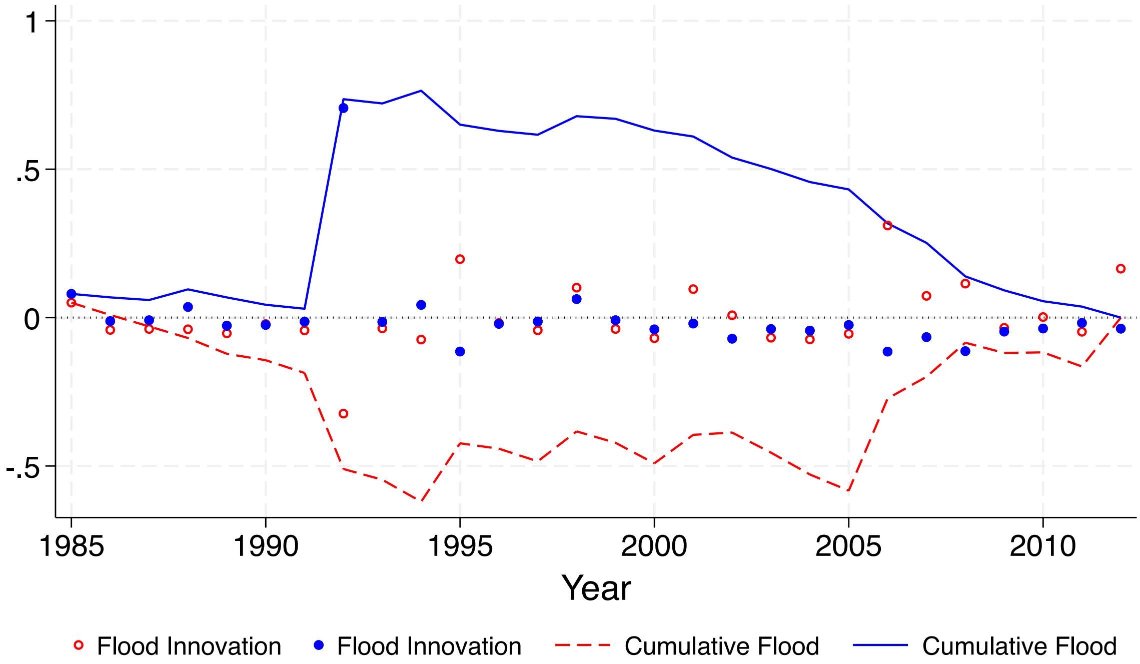
Notes: The graph plots the (log) change in aggregate output due to flood risk as outlined in Equation (VI.3) across 3-digit ISIC sectors, keeping only six regencies in which all 25 3-digit ISIC sectors are situated. This represents the Flood Risk Only scenario where in the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The sectors are ranked from left to right in the increasing order of their respective capital intensities.

Figure B.16: **Change in output across sectors due to flood risk and equilibrium**



Notes: The graph plots the (log) change in aggregate output due to flood risk in blue and sum of (log) change in aggregate output due to flood risk and (log) change in aggregate output due to equilibrium wage adjustments and firm entry in red as outlined in Equation (VI.3) across 3-digit ISIC sectors, keeping only six regencies in which all 25 3-digit ISIC sectors are situated. In the counterfactual world with flood defenses, all the regencies above 80th percentile on the flood exposure distribution are assigned the median value of the distribution. The sectors are ranked from left to right in the increasing order of their respective capital intensities.

Figure B.17: **Flood innovations and cumulative flood shocks**



Notes: The graph presents the evolution of cumulative flood shock variable over the years for two sample regencies in Indonesia. Circles represent flood innovations, which are generated as residuals from estimating Equation (A.1) for all the regencies in the period 1985-2012. The lines show the running sum of these flood innovations for each regency over time. Red (hollow circle and dashed line) represents the Bogor regency, which experienced runs of low flooding during this period. One the other hand, Yogyakarta city, represented in blue (solid circle and solid line) experienced runs of high flooding.

Table B.6: **Summary statistics on cumulative flood variables**

(1) Mean	(2) Std. Dev.	(3) 50th Pctile	(4) 90th Pctile	(5) 95th Pctile
Panel 1: Regency-level Cumulative Flood Shocks				
0.009	0.36	-0.027	0.348	0.69
Panel 2: Firm-level Cumulative Flood Shocks				
0.07	0.092	0.041	0.197	0.252

Notes: The table presents the summary statistics on cumultive flood variables used in the reduced-form analysis of long-run effects of flooding. Panel 1 and Panel 2 report statistics on cumulative flood variables at the regency and plant level respectively.

Table B.7: **Long-run effect of flooding on aggregate variables**

	(1) D.ln(VA)	(2) D.ln(K)	(3) D.ln(L)	(4) D.ln(VA)	(5) D.ln(K)	(6) D.ln(L)
D.CumulativeFlood	-0.563**	-0.703***	-0.803***	-0.544***	-0.600***	-0.464***
	(0.229)	(0.253)	(0.172)	(0.146)	(0.166)	(0.123)
Observations	246	246	246	1,320	1,320	1,320
R-squared	0.012	0.017	0.042	0.062	0.058	0.073
2-digit ISIC FE	-	-	-	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (A.2) using the first-difference estimator for aggregate variables i.e., logarithms of total value-added, capital stock, and labor employment at the regency or sector-regency level for the years 1994 and 2008. To get to the aggregate variables from firm-level information, following steps are undertaken. First, The un-logged version of all the monetary variables are deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values. Second, the tails on both ends of the resulting variables are trimmed by 1% for each year to address measurement error issues. Third, the variables are then summed across regency or sector-regency for each year. Finally, the variables are log-transformed and used in the regressions. Columns 1, 2 and 3 (4, 5, and 6) show the results for value-added, capital stock, and labor employment respectively when the firm data is collapsed at the regency (sector-regency) level. Standard errors are clustered at the regency level.

Table B.8: **Long-run effect of flooding on firm exit and entry**

	(1) Pr(exit)	(2) D.ln(#Plants)	(3) D.ln(#Plants)
CumulativeFlood	0.013*		
	(0.007)		
D.CumulativeFlood		-0.189***	-0.245***
		(0.070)	(0.085)
Observations	166,176	1,474	254
Adj R-squared	0.076	0.049	0.011
Regency FE	Y	-	-
2-digit ISIC FE	-	Y	-
Province $\times$ year FE	Y	-	-
2-digit ISIC $\times$ year FE	Y	-	-

Standard errors clustered at the regency level are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Column 1 presents the results of estimating Equation (A.3) for firm exit dummy where the dummy variable takes a value of 1 in the last year of the firm observation in the data. All the firms that start and end their operations in the period 1994-2008 are included in the estimation. Columns 2 and 3 present the results of estimating Equation (A.2) with the difference in the logarithm of number of firms operating in a regency or sector-regency for years 1994 and 2008 as the dependent variable. Column 1 controls for regency, province $\times$ year, and sector $\times$ year fixed effects. Column 2 controls for industry fixed effects. Standard errors are clustered at the regency level.

C Corresponding Tables for Figures in the Main Paper

Table C.1: **Effect of flooding on aggregate variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
25th Pctile Flood	-0.039	-0.072***	-0.028	-0.005
	(0.027)	(0.028)	(0.019)	(0.026)
Observations	34,359	34,359	34,359	12,225
Adj R-squared	0.510	0.489	0.509	0.278
Dep. var mean	10.802	10.845	5.988	0.938
Regency FE	Y	Y	Y	Y
Year FE	-	-	-	-
2-digit ISIC × year FE	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
50th Pctile Flood	-0.045	-0.098***	-0.039*	-0.002
	(0.032)	(0.032)	(0.023)	(0.028)
Observations	31,085	31,085	31,085	11,292
Adj R-squared	0.507	0.488	0.507	0.277
Dep. var mean	10.770	10.823	5.965	0.938
Regency FE	Y	Y	Y	Y
Year FE	-	-	-	-
2-digit ISIC × year FE	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
75th Pctile Flood	-0.083**	-0.117***	-0.051*	0.021
	(0.040)	(0.042)	(0.029)	(0.041)
Observations	27,598	27,598	27,598	9,789
Adj R-squared	0.504	0.485	0.506	0.276
Dep. var mean	10.750	10.804	5.942	0.931
Regency FE	Y	Y	Y	Y
Year FE	-	-	-	-
2-digit ISIC × year FE	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
90th Pctile Flood	-0.196**	-0.254***	-0.150***	0.007
	(0.077)	(0.075)	(0.055)	(0.067)
Observations	25,282	25,282	25,282	8,796
Adj R-squared	0.492	0.476	0.494	0.276
Dep. var mean	10.669	10.729	5.882	0.938
Regency FE	Y	Y	Y	Y
Year FE	-	-	-	-
2-digit ISIC × year FE	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.1) for aggregate variables i.e., logarithms of total value-added, capital stock, and labor employment at the regency or sector-regency level. To get to the aggregate variables from firm-level information, following steps are undertaken. First, The un-logged version of all the monetary variables are deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values. Second, the tails on both ends of the resulting variables are trimmed by 1% for each year to address measurement error issues. Third, the variables are then summed across regency or sector-regency for each year using labor share weights. Finally,

Table C.2: **Effect of flooding on firm-level variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]	(5) ln(VA)	(6) ln(K)	(7) ln(L)	(8) ln(L) [Temp]
25th Pctile Flood	0.010	-0.016	0.002	0.027***	0.007	-0.016	-0.001	0.027***
	(0.014)	(0.011)	(0.004)	(0.008)	(0.014)	(0.011)	(0.004)	(0.008)
Observations	298,730	298,730	298,730	298,695	298,730	298,730	298,730	298,695
Adj R-squared	0.848	0.842	0.906	0.660	0.848	0.842	0.907	0.660
Dep. var mean	8.585	8.642	4.122	0.239	8.585	8.642	4.122	0.239
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y	Y	Y	Y	Y
Plant-level controls	-	-	-	-	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]	(5) ln(VA)	(6) ln(K)	(7) ln(L)	(8) ln(L) [Temp]
50th Pctile Flood	-0.012	-0.028*	0.002	0.038***	-0.014	-0.028*	-0.000	0.037***
	(0.015)	(0.016)	(0.005)	(0.013)	(0.014)	(0.016)	(0.004)	(0.012)
Observations	265,704	265,704	265,704	265,673	265,704	265,704	265,704	265,673
Adj R-squared	0.848	0.843	0.906	0.656	0.849	0.843	0.907	0.656
Dep. var mean	8.574	8.646	4.123	0.244	8.574	8.646	4.123	0.244
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y	Y	Y	Y	Y
Plant-level controls	-	-	-	-	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]	(5) ln(VA)	(6) ln(K)	(7) ln(L)	(8) ln(L) [Temp]
75th Pctile Flood	-0.014	-0.058***	0.007	0.060***	-0.020	-0.058***	0.001	0.059***
	(0.022)	(0.022)	(0.008)	(0.020)	(0.021)	(0.022)	(0.007)	(0.020)
Observations	230,579	230,579	230,579	230,550	230,579	230,579	230,579	230,550
Adj R-squared	0.848	0.842	0.906	0.650	0.849	0.842	0.908	0.650
Dep. var mean	8.576	8.650	4.126	0.240	8.576	8.650	4.126	0.240
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y	Y	Y	Y	Y
Plant-level controls	-	-	-	-	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]	(5) ln(VA)	(6) ln(K)	(7) ln(L)	(8) ln(L) [Temp]
90th Pctile Flood	-0.041	-0.057*	0.005	0.055**	-0.049*	-0.057*	-0.004	0.053**
	(0.028)	(0.032)	(0.013)	(0.022)	(0.027)	(0.032)	(0.012)	(0.021)
Observations	201,592	201,592	201,592	201,566	201,592	201,592	201,592	201,566
Adj R-squared	0.848	0.843	0.904	0.656	0.849	0.843	0.905	0.656
Dep. var mean	8.505	8.590	4.096	0.241	8.505	8.590	4.096	0.241
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y	Y	Y	Y	Y
Plant-level controls	-	-	-	-	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.2) for firm-level variables i.e., logarithms of value-added, capital stock, permanent labor employment, and temporary labor employment. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Columns 1, 2, 3, and 4 (5, 6, 7, and 8) show the results for value-added, capital stock, temporary labor employment, and permanent labor employment respectively without (with) firm age controls. Results reported in all eight columns control for firm, province × year, and sector × year fixed effects. Standard errors are clustered at the regency level.

Table C.3: **Effect of flooding on firm-level capital categories**

	(1) ln(Structure)	(2) ln(Land)	(3) ln(Vehicle)	(4) ln(Equipment)
25th Pctile Flood	-0.020*	-0.013	0.004	0.008
	(0.011)	(0.012)	(0.010)	(0.011)
Observations	273,035	261,934	243,438	275,541
Adj R-squared	0.838	0.787	0.721	0.869
Dep. var mean	7.271	7.223	6.500	7.070
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(Structure)	(2) ln(Land)	(3) ln(Vehicle)	(4) ln(Equipment)
50th Pctile Flood	-0.034**	-0.023	0.000	0.012
	(0.017)	(0.017)	(0.015)	(0.016)
Observations	242,913	232,802	216,001	244,572
Adj R-squared	0.839	0.786	0.719	0.868
Dep. var mean	7.274	7.221	6.509	7.082
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(Structure)	(2) ln(Land)	(3) ln(Vehicle)	(4) ln(Equipment)
75th Pctile Flood	-0.053**	-0.076***	-0.020	-0.002
	(0.022)	(0.022)	(0.022)	(0.023)
Observations	210,945	202,016	187,908	212,278
Adj R-squared	0.838	0.784	0.717	0.867
Dep. var mean	7.277	7.217	6.511	7.088
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) ln(Structure)	(2) ln(Land)	(3) ln(Vehicle)	(4) ln(Equipment)
90th Pctile Flood	-0.083**	-0.118***	0.037	0.027
	(0.034)	(0.037)	(0.033)	(0.036)
Observations	184,845	177,084	163,454	184,864
Adj R-squared	0.839	0.782	0.716	0.867
Dep. var mean	7.222	7.163	6.486	7.010
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
2-digit ISIC × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.2) for four different capital categories at the firm level. Column 1 reports the results for value of structures, which include buildings and all man-made constructions to support the manufacturing activities within the firm. Column 2 shows results on land, which is the total value of land occupied by the manufacturing firm. Column 3 reports results on the value of vehicles and other transportation equipment owned by the firm. The last column shows results for value of machinery and other production equipment employed in the firm. As mentioned in the data section, the reporting on different capital categories is not consistent over time, and that is why the number of observations are different across all four columns. Results reported in all the four columns control for firm, province × year, and sector × year fixed effects along with firm age controls. Standard errors are clustered at the regency level.

Table C.4: **Effect of 25th percentile flood on firm level variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
Food Processing × 25th Pctile Flood	0.019	-0.031**	-0.001	0.039***
	(0.019)	(0.015)	(0.005)	(0.011)
Alcoholic Beverages × 25th Pctile Flood	-0.039	0.011	-0.007	0.006
	(0.039)	(0.035)	(0.012)	(0.015)
Tobacco × 25th Pctile Flood	-0.040	0.027	-0.030**	0.042**
	(0.047)	(0.033)	(0.013)	(0.017)
Textiles × 25th Pctile Flood	0.045**	-0.008	0.003	0.032***
	(0.021)	(0.017)	(0.005)	(0.010)
Leather Products × 25th Pctile Flood	0.046	0.073**	0.022	0.011
	(0.043)	(0.033)	(0.014)	(0.015)
Footwears × 25th Pctile Flood	0.064**	0.019	-0.011	0.030***
	(0.032)	(0.032)	(0.017)	(0.011)
Furniture and Wood Products × 25th Pctile Flood	-0.002	-0.043**	-0.008	0.021*
	(0.022)	(0.020)	(0.008)	(0.012)
Paper Products × 25th Pctile Flood	-0.021	-0.001	-0.004	0.024*
	(0.030)	(0.023)	(0.010)	(0.013)
Chemical Products × 25th Pctile Flood	-0.009	-0.026	0.006	0.036***
	(0.024)	(0.017)	(0.011)	(0.013)
Rubber and Plastic Products × 25th Pctile Flood	-0.008	-0.060**	0.009	0.034***
	(0.022)	(0.024)	(0.011)	(0.012)
Ceramics, Glass and Clay Products × 25th Pctile Flood	0.015	0.032*	0.002	0.065***
	(0.030)	(0.019)	(0.010)	(0.019)
Cement and Lime × 25th Pctile Flood	0.021	-0.027	-0.005	0.044***
	(0.025)	(0.025)	(0.009)	(0.014)
Iron and Steel × 25th Pctile Flood	-0.026	-0.118***	0.024	0.071***
	(0.032)	(0.045)	(0.015)	(0.013)
Metal Products × 25th Pctile Flood	0.006	-0.005	-0.006	0.043***
	(0.024)	(0.027)	(0.011)	(0.012)
Machines and Repair × 25th Pctile Flood	-0.027	-0.029	0.010	0.037***
	(0.030)	(0.025)	(0.010)	(0.011)
Electronics × 25th Pctile Flood	0.033	0.040	0.000	0.052***
	(0.026)	(0.034)	(0.013)	(0.015)
Observations	298,730	298,730	298,730	298,695
Adj R-squared	0.845	0.841	0.906	0.642
Dep. var mean	8.585	8.642	4.122	0.239
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added, capital stock, permanent labor employment, and temporary labor employment using 25th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Results reported in all the four columns control for firm and province × year fixed effects. Standard errors are clustered at the regency level.

Table C.5: **Effect of 50th percentile flood on firm level variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
Food Processing × 50th Pctile Flood	-0.002	-0.034*	0.001	0.058***
	(0.018)	(0.020)	(0.005)	(0.017)
Alcoholic Beverages × 50th Pctile Flood	-0.062	0.011	-0.003	0.026
	(0.047)	(0.040)	(0.017)	(0.020)
Tobacco × 50th Pctile Flood	-0.091*	0.002	-0.040***	0.051**
	(0.050)	(0.038)	(0.011)	(0.025)
Textiles × 50th Pctile Flood	0.017	-0.021	0.002	0.047**
	(0.025)	(0.021)	(0.006)	(0.019)
Leather Products × 50th Pctile Flood	0.051	0.067*	0.012	0.025
	(0.049)	(0.036)	(0.016)	(0.023)
Footwears × 50th Pctile Flood	0.062*	-0.005	-0.020	0.051**
	(0.035)	(0.029)	(0.016)	(0.021)
Furniture and Wood Products × 50th Pctile Flood	-0.016	-0.059**	-0.002	0.033
	(0.025)	(0.023)	(0.008)	(0.021)
Paper Products × 50th Pctile Flood	-0.071***	-0.015	-0.005	0.043**
	(0.022)	(0.028)	(0.010)	(0.020)
Chemical Products × 50th Pctile Flood	-0.041	-0.058***	0.004	0.052**
	(0.026)	(0.021)	(0.011)	(0.022)
Rubber and Plastic Products × 50th Pctile Flood	-0.016	-0.063**	0.011	0.049**
	(0.024)	(0.029)	(0.011)	(0.020)
Ceramics, Glass and Clay Products × 50th Pctile Flood	0.044**	0.066**	0.012	0.120**
	(0.021)	(0.029)	(0.018)	(0.050)
Cement and Lime × 50th Pctile Flood	0.014	-0.025	0.004	0.056***
	(0.029)	(0.028)	(0.011)	(0.019)
Iron and Steel × 50th Pctile Flood	-0.010	-0.106**	0.040**	0.090***
	(0.042)	(0.051)	(0.016)	(0.020)
Metal Products × 50th Pctile Flood	-0.002	0.005	-0.006	0.065***
	(0.025)	(0.030)	(0.012)	(0.021)
Machines and Repair × 50th Pctile Flood	-0.061**	-0.051*	0.019*	0.055***
	(0.026)	(0.027)	(0.010)	(0.019)
Electronics × 50th Pctile Flood	-0.010	0.027	-0.012	0.065***
	(0.032)	(0.033)	(0.012)	(0.022)
Observations	265,704	265,704	265,704	265,673
Adj R-squared	0.846	0.841	0.906	0.638
Dep. var mean	8.574	8.646	4.123	0.244
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added, capital stock, permanent labor employment, and temporary labor employment using 50th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Results reported in all the four columns control for firm and province × year fixed effects. Standard errors are clustered at the regency level.

Table C.6: **Effect of 75th percentile flood on firm level variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
Food Processing × 75th Pctile Flood	0.008	-0.067**	0.008	0.092***
	(0.026)	(0.027)	(0.009)	(0.030)
Alcoholic Beverages × 75th Pctile Flood	-0.058	-0.019	-0.002	0.059*
	(0.066)	(0.055)	(0.021)	(0.032)
Tobacco × 75th Pctile Flood	-0.162***	-0.035	-0.050***	0.039
	(0.062)	(0.046)	(0.016)	(0.040)
Textiles × 75th Pctile Flood	0.025	-0.065**	0.006	0.084***
	(0.036)	(0.027)	(0.009)	(0.031)
Leather Products × 75th Pctile Flood	0.060	-0.029	0.027	0.055
	(0.059)	(0.054)	(0.018)	(0.039)
Footwears × 75th Pctile Flood	0.057	-0.037	-0.009	0.077**
	(0.052)	(0.045)	(0.016)	(0.033)
Furniture and Wood Products × 75th Pctile Flood	-0.008	-0.102***	-0.004	0.062*
	(0.038)	(0.035)	(0.011)	(0.034)
Paper Products × 75th Pctile Flood	-0.057*	-0.048	-0.004	0.075**
	(0.030)	(0.039)	(0.012)	(0.031)
Chemical Products × 75th Pctile Flood	-0.052	-0.105***	0.004	0.079**
	(0.042)	(0.029)	(0.012)	(0.036)
Rubber and Plastic Products × 75th Pctile Flood	-0.004	-0.080**	0.015	0.081**
	(0.033)	(0.033)	(0.012)	(0.032)
Ceramics, Glass and Clay Products × 75th Pctile Flood	0.036	0.021	0.009	0.187**
	(0.034)	(0.041)	(0.018)	(0.081)
Cement and Lime × 75th Pctile Flood	0.038	-0.074**	-0.001	0.098***
	(0.043)	(0.036)	(0.014)	(0.029)
Iron and Steel × 75th Pctile Flood	-0.022	-0.206***	0.048***	0.123***
	(0.059)	(0.058)	(0.013)	(0.037)
Metal Products × 75th Pctile Flood	-0.010	-0.061*	-0.009	0.088***
	(0.041)	(0.036)	(0.017)	(0.033)
Machines and Repair × 75th Pctile Flood	-0.047	-0.113***	0.040***	0.092***
	(0.033)	(0.036)	(0.015)	(0.032)
Electronics × 75th Pctile Flood	-0.012	0.011	0.025	0.101***
	(0.044)	(0.037)	(0.018)	(0.036)
Observations	230,579	230,579	230,579	230,550
Adj R-squared	0.846	0.840	0.907	0.633
Dep. var mean	8.576	8.650	4.126	0.240
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added, capital stock, permanent labor employment, and temporary labor employment using 75th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Results reported in all the four columns control for firm and province × year fixed effects. Standard errors are clustered at the regency level.

Table C.7: **Effect of 90th percentile flood on firm level variables**

	(1) ln(VA)	(2) ln(K)	(3) ln(L)	(4) ln(L) [Temp]
Food Processing × 90th Pctile Flood	0.005	-0.070*	0.004	0.069**
	(0.037)	(0.038)	(0.014)	(0.034)
Alcoholic Beverages × 90th Pctile Flood	-0.068	-0.089	-0.009	0.084*
	(0.138)	(0.117)	(0.042)	(0.047)
Tobacco × 90th Pctile Flood	-0.222**	-0.062	-0.047**	0.039
	(0.092)	(0.050)	(0.021)	(0.039)
Textiles × 90th Pctile Flood	0.007	-0.075**	0.012	0.068*
	(0.036)	(0.036)	(0.016)	(0.040)
Leather Products × 90th Pctile Flood	0.028	-0.102	0.044	0.055
	(0.098)	(0.096)	(0.040)	(0.049)
Footwears × 90th Pctile Flood	0.022	-0.022	-0.067*	0.097**
	(0.065)	(0.091)	(0.038)	(0.045)
Furniture and Wood Products × 90th Pctile Flood	-0.044	-0.115*	-0.024	0.093**
	(0.046)	(0.064)	(0.019)	(0.039)
Paper Products × 90th Pctile Flood	-0.049	-0.046	-0.028	0.109**
	(0.039)	(0.070)	(0.023)	(0.043)
Chemical Products × 90th Pctile Flood	-0.069	-0.070	0.014	0.107**
	(0.063)	(0.047)	(0.025)	(0.050)
Rubber and Plastic Products × 90th Pctile Flood	0.076	0.065	0.015	0.109**
	(0.046)	(0.061)	(0.019)	(0.044)
Ceramics, Glass and Clay Products × 90th Pctile Flood	0.059	-0.112**	-0.008	0.122**
	(0.064)	(0.053)	(0.032)	(0.052)
Cement and Lime × 90th Pctile Flood	0.076	-0.052	0.055***	0.123***
	(0.054)	(0.052)	(0.021)	(0.037)
Iron and Steel × 90th Pctile Flood	-0.089	-0.241**	0.070**	0.152**
	(0.121)	(0.113)	(0.032)	(0.070)
Metal Products × 90th Pctile Flood	-0.106**	-0.064	-0.020	0.112**
	(0.052)	(0.073)	(0.027)	(0.044)
Machines and Repair × 90th Pctile Flood	-0.034	-0.004	0.050*	0.150***
	(0.051)	(0.068)	(0.030)	(0.043)
Electronics × 90th Pctile Flood	-0.137	0.104	0.017	0.197***
	(0.091)	(0.071)	(0.028)	(0.050)
Observations	201,592	201,592	201,592	201,566
Adj R-squared	0.845	0.841	0.904	0.639
Dep. var mean	8.505	8.590	4.096	0.241
Firm FE	Y	Y	Y	Y
Province × year FE	Y	Y	Y	Y
Plant-level controls	Y	Y	Y	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.3) for firm-level variables i.e., logarithms of value-added, capital stock, permanent labor employment, and temporary labor employment using 90th percentile flood dummy. The un-logged version of all the monetary variables have been deflated by the wholesale price index at the 5-digit ISIC level to reflect their real values and the log-transformed variables are trimmed by 1% for each year to address measurement error issues. Results reported in all the four columns control for firm and province × year fixed effects. Standard errors are clustered at the regency level.

Table C.8: **Effect of flooding on firm exit and entry**

	(1) Pr(exit)	(2) ln(#Entrants)
25th Pctile Flood	0.005	-0.060**
	(0.006)	(0.026)
Observations	443,491	13,036
Adj R-squared	0.379	0.310
Dep. var mean	0.119	0.736
Regency FE	Y	Y
Year FE	-	-
Province × year FE	Y	-
2-digit ISIC × year FE	Y	Y
Plant-level controls	Y	-

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Pr(exit)	(2) ln(#Entrants)
50th Pctile Flood	0.004	-0.072**
	(0.007)	(0.029)
Observations	395,082	11,992
Adj R-squared	0.352	0.316
Dep. var mean	0.117	0.741
Regency FE	Y	Y
Year FE	-	-
Province × year FE	Y	-
2-digit ISIC × year FE	Y	Y
Plant-level controls	Y	-

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Pr(exit)	(2) ln(#Entrants)
75th Pctile Flood	0.003	-0.041
	(0.010)	(0.034)
Observations	341,844	10,454
Adj R-squared	0.259	0.317
Dep. var mean	0.104	0.737
Regency FE	Y	Y
Year FE	-	-
Province × year FE	Y	-
2-digit ISIC × year FE	Y	Y
Plant-level controls	Y	-

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Pr(exit)	(2) ln(#Entrants)
90th Pctile Flood	-0.001	-0.204***
	(0.015)	(0.049)
Observations	294,315	9,224
Adj R-squared	0.264	0.308
Dep. var mean	0.107	0.699
Regency FE	Y	Y
Year FE	-	-
Province × year FE	Y	-
2-digit ISIC × year FE	Y	Y
Plant-level controls	Y	-

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Column 1 presents the results of estimating Equation (III.4) for firm exit dummy, where the dummy variable takes a value of 1 in the last year of firm observation in the data. Columns 2 presents the results of estimating Equation (III.1) with the logarithm of number of firms operating in a sector-regency in each year as the dependent variable. Column 1 controls for regency, province × year, and sector × year fixed effects. Column 2 controls for regency and sector × year fixed effects. Standard errors are clustered at the regency level.

Table C.9: **Effect of flooding on firm capacity utilization**

	(1) % PCU	(2) % PCU
25th Pctile Flood	-0.952**	-0.964**
	(0.425)	(0.427)
Observations	298,516	298,516
Adj R-squared	0.294	0.295
Dep. var mean	68.419	68.419
Firm FE	Y	Y
Province × year FE	Y	Y
2-digit ISIC × year FE	Y	Y
Plant-level controls	-	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) % PCU	(2) % PCU
50th Pctile Flood	-1.729***	-1.739***
	(0.473)	(0.474)
Observations	265,506	265,506
Adj R-squared	0.293	0.293
Dep. var mean	68.415	68.415
Firm FE	Y	Y
Province × year FE	Y	Y
2-digit ISIC × year FE	Y	Y
Plant-level controls	-	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) % PCU	(2) % PCU
75th Pctile Flood	-3.236***	-3.249***
	(0.782)	(0.786)
Observations	230,429	230,429
Adj R-squared	0.292	0.293
Dep. var mean	68.570	68.570
Firm FE	Y	Y
Province × year FE	Y	Y
2-digit ISIC × year FE	Y	Y
Plant-level controls	-	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) % PCU	(2) % PCU
90th Pctile Flood	-3.559***	-3.576***
	(1.066)	(1.073)
Observations	201,464	201,464
Adj R-squared	0.289	0.289
Dep. var mean	68.836	68.836
Firm FE	Y	Y
Province × year FE	Y	Y
2-digit ISIC × year FE	Y	Y
Plant-level controls	-	Y

Standard errors clustered at the regency level are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table presents the results of estimating Equation (III.2) for firm-level production capacity utilization (PCU). PCU measures the percentage of the potential firm capacity, in terms of production, that is realized in a given year. Columns 1 (2) show the results without (with) firm age controls. Results reported in both the columns control for firm, province × year, and sector × year fixed effects. Standard errors are clustered at the regency level.

D Detailed Proofs of the Theory

D.1 Flood Risk (τ_{srt})

The distribution of the share variable x is as follows:

$$G_{rt}(x) = \begin{cases} 1 - \left(\frac{1}{x}\right)^{\phi_{rt}} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

Firms maximize expected profits by choosing the optimal capital to install in a given period taking expectations on the random variable x_{it}. Firm's optimization problem is as below:

$$K_{it} = \text{argmax}\left\{\Gamma_{it}\mathbb{E}\left[\left(\frac{K_{it}}{x_{it}}\right)^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\right] - \rho K_{it}\right\}$$

where $\Gamma_{it}(\theta, w) \equiv [1-(1-\alpha_s)\eta_s]\,\theta_i^{\frac{1}{1-(1-\alpha_s)\eta_s}}\left\{\frac{w_t}{(1-\alpha_s)\eta_s}\right\}^{-\frac{(1-\alpha_s)\eta_s}{1-(1-\alpha_s)\eta_s}}$.

The above problem can be written as:

$$K_{it} = \text{argmax}\left\{\Gamma_{it}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\int_1^\infty x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}g(x_{it})dx_{it} - \rho K_{it}\right\}$$

Putting the p.d.f of share distribution follows:

$$\begin{aligned} K_{it} &= \text{argmax}\left\{\Gamma_{it}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\int_1^\infty x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\phi_{rt}x_{it}^{-\phi_{rt}-1}dx_{it} - \rho K_{it}\right\} \\ &= \text{argmax}\left\{\Gamma_{it}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\phi_{rt}\int_1^\infty x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}-\phi_{rt}-1}dx_{it} - \rho K_{it}\right\} \\ &= \text{argmax}\left\{\frac{\phi_{rt}}{\phi_{rt}+\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\Gamma_{it}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} - \rho K_{it}\right\} \\ &= \text{argmax}\left\{\tau_{srt}\Gamma_{it}K_{it}^{\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} - \rho K_{it}\right\} \end{aligned}$$

D.2 Expected Aggregate Equilibrium Output $(\overline{Y}_{srt})$

The general expression for aggregate output is as follows:

$$\overline{Y}_{srt} = \int_{\theta^*_{srt}}^\infty Y_{it}(\theta)\mu_{srt}(\theta)d\theta$$

Putting the expression for equilibrium firm-level output from Equation (IV.8):

$$\overline{Y}_{srt} = \Lambda_{st}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \int_{\theta^*_{srt}}^{\infty} \theta^{\frac{1}{1-\eta_s}} \mu_{srt}(\theta) d\theta$$

Adding the equilibrium productivity distribution from Equation (IV.11):

$$\overline{Y}_{srt} = \Lambda_{st}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \frac{\xi\overline{\theta}_r^{\xi}}{1-H_r(\theta^*_{srt})} \int_{\theta^*_{srt}}^{\infty} \theta^{\frac{1}{1-\eta_s}-\xi-1} d\theta$$

Integrating the above, follows:

$$\overline{Y}_{srt} = \Lambda_{st}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \frac{\xi\overline{\theta}_r^{\xi}(1-\eta_s)}{\xi(1-\eta_s)-1} \frac{(\theta^*_{srt})^{\frac{1}{1-\eta_s}-\xi}}{1-H_r(\theta^*_{srt})}$$

Using the initial productivity distribution to compute $(1 - H_r(\theta^*_{srt}))$ and assuming $(\xi(1-\eta_s) > 1)$, the formula for aggregate output in Equation (IV.12) is obtained as below:

$$\overline{Y}_{srt} = \Lambda_{st}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1} (\theta^*_{srt})^{\frac{1}{1-\eta_s}}$$

D.3 Labor Market Clearing (w_t)

The labor market clearing takes place at the regency level with the total labor employed by all the firms equal to the aggregate (exogenous, time-invariant) labor supply in the regency as follows:

$$\overline{L}_t = \int_{i \in t} L di$$

Using Equation (IV.7), LHS can be expanded as follows:

$$\overline{L}_t = \sum_{r=1}^{R}\sum_{s=1}^{S} \int \frac{(1-\alpha_s)\eta_s}{w_t} \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st}\theta_i^{\frac{1}{1-\eta_s}} x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} f(x,\theta) dx d\theta$$

Using the independence of stochastic processes for θ and x, $f(x,\theta)$ can written as the product of respective densities as follows:

$$\overline{L} = \sum_{r=1}^{R}\sum_{s=1}^{S} \int_{\theta^*_{srt}}^{\infty} \int_{1}^{\infty} \frac{(1-\alpha_s)\eta_s}{w_t} \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st}\theta_i^{\frac{1}{1-\eta_s}} x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} g(x)\mu_{srt}(\theta) dx d\theta$$

$$= \sum_{r=1}^{R}\sum_{s=1}^{S} \frac{(1-\alpha_s)\eta_s}{w_t} \tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}} \Lambda_{st} \left(\int_{\theta^*_{srt}}^{\infty} \theta_i^{\frac{1}{1-\eta_s}} \mu_{srt}(\theta) d\theta\right) \left(\int_{1}^{\infty} x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}} g(x) dx\right)$$

Using the definitions of the respective distributions, the above can be written as follows:

$$\overline{L} = \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}\left(\frac{\xi\overline{\theta}_r^{\xi}}{1-H_r(\theta^*_{srt})}\int_{\theta^*_{srt}}^{\infty}\theta_i^{\frac{1}{1-\eta_s}-\xi-1}d\theta\right)\left(\int_1^{\infty}x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\phi_{rt}x_{it}^{-\phi_{rt}-1}dx\right)$$

$$= \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}\left(\frac{\xi\overline{\theta}_r^{\xi}}{1-H_r(\theta^*_{srt})}\int_{\theta^*_{srt}}^{\infty}\theta_i^{\frac{1}{1-\eta_s}-\xi-1}d\theta\right)\left(\phi_{rt}\int_1^{\infty}x_{it}^{-\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}-\phi_{rt}-1}dx\right)$$

$$= \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}\left(\frac{\xi\overline{\theta}_r^{\xi}(1-\eta_s)}{\xi(1-\eta_s)-1}\frac{(\theta^*_{srt})^{\frac{1}{1-\eta_s}-\xi}}{1-H_r(\theta^*_{srt})}\right)\left(\frac{\phi_{rt}}{\phi_{rt}+\frac{\alpha_s\eta_s}{1-(1-\alpha_s)\eta_s}}\right)$$

$$= \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}^{\frac{1-\eta+\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}(\theta^*_{srt})^{\frac{1}{1-\eta_s}}$$

Expanding the θ^*_{srt} using Equation (IV.13) follows:

$$\overline{L}_t = \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}^{\frac{1-\eta+\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}\left\{\frac{f}{\left[1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}\right]\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}}\Lambda_{st}}\right\}$$

$$= \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s}{w_t}\tau_{srt}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}\left\{\frac{f}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}}\right\}$$

The above delivers the equilibrium wage equation as follows:

$$w_t = \frac{f}{\overline{L}}\sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s\tau_{srt}}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}$$

The expression for cutoff productivity derived in Equation (IV.13) combined with the definition of Λ_{st} gives:

$$\theta^*_{srt} = \left\{\frac{f}{\left[1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}\right]\tau_{srt}^{\frac{\alpha_s\eta_s}{1-\eta_s}}\left\{\frac{w_t}{(1-\alpha_s)\eta_s}\right\}^{-\frac{(1-\alpha_s)\eta_s}{1-\eta_s}}\left\{\frac{\rho}{\alpha_s\eta_s}\right\}^{-\frac{\alpha_s\eta_s}{1-\eta_s}}}\right\}^{1-\eta_s}$$

$$= \frac{f^{1-\eta_s}\rho^{\alpha_s\eta_s}w_t^{(1-\alpha_s)\eta_s}}{\left(1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}\right)^{1-\eta_s}\tau_{srt}^{\alpha_s\eta_s}\left((1-\alpha_s)\eta_s\right)^{(1-\alpha_s)\eta_s}\left(\alpha_s\eta_s\right)^{\alpha_s\eta_s}}$$

Endogenizing wages using the equilibrium expression derived above:

$$\theta^*_{srt} = \frac{f^{1-\eta_s}\rho^{\alpha_s\eta_s}}{(1-(1-\alpha_s)\eta_s - \alpha_s\eta_s\tau_{srt})^{1-\eta_s}\,\tau_{srt}^{\alpha_s\eta_s}\,((1-\alpha_s)\eta_s)^{(1-\alpha_s)\eta_s}\,(\alpha_s\eta_s)^{\alpha_s\eta_s}}$$
$$\times\left\{\frac{f}{\overline{\overline{L}}}\sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s\tau_{srt}}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}\right\}^{(1-\alpha_s)\eta_s}$$

Simplifying further delivers the equilibrium cutoff productivity expression:

$$\theta^*_{srt} = \frac{f^{1-\alpha_s\eta_s}}{\left((1-\alpha_s)\eta_s\overline{L}\right)^{(1-\alpha_s)\eta_s}}\left(\frac{\rho}{\alpha_s\eta_s\tau_{srt}}\right)^{\alpha_s\eta_s}$$
$$\times\left\{\frac{1}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}}\right\}^{1-\eta_s}$$
$$\times\left\{\sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(1-\alpha_s)\eta_s\tau_{srt}}{1-(1-\alpha_s)\eta_s-\alpha_s\eta_s\tau_{srt}}\frac{\xi(1-\eta_s)}{\xi(1-\eta_s)-1}\right\}^{(1-\alpha_s)\eta_s}$$

D.4 MLE Estimator for Regency Flood Exposure (ϕ_{rt})

The distribution of the share variable x is as follows:

$$G_{rt}(x) = \begin{cases} 1-\left(\frac{1}{x}\right)^{\phi_{rt}} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

This gives the p.d.f.

$$g_{rt}(x) = \frac{\phi_{rt}}{x^{\phi_{rt}+1}}$$

The likelihood function for a sample of firms in period t $(x_{1t}, x_{2t}, x_{3t}, \cdots, x_{Nt})$ located in regency r can be written as:

$$L(\phi) = \prod_{i=1}^{N_{rt}} \frac{\phi_{rt}}{x_{it}^{\phi_{rt}+1}}$$

The log-likelihood function becomes:

$$\ln(L(\phi)) = N_{rt}\ln(\phi_{rt}) - (\phi_{rt}+1)\sum_{i=1}^{N_{rt}}\ln(x_{it})$$

To find the MLE for ϕ_{rt}, take the derivative of the log-likelihood with respect to ϕ_{rt}:

$$\frac{d\ln(L(\phi))}{d\phi} = \frac{N_{rt}}{\phi_{rt}} - \sum_{i=1}^{N_{rt}} \ln(x_{it})$$

Setting the derivative to zero delivers the estimator:

$$\widehat{\phi}_{rt} = \frac{N_{rt}}{\sum_{i=1}^{N_{rt}} \ln(x_{it})}$$