Firms Under Water: Floods, Adaptation and Performance [*]

Roberta Muri

University of Bologna and Prometeia Think Tank

roberta.muri2@unibo.it

This version: October 23, 2026

Click here for the latest version

Abstract

When a flood arrives, how much of the damage depends on protective infrastructure built before it? This paper estimates the causal effect of ex-ante adaptation on firm revenues. I build a novel dataset linking administrative records on public investment projects (2003–2024) to river and canal geometries, geocoded firms and balance-sheet outcomes, high-frequency rainfall, and flood-extent maps. I develop an original LLM-based classification pipeline to identify hydrological adaptation projects, distinguish ex-ante (preventive) from ex-post (reactive) interventions, and construct firm-level exposure to protection. I apply the framework to Emilia–Romagna, an Italian region that experienced a severe flood in May 2023. To address endogenous targeting of adaptation towards high-risk areas, identification exploits systematic differences in completion speed across funding pipelines. Firms inside the flood perimeter lose about 4 percent of their revenue in the year of the event, relative to comparable firms outside it. The effect decays sharply with distance, and beyond 250 metres from the perimeter it is indistinguishable from zero. Completed ex-ante adaptation attenuates the losses associated with extreme rainfall: for every additional 10 mm of extreme rainfall, €1 million of works completed near a firm preserves about €104,000 of revenue, roughly 80 percent of the loss that the same rainfall causes in the absence of protection. No such protection is found for restorative works built to repair past damage rather than to reduce future risk, suggesting that prevention and repair are not substitutes.

JEL Codes: Q54, Q58, L25, D22, H54
Keywords: public adaptation, floods, firm performance

[*]I am deeply grateful to Lorenzo Forni and Vincenzo Scrutinio for invaluable guidance throughout this project. I thank Andrea Camilli, Hannah Druckenmiller, Alessandro Palma, and Guglielmo Zappalà for helpful comments and discussions. I also thank participants at the 67th Annual Meeting of the Italian Economic Association (SIE) in Cagliari (2026), the AECCE Workshop in Urbino (2026), LSE Environment Week (2026), the Italian Association of Environmental and Resource Economists (IAERE) Conference in Trento (2026), the Bologna Environmental and Experimental Group (2026), Mapping Environmental Vulnerabilities: Data, Space and Risk in Rome (2025), and the Verona Early Career Workshop in Economics (2025). I am grateful to Elisa Mahfouz, Emanuela Muri, Damiano Paoli, and Pietro Scala for their support. All errors are my own.

1 Introduction

Between 1980 and 2024, floods were the costliest weather- and climate-related hazard in the European Union, causing an estimated €386 billion in economic losses. Nearly 30 percent of these losses occurred between 2021 and 2024 alone (European Environment Agency, 2025). Global warming and continued development in flood-prone areas are expected to increase future losses. By 2100, annual river-flood damage in the EU and UK could reach six times current levels under 3°C of global warming and no further adaptation (Dottori et al., 2020).

To contain these losses, governments will have to commit more resources to adaptation. Public money can go to two uses: ex-ante adaptation, such as levees, retention basins and drainage built before the water arrives, and ex-post relief and reconstruction after it (Carleton et al., 2024). The two differ in how they are paid for and in what they return. Prevention is paid upfront to reduce expected damage. Relief is paid only if an event occurs, to recover from realised damage. Little is known about what prevention returns, particularly to economic activity.

The binding constraint is measurement at the firm level. To date, no dataset links adaptation projects to the natural features they are intended to protect and, in turn, to economic outcomes such as firm performance. In the absence of such a measure, protection can be evaluated only where it can be aggregated, across administrative units or through asset prices, neither of which recovers the losses a firm incurs, or avoids, when a flood occurs. What, then, are the economic benefits of public ex-ante adaptation for private firms? Does public flood protection reduce firms' revenue losses in a flood?

This paper addresses these questions using the May 2023 floods in Emilia-Romagna, a major manufacturing region in northern Italy in which three quarters of productive establishments sit in areas exposed to hydrogeological hazard (Loberto and Russo, 2024). Exceptionally heavy rainfall caused widespread river overflows, affecting residential areas, infrastructure, and economic activity: twenty-three watercourses were overtopped and 540 square kilometres of plain were inundated (Brath et al., 2023). The floods followed two decades of recorded public investment in the region's rivers and canals. During this period, public authorities planned and implemented numerous hydraulic interventions intended to reduce flood risk. This setting allows me to examine whether preventive public investments reduced firms' economic losses during a major flood.

I combine detailed records of public hydraulic investment with geocoded firm accounts, inundation maps, and high-resolution precipitation data. A central feature of the data is the ability to observe public adaptation at the level of individual investment projects. Rather than relying on aggregate expenditure, I recover what each project was intended to achieve, where it was located, and whether it had reached completion before the floods.

I construct the investment dataset using the OpenCUP registry, which contains administrative information on public investment projects. I employ a large language model to classify project descriptions, identifying hydraulic interventions and distinguishing preventive works from projects primarily intended to repair damage from previous events. I then match the identified projects to the regional river and canal network and recover their status before May 2023. Financial closure provides an administrative proxy for project completion. The classification is cost-effective, about €55 for the

600,823 records of the registry, and I validate it in three ways to assess its reliability, with a precision of about 90 percent.

I link these investment records to individual firms. For each firm, I construct measures of planned and completed preventive investment along watercourse segments within 250 metres. Project costs are expressed in constant prices and weighted by distance. This approach provides firm-specific measures of nearby public protection, distinguishing the scale of investment planned in a location from the portion completed before the flood.

The empirical analysis proceeds in two steps. First, I examine the economic consequences of the May 2023 floods using a difference-in-differences design. I compare changes in revenues between firms located inside and outside the inundation perimeters. Event-study estimates trace these differences before and after the disaster, while additional specifications examine how the impact varies with distance from the inundated areas.

Then, the main analysis examines whether preventive public investment attenuates the economic consequences of extreme rainfall. An important consideration is that realised inundation is itself potentially affected by flood protection. If an investment prevents a location from flooding, conditioning on observed inundation would exclude part of the benefit that the investment is intended to provide. I therefore use the intensity of the meteorological shock, measured by the maximum precipitation accumulated over three consecutive hours during May 2023, rather than realised inundation. I estimate whether the relationship between rainfall and post-flood revenues varies with nearby completed preventive investment.

Identifying the protective effect of public investment is challenging because infrastructure is not randomly allocated. Public authorities direct preventive works towards locations facing greater expected flood damage. Firms located near more investment may therefore face greater underlying flood risk, making a direct comparison between protected and unprotected firms potentially misleading. I address this concern by distinguishing between planned and completed investment. The analysis compares, before and after the flood, firms with similar planned investment nearby but different levels of completed adaptation by the time of the shock. The empirical specification includes interactions between rainfall and both planned and completed protection, together with firm and province-by-year fixed effects.

Completion may nevertheless be endogenous. Projects may advance at different speeds because of local administrative capacity, political priorities, or other factors correlated with firms' economic outcomes. To address this concern, I instrument completed protection using variation in completion rates across funding sources and project vintages. State and EU funding channels exhibit lower completion rates than regional funding channels, conditional on project size and location.

For each firm, I construct an index of predicted completed investment by weighting the completion rates of its project cohorts by nearby planned expenditure. These rates are computed using other projects, excluding those located in the firm's immediate surroundings. The resulting leave-one-out instrument predicts the completion of the planned investment portfolio. I use this index to instrument completed protection and its interaction with rainfall in the post-flood period. Identification requires

that differences in completion rates across funding sources and vintages affect firms' post-flood revenues only through completed protection, conditional on planned investment and the other controls. I run several placebo and additional tests to assess the credibility of this assumption.

I find five main results. First, the May 2023 floods caused substantial revenue losses for directly exposed firms. Firms located within the inundation perimeters experienced an approximately 4 percent decline in revenues in 2023 relative to firms outside these areas. The impact was spatially concentrated: estimated losses decline sharply with distance from the inundated area and, beyond 250 metres, are negligible and no longer statistically distinguishable from zero.

Second, completed preventive investment attenuates the negative relationship between extreme rainfall and firm revenues. Instrumental-variable estimates indicate that, for every additional 10 mm of extreme rainfall, €1 million of completed preventive works within 250 metres of a firm preserves approximately €104,000 of revenue, about 80 percent of the loss that the same rainfall causes to a firm without protection. At the average stock of completed works in the sample, €0.89 million, the share offset is about 71 percent. The estimated protective effect remains positive under alternative measures of hydraulic adaptation based on rivers and canals.

Third, the protective effects extend to other measures of firm performance. Completed preventive investment also attenuates the adverse effects of extreme rainfall on value added and operating profits. By contrast, the estimates for employment, labour costs, and balance-sheet outcomes do not establish similarly clear effects. These findings are consistent with public protection helping firms maintain economic activity during a flood, although annual financial accounts do not identify the precise physical mechanisms through which protection operates.

Fourth, the damage travels down the river network but the protection does not. Rainfall falling upstream of a firm is as harmful as rainfall falling on the firm itself, at €138,447 per 10 mm of rainfall. Local adaptation, however, does not measurably protect against it: its attenuation of upstream rainfall is smaller and not distinguishable from zero. Protection appears to be local in a strong sense, holding back the water that falls nearby and doing much less against a volume arriving already channelled from above, which calls for coordination at the level of the basin rather than the site.

Fifth, the evidence of protection is concentrated in projects classified as preventive rather than remedial. When preventive and remedial investments enter the specification jointly, completed preventive works retain a positive and statistically significant protective interaction, while the corresponding coefficient for remedial works is not statistically distinguishable from zero. This comparison suggests that the stated purpose of public investment matters for understanding its economic effects, although the estimates do not rule out protective benefits from remedial interventions.

The paper contributes to two literatures. The first is the economics of adaptation. Carleton et al. (2024) distinguishes the two channels through which adaptation operates, before the hazard and after it, and I build the classification on this distinction. A large empirical literature infers adaptation from how the damage caused by climatic shocks of given intensity, such as extremely hot days, changes over time (e.g. Barreca et al., 2016; Burke and Emerick, 2016; Carleton et al., 2022).

My first contribution is to measure adaptation directly. I develop a procedure that reads the free-

text descriptions of an administrative registry of public investment, identifies the projects that address hydrological risk, and classifies each as ex ante or ex post.

The empirical evidence on the returns to adaptation is thinner. Models calibrate its effectiveness rather than estimate it. Catalano et al. (2020) distinguish, as this paper does, preventive measures such as infrastructure from remedial ones such as relief and reconstruction, and study how fiscal policy should combine them; Fried (2022) shows that federal disaster aid weakens the incentive to adapt in advance, an effect that federal subsidies to adaptation more than offset. Empirical estimates follow three routes. The first values protection through the prices of the assets behind it: Bradt and Aldy (2026) find that federal levees raise the value of protected homes by 3 to 4 percent, and lower it in neighbouring areas onto which they shift the risk. The second measures the damage that protection avoids: Taylor and Druckenmiller (2022) estimate from flood insurance claims what a hectare of wetland saves each year. The third exploits aggregate data: Ficarra and Mari (2025) find that spending on flood defences across 309 English counties reduces the likelihood of flooding but is less effective at limiting economic damage once a flood occurs. None of these studies provides direct causal evidence on the ability of adaptation policies to protect local economic activity, that is, the production that firms lose, or keep, when the water arrives.

Therefore, my second contribution is to estimate how much revenue completed public protection preserved for firms in a realised flood, comparing firms with the same planned works nearby but a different share of them finished when the flood came, and instrumenting completion with the administrative speed of the funding pipelines. To my knowledge, this is the first causal evidence on the direct effect of preventive policies on the economic activity of firms exposed to a rainfall shock. It also shows that the return is specific to works built to prevent: works built to repair earlier damage show no comparable effect.

The second studies floods and firm performance. Floods damage firms through the balance sheet (Hosono et al., 2016; Basker and Miranda, 2018) and through supply chains (Barrot and Sauvagnat, 2016; Carvalho et al., 2021), reaching firms the water never touched, and it differs sharply across firms (Tarsia, 2026). Work on recovery asks what helps once the damage is done. Collier et al. (2024) follow more than 167,000 US firms that applied for federal disaster recovery loans and find that the loans reduce exit and bankruptcy and raise employment over up to seven years, in part by crowding in private credit that firms struggle to obtain after a disaster. Erda (2026) shows that large floods set off creative destruction among US manufacturing plants: the least productive exit, productive entrants acquire their machinery, and survivors upgrade their equipment as they rebuild, while federal disaster spending sustains this reallocation by easing the financing of young, small firms. All of this takes the event as its starting point. I look backwards from it, asking what protection already standing did to the losses that materialised, and establish that completed ex-ante works offset a substantial share of the revenue loss from extreme rainfall, while works built to repair earlier damage do not.

The remainder of the paper is structured as follows. Section 2 describes the institutional setting. Section 3 presents the data and the construction of the adaptation exposure measure. Section 4 develops the empirical strategy and identification. Section 5 presents the results and discusses mechanisms and limitations. Section 6 concludes.

2 Institutional Framework

Italy is one of the European countries most exposed to hydrogeological risk. Some 94.5 percent of municipalities are classified as at risk from landslides, floods, avalanches or coastal erosion, and 6.8 million people live in areas of medium flood hazard, defined by return periods between 100 and 200 years; flood damage alone costs an estimated one to three billion euros a year (ISPRA, 2024). In response to risks of this kind, Italy adopted a national plan for adaptation to climate change in December 2023 (Ministero dell'Ambiente e della Sicurezza Energetica, 2023)[1].

2.1 Flood Risk in Emilia-Romagna

Within this already exposed country, Emilia-Romagna stands out[2]. It is one of Italy's most economically productive regions, contributing approximately 9 percent of national GDP (ISTAT, 2025) and hosting dense concentrations of manufacturing activity in the mechanical, ceramic, food-processing, and packaging sectors. At the same time, it consistently ranks among the regions with the highest values of population and economic assets at hydrogeological risk: nearly 76 percent of regional manufacturing establishments operate in areas classified as flood-prone (Loberto and Russo, 2024). The result is an unusually high concentration of productive activity in territory that is structurally vulnerable to inundation, a combination that makes Emilia-Romagna a natural laboratory for studying the economic consequences of flood risk and the returns to public adaptation[3].

2.2 Public Adaptation: Governance and Investment

Flood protection in Italy operates through a fragmented, multi-level governance structure in which planning, financing, and construction responsibilities are distributed across institutions at the national, interregional, regional, and local levels. Understanding this structure is important for interpreting both the data and the identification strategy.

[1]The plan sets out 361 measures across eighteen sectors, of which twenty-nine address geological, hydrological and hydraulic instability. Two features are relevant here. First, the plan classifies its measures as grey (structural), green (nature-based), or soft (informational, organisational and governance-related), and 274 of the 361, roughly three quarters, are soft; the measures listed under hydrogeological instability are non-structural throughout. The national planning framework therefore does not operate on the margin this paper studies, which is the stock of physical works already built. Second, although the plan was approved by ministerial decree in December 2023, seven months after the events studied here, its technical annexes on sectoral actions date from 2018 and were not revised in response to them. The adaptation planning observed in this paper is thus a regional and local activity, recorded project by project in the public investment registry, rather than an implementation of the national plan.

[2]The region has some 4.45 million residents (ISTAT, 2024) over roughly 22,500 square kilometres, a density of about 200 inhabitants per square kilometre.

[3]The hydrographic network underlying this risk drains the northern slopes of the Apennines through a dense system of rivers flowing towards the Po and the Adriatic , including, from west to east, the Trebbia, Taro, Parma, Enza, Secchia, Panaro, Reno, Santerno, Senio, Lamone, Montone, Rabbi, Ronco, and Savio, as well as an extensive network of reclamation canals covering the Po Valley plain. Apennine rivers are characterised by steep gradients, short catchment response times, and highly variable seasonal regimes, making rapid inundation a recurrent risk during intense rainfall events. The Po Valley plain sits at or below sea level in many areas and depends on pumping infrastructure for drainage, making it particularly vulnerable to prolonged inundation once levees are overtopped or breached.

Governance. At the national level, the Ministry of Environment and Energy Security (MASE) sets regulatory standards and allocates state funding for hydrogeological risk mitigation under the national remote-sensing programme for hydrogeological risk (*Piano Straordinario di Telerilevamento*) and successive programming frameworks. The Po River Basin Authority (*Autorità di Bacino Distrettuale del Fiume Po*, AdbPo)[4], based in Parma, is responsible for basin-wide flood risk management planning under the EU Floods Directive (2007/60/EC, transposed into Italian law by D.Lgs. 49/2010). Its plans identify priority intervention zones and guide the allocation of structural protection works across the basin.

At the interregional level, the Interregional Agency for the Po River (AIPO)[5], established in 2003 as an instrumental body of the four Po basin regions (Piemonte, Lombardia, Emilia-Romagna, and Veneto), manages the main Po stem and its principal tributaries. AIPO is responsible for hydraulic safety works on the primary hydrographic network, including levee maintenance, bank protection, and emergency interventions.

At the regional level, the Regional Agency for Territorial Safety and Civil Protection (*Agenzia per la Sicurezza Territoriale e la Protezione Civile*) coordinates civil protection activities, manages the regional early-warning system, and oversees investment programmes for territorial safety. The regional government co-finances works through its own budget and administers EU structural funds (FESR and cohesion funds) allocated to flood risk reduction.

The primary operational bodies for the secondary and tertiary hydrographic network are the eight first-degree Reclamation Consortia (*Consorzi di Bonifica*) operating in Emilia-Romagna, plus one second-degree consortium[6]. The consortia manage and maintain approximately 19,800 km of drainage and irrigation canals, of which roughly 5,700 km are exclusively drainage, and operate 511 pumping stations with a total capacity of around 2,000 m^3/second[7]. They are funded by a combination of levy contributions from landowners in their district and public grants from the regional and national governments. The Region finances the construction of new or restructured works; routine maintenance is covered by the consorzio levy.

Funding pipelines and project execution. A hydraulic intervention typically originates as a project entry in the OpenCUP registry[8], the national public investment tracking system, assigned a unique project identifier (CUP) and classified by nature, typology, and funding source. Funding sources include state transfers from MASE, regional budget allocations, EU structural funds channelled through the regional authority, and for maintenance works, direct consortia expenditure. Once a project is approved and financed, it is assigned to a contracting authority, typically the relevant Reclamation Consortium (*Consorzio di Bonifica*) or AIPO, and enters the procurement pipeline. The

[4] https://www.adbpo.it/

[5] https://www.agenziapo.it/

[6] The regional association of consortia is *ANBI Emilia-Romagna*: https://www.anbiemiliaromagna.it/.

[7] Figures from the Emilia-Romagna regional government: https://ambiente.regione.emilia-romagna.it/it/suolo-bacino/argomenti/bonifica-idraulica-e-irrigazione.

[8] The OpenCUP portal (*Portale dei dati aperti sugli investimenti pubblici*) is maintained by the Department for Economic Planning and Coordination (*DIPE*) of the Presidency of the Council of Ministers: https://www.opencup.gov.it/.

time between approval and final financial closure varies substantially across funding sources. As documented in Section 5, state- and EU-funded projects face more approval layers, centralised auditing, and co-financing conditionality requirements that systematically slow execution relative to purely regional projects. This differential is the source of variation exploited in the identification strategy.

Spatial distribution of protection. Flood protection investment is not randomly distributed across the territory. Historically, works have been concentrated along the major Apennine rivers that drain the most exposed and economically valuable corridors, the foothills and plains of the Romagna sub-region, the Reno basin east of Bologna, and the area around Modena and Reggio Emilia. Areas at higher observed flood frequency and with greater economic assets at risk tend to receive more planned investment, a pattern that is visible in the data and that constitutes the primary identification challenge addressed in Section 4.

Insurance and compensation. Italy did not have mandatory flood insurance for businesses at the time of the event[9]. Private cover existed but reached roughly 5 percent of firms at the national level (ANIA, 2024). Compensation after an event rests on public intervention[10]. It operates through the declaration of a state of calamity, which activates emergency transfers (*ristori*) to households and firms in the affected areas, claimed through their own municipalities, and grants access to the Extraordinary Wage Support Fund (*Cassa Integrazione Guadagni Straordinaria*, CIGS) for workers in affected establishments. These transfers are administered ex post, typically with substantial delays.

This compensation framework is directly relevant to the empirical design. In the aftermath of the May 2023 floods, the government declared a state of calamity for Emilia-Romagna on 23 May 2023. However, the first compensation transfers to affected firms were not disbursed until 2024. The 2023 revenue estimates in Section 5 therefore capture unmitigated economic damage, prior to any policy-induced recovery. Compensation could potentially confound the 2024 event-study coefficients, the more so because Italian local governments' expenditure response to floods and earthquakes persists for a decade or more after the event (Masiero and Santarossa, 2020; Lodi et al., 2023); I flag this when discussing the post-period dynamics.

[9] An obligation was introduced afterwards. The 2024 Budget Law (Law 213 of 30 December 2023, art. 1, commi 101–112) requires all firms registered in the Italian business register, other than agricultural ones, to insure land, buildings, plant, machinery and equipment against earthquake, flood, landslide and inundation, on pain of an administrative fine and exclusion from public grants, including disaster relief. The deadline has been deferred repeatedly and by firm size: from 31 December 2024 to 31 March 2025 for large firms, then to 30 September 2025 for medium-sized firms and 31 December 2025 for micro and small ones, with further extensions into 2026 for tourism, food service, fishing and aquaculture. The obligation was therefore neither in force nor binding in May 2023. Take-up has since risen but remains partial: a year and a half after the first deadline, fewer than one firm in five was covered (IVASS, 2026).

[10] How disaster losses are financed, privately and publicly, has been studied mainly for the United States. Gallagher (2014) documents that private flood-insurance take-up rises after an event and decays over the following decade, and Deryugina (2017) that the public fiscal response runs mainly through social insurance rather than through disaster aid. For Italy the demand side, and firms' insurance decisions in particular, has not been documented.

2.3 The May 2023 Floods

The May 2023 episode has no precedent in the region's hydrological record, which begins in 1921 (Brath et al., 2023). It comprised two temporally distinct storm cells that struck the same territory within fifteen days, compounding damage and leaving soils fully saturated between the two events.

Event 1 (2–4 May). Persistent rainfall fell on the eastern Apennines from the evening of 1 May through 3 May, with cumulative totals above 200 mm in the hills between Bologna and Forlì-Cesena[11]. Rivers overflowed and levees breached in the provinces of Forlì-Cesena, Ravenna and Bologna. The official perimeters cover 86 km^2, 0.38 percent of the regional surface. Two people died and several thousand were evacuated.

Event 2 (16–17 May). A second and far larger system arrived on 16 May, before the soils had drained. Totals over 48 hours set records for the month, reaching 254.8 mm at the Trebbio rain gauge, in the municipality of Modigliana, and 254.6 mm at San Cassiano, in Brisighella, both in the Lamone basin[12]. Figure 1 places the episode in the distribution of daily regional rainfall since 2001: the six rainiest days of May 2023 lie far into the right tail.

By 17 May twenty-three watercourses had burst their banks, inundating 100 municipalities and triggering 65,598 landslides on the slopes above (Brath et al., 2023). The perimeters for the second event cover 853 km^2, 3.80 percent of the regional surface, nearly ten times the area of the first[13]. Figure 2 maps both footprints.

In total, the two episodes caused seventeen deaths, displaced more than 36,000 residents, closed over 500 roads, and generated €8.5 billion in damages certified to the European Union by the regional government (European Parliament, 2024)[14]. The economic impact was concentrated in the Romagna industrial districts, where manufacturing plants, logistics warehouses, and food processing facilities, many operating on just-in-time production models, suffered direct physical damage to premises, machinery, and inventories, as well as access disruption through flooded road corridors.

The severity of the episode is not incidental to the research design. Severe events are the part of the hazard distribution that is becoming less rare, since heavy precipitation intensifies with warming (IPCC, 2021). The estimates that follow are therefore returns to protection under a tail realisation

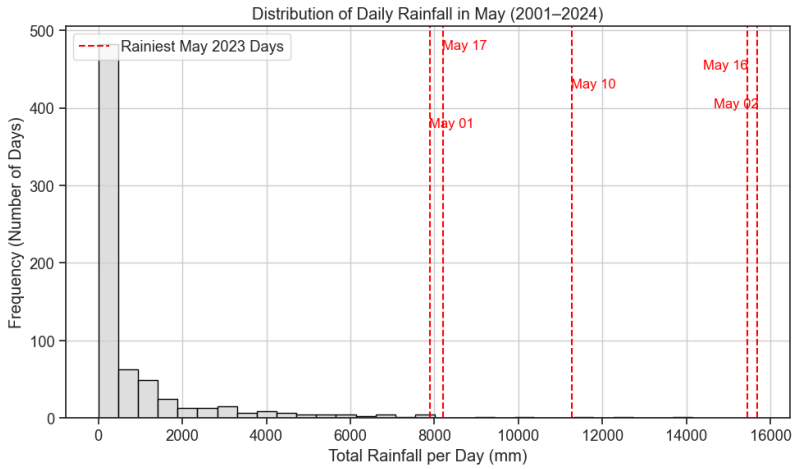
[11] ARPAE Emilia-Romagna, *L'evento meteo-idrogeologico del 1–4 maggio 2023*, https://www.arpae.it/it/notizie/levento-meteo-idrogeologico-del-1-4-maggio.

[12] ARPAE Emilia-Romagna, *Rapporto IdroMeteoClima Emilia-Romagna, Dati 2023*, https://www.arpae.it/it/notizie/speciale-alluvione-idrometeoclima-2023.pdf. Brath et al. (2023) estimate return periods of 100–500 years for the individual events and greater than 1,000 years for the combination.

[13] This figure is not directly comparable with the approximately 540 km^2 of inundated plain reported by Brath et al. (2023). The commission reconstructs the volume of water released by the overflow of the twenty-three watercourses and the extent of plain it flooded; the perimeters used here are mapped by photo-interpretation of remote-sensing imagery and cross-validated against ground reports, and cover inundation from all causes. The two are different quantities produced by different methods, and the analysis uses the perimeters throughout.

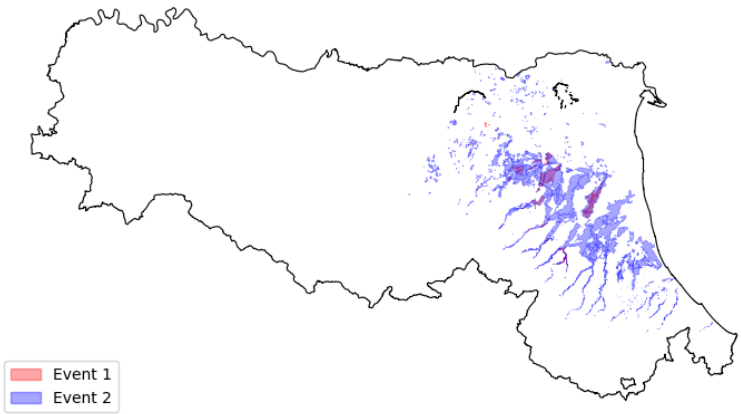
[14] See also Agenzia per la Sicurezza Territoriale e la Protezione Civile, Regione Emilia-Romagna, *Alluvione, un anno dopo*, May 2024: https://protezionecivile.regione.emilia-romagna.it/notizie/2024/maggio/alluvione-un-anno-dopo.

Figure 1: Distribution of daily regional rainfall in May, 2001–2024.



Notes: The x-axis reports total daily precipitation summed across all 1,131 grid cells of the ARPAE ERG5 product (5×5 km resolution), expressed in mm. Red dashed lines mark the six rainiest days of May 2023. All six lie far into the right tail of the 2001–2024 distribution, confirming the exceptional meteorological intensity of the two flood events.

Figure 2: Spatial footprint of the May 2023 floods, Emilia-Romagna.



Notes: Pink polygons denote areas inundated during Event 1 (2–4 May); blue polygons denote areas inundated during Event 2 (16–17 May). Event 1 covers approximately 86 km^2 (0.38 percent of the regional surface); Event 2 covers approximately 853 km^2 (3.80 percent), nearly an order of magnitude greater in extent and substantially more spatially fragmented. Source: Agenzia per la Sicurezza Territoriale e la Protezione Civile, Regione Emilia-Romagna, version V6.

rather than under typical conditions, which is what makes them informative and, at the same time, bounds what they can be taken to imply for a warmer future. I return to that question in Section 5.5.

The spatial concentration of the damage, the sharp meteorological trigger, and the availability of high-resolution flood perimeters from the regional civil protection agency make this episode well suited for the quasi-experimental identification strategy developed in the following sections.

3 Data

The design requires assigning to each firm four quantities that no single source records: the stock of hydraulic protection near it before the flood, the rainfall that fell on it, whether it was inundated, and its revenue. The three subsections that follow construct them in turn. Section 3.1 isolates, from the registry of public investment projects, the works that address hydrological risk; separates those undertaken in anticipation of a hazard from those undertaken in its aftermath; matches each to the watercourse it was built on; and distinguishes what had been approved from what had been finished by May 2023, which is the distinction the identification strategy turns on. Section 3.2 assembles the hydrographic network on which those works sit, the official perimeters of the areas that flooded, and the gridded rainfall of the event. Section 3.3 geocodes the firms and places each one relative to both. The result is a firm-level panel in which exposure to the shock and exposure to protection are measured separately, at the address of the establishment.

3.1 Public Adaptation Data

I measure public hydraulic adaptation using the OpenCUP portal[15] (September 2025 release), Italy's centralised registry of public investment projects managed by the Department for Programming and Coordination of Economic Policy (DIPE). Each record carries a unique project identifier (CUP), structured metadata on location, financial volume, and implementing authority, and free-text fields describing the intervention's nature and objectives (see Figure A.1). After restricting to Emilia-Romagna and applying cleaning steps that remove non-investment entries (reported in Table A.1), the working dataset contains 600,823 unique CUPs. The registry structure and the variable definitions are described in Appendix 1.

Classification pipeline. For classification, I use three standardised administrative fields and the short narrative description written by the implementing authority. The administrative fields assign projects to broad spending categories that do not distinguish hydraulic works from other public investment. The narrative description does, but it is unstructured, in Italian, and covers 600,823 records, so it cannot be read by hand.

I therefore read the four fields together with a large language model[16]. The pipeline sends the concatenated text to OpenAI's GPT-4o-mini in a zero-shot setting, with no labelled training data,

[15] https://www.opencup.gov.it/portale/web/opencup/home.

[16] The fields are `natura_intervento`, `tipologia_intervento`, `categoria_intervento` and `descrizione_sintetica_cup`; the first three are the standardised administrative labels, the fourth the narrative

and asks two questions in a single prompt (Appendix 1.2): whether the intervention addresses hydrogeological risk, and, conditional on it doing so, whether it is ex ante or ex post. The distinction follows the taxonomy of Carleton et al. (2024), with two differences. Ex-ante works are forward-looking measures that reduce future risk, such as levees, slope stabilisation and drainage upgrades, while ex-post works repair damage and restore functionality after an event. The first difference is that the categories are mutually exclusive by construction: a project cannot be both reactive and preventive, so a repair that also raises the standard of protection is recorded on one margin only, according to the purpose the implementing authority declares in the text. The second difference is the unit. The taxonomy classifies adaptation itself, whereas I classify public projects, and a project's cost need not be adaptive capital in its entirety. When an embankment road is rebuilt, part of the outlay replaces worn-out capital and part of it raises resilience to inundation; the registry reports only the total, so the two cannot be separated. What I measure is therefore spending on projects whose declared purpose is protective, which is an upper bound on the adaptive capital they deliver.

The model returns one of three mutually exclusive answers, which a deterministic rule converts into three binary indicators, hydrogeological relevance, preparedness and remediation, that are then merged with the project's structured metadata (Table A.2). One call per project classified the full sample for approximately €55. Reading the narrative rather than matching keywords also avoids the error that dictionaries make systematically, flagging records in which a river name appears in a firm name or a street address rather than in a description of works.

Validation. I validate the measure in three ways. First, I compare the algorithmic labels with a hand-coded subsample, assessing the three indicators by hand and treating the manual assessment as the reference wherever it disagrees with the model (Appendix 1.3.1). Exact-match accuracy across the three indicators is 94.6 percent. Precision, the share of the projects the classifier calls hydraulic that the manual review confirms, is 0.918 for hydrogeological relevance, 0.893 for preparedness and 0.888 for remediation. The classifier therefore works well, but roughly one project in ten of those it selects may be mislabelled. For preparedness the errors are over-inclusion only: the audit finds projects that the review would have excluded, and none that it would have kept. Remediation is the only margin on which the classifier also omits projects, missing 5.4 percent of those the review keeps, although there too over-inclusion is the more common error.

The main analysis uses preventive works alone. If over-inclusion is a threat there, it means I am adding to a firm's exposure projects that deliver no protection, which would pull the estimated effect per euro towards zero; to that extent the magnitudes below would be a lower bound.

Second, I compare the pipeline with two cheaper alternatives, as in Geyiktepe and Rodrik (2026): i) a keyword dictionary applied to the project descriptions and ii) the administrative sector codes already recorded in the registry. The dictionary over-reads geography, flagging records in which a river name appears in an address or a firm name rather than in a description of works. The sector

written by the implementing authority. They are concatenated in that order into a single string, so that a record reaches the model as, for example, *Construction of public works – other – watercourses – basins rivers Montone, Rabbi. Works for the stabilisation of riverbanks and riverbeds.* Table A.13 reproduces this and two further records in full.

codes have almost perfect precision but miss 27 percent of genuine hydraulic projects, most of them post-flood repairs filed under road infrastructure. Neither recovers the distinction between works built in anticipation and works built in response (Appendix 1.3.2).

Third, I test the classified series against realised weather, to see whether the two margins behave as the adaptation literature holds they should, with prevention responding to expectations and repair to realisations (Carleton et al., 2024; Fried, 2022; Van der Straten, 2023). Estimating local projections of quarterly project registrations on extreme-rainfall days across municipalities, remedial registrations rise one quarter after the shock, by 0.229 projects per extreme-rain day ($p < 0.001$), and remain elevated thereafter, while preventive registrations show no positive short-run response under any definition of the shock (Appendix 1.3.3).

The three checks together support the measure. The labels are not error-free, but they are more accurate than the alternatives available at comparable cost, and only the language model recovers the distinction between works built in anticipation and works built in response, which is what the analysis requires and what neither keywords nor sector codes supply. Reading the full registry cost approximately €55, so the same procedure can be applied to other policy domains where projects are described in free text.

Scale and composition. The registry lists a project once for each municipality it covers, so a project with works in several municipalities occupies several rows. The classification identifies 118,310 hydro-related rows out of 2,414,688 total observations (4.90 percent). Collapsing to one observation per CUP yields 76,902 hydro-related projects out of 1,052,460 unique projects (7.31 percent; Table 1)[17]. In real terms, hydro-related investments account for €47.4 billion out of €290.4 billion total (16.3 percent; Table 2). The portfolio is predominantly preventive: ex-ante projects represent 54.8 percent of hydro CUPs and 73.4 percent of hydro spending, with ex-post interventions accounting for the remaining 45.2 percent of projects and 26.6 percent of expenditure (Section 1.4.1 and Table A.6).

Figure 3 traces the annual evolution of this portfolio. In terms of project counts (left panel), ex-ante interventions outnumber ex-post ones until 2022, with both series trending upward from the mid-2010s; after 2022 ex-post projects surge, exceeding 5,000 in 2023 and 6,000 in 2024. In terms of spending (right panel), ex-ante investment dominates throughout, reaching €3.6 billion in 2024, while ex-post spending remains systematically lower and peaks at approximately €1.8 billion in 2023. Changepoint detection on the quarterly series dates these movements, selecting three level shifts without prior information on their number or their location (Figure A.7). The first, in 2010Q3, appears in every category at once and reflects the registry rather than investment: Law 136/2010, in force that quarter, made the CUP mandatory on public-works payments. The second, in 2018Q1, falls on preventive registrations alone and coincides with the opening of a multi-year national fund for municipal works on hydrogeologically fragile land. The third, in 2023Q2, the quarter of the flood, falls on remedial registrations alone, with no corresponding shift in the preventive series. The two margins therefore

[17] Shares are computed over all 1,052,460 projects with valid registration dates. The companion methodological chapter reports the identical hydro count as 12.8 percent of the 600,823 projects that remain after standardised non-investment entries are removed prior to classification, so the two figures differ only in the denominator.

Table 1: Hydro-related projects and watercourse-assigned policies

	Hydro		Hydro on watercourse		Non-hydro	
	Count	Share (%)	Count	Share (%)	Count	Share (%)
Total rows	118,310	4.90	14,326	0.59	2,296,378	95.10
Unique CUPs	76,902	7.31	8,680	0.82	975,558	92.69

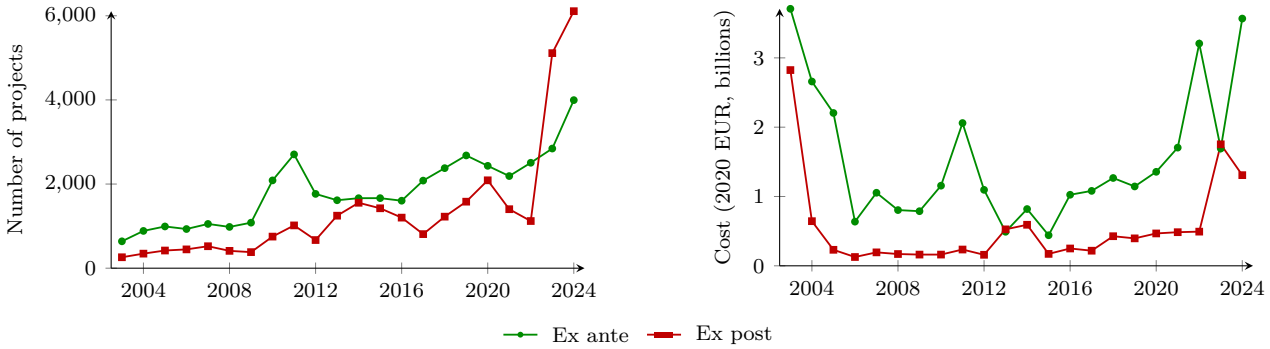
Notes: "Hydro" comprises all hydro-related projects in the Emilia-Romagna extract. "Hydro on watercourse" denotes the subset matched to at least one river or canal segment. Shares are computed relative to the full universe of Emilia-Romagna projects with a valid registration date (2,414,688 rows; 1,052,460 unique CUPs), of which the 600,823 projects retained for classification (Section 3.1) are a subset.

Table 2: Project-level cost shares of hydro-related interventions (2020 euros)

	Total cost	Share of total (%)	Share of hydro (%)
All projects	290,430,863,566	100.00	–
of which: Hydro (all)	47,361,390,546	16.31	100.00
of which: Hydro on watercourse	8,317,250,642	2.86	17.56

Notes: Cost figures deflated to constant 2020 euros using the Italian CPI (AMECO, European Commission). One observation per CUP. "All projects" is the full universe of 1,052,460 Emilia-Romagna CUPs with a valid registration date, including standardised non-investment entries not sent to classification; the companion methodological chapter instead reports composition over the 600,823-project classification sample (after those entries are removed), where the totals are correspondingly smaller. The two bases coincide to the euro on like-for-like sums.

Figure 3: Annual ex-ante and ex-post hydraulic adaptation in Emilia-Romagna, 2003–2024.



Notes: The left panel shows the annual number of hydro-related projects classified as ex ante (preventive, in green) and ex post (remedial, in red) by the LLM pipeline. The right panel shows the corresponding annual spending in constant 2020 euros. Classification based on the zero-shot GPT-4o-mini pipeline described above; source data from OpenCUP (September 2025 release).

keep different time: repair follows a realisation of the hazard, while prevention tracks funding rules (Section 1.4.2). The preventive expansion of 2024 is consistent with belief updating after the flood, but equally with funding that became available on its own schedule, and the aggregate series cannot separate the two.

Watercourse matching. Each hydro-related project is matched to specific watercourse segments by searching its description for the names of the rivers and canals recorded in the official hydrographic network. The match requires a harmonisation step on both sides. On the network side, watercourse names are standardised, abbreviations expanded and a small number of entries with no reliable name attribution removed (Appendix 1.6.1). On the project side, the descriptions are cleaned before matching: Italian administrative text names water bodies with a plural generic term followed by a list of proper names, so that *fiumi uso, marecchia, ausa* must be expanded into the singular forms under which each watercourse is registered, and truncated or misspelt forms must be repaired. Table A.13 gives worked examples and Appendix 1.6.2 documents the rules and the quality checks. As shown in Tables 1 and 2, 8,680 CUPs (11.3 percent of hydro-related projects) are successfully matched to at least one river or canal segment, corresponding to €8.3 billion in investment (17.6 percent of hydro expenditure). Among these, 6,529 projects (75.2 percent) are associated exclusively with rivers, 1,816 (20.9 percent) exclusively with canals, and 335 (3.9 percent) with both. River-linked projects dominate both in frequency and financial scale, with an average project cost close to €1 million; canal-linked projects are smaller and more homogeneous. Table A.16 reports the full distribution of project-level costs by watercourse type.

Planned and closed adaptation. OpenCUP records a closure date alongside the registration date, so for each matched project I can distinguish planned from closed adaptation by the timing of the flood:

1. **Planned adaptation:** projects approved before May 2023, regardless of completion status. Planned spending records where the authorities committed funds, and so potentially where they judged flood risk to be highest;
2. **Closed adaptation:** projects with confirmed financial closure before May 2023, representing physical protection actually in place during the flood.

Completion enters as a binary: a project is either closed before May 2023 or it is not. The registry reports financial closure but no intermediate progress milestones, so I cannot observe how far an unfinished work had advanced, nor weight it by the protection that partial construction may already provide. If works under way protect in part, closed adaptation understates the protection in place at the time of the flood.

I construct a firm-level adaptation exposure index as an inverse-distance weighted average of cu-

mulative pre-flood closed expenditure on watercourse segments within 250 metres of each firm:

$$\text{Adapt}_i = \sum_{r \in \mathcal{R}_i} \omega_{ir}\, C_r, \qquad \omega_{ir} = \frac{1/d_{ir}}{\sum_{k \in \mathcal{R}_i} 1/d_{ik}}, \qquad \mathcal{R}_i = \{r : d_{ir} \leq 250 \text{ m}\} \tag{1}$$

where C_r is cumulative closed expenditure in constant 2020 euros on segment r and d_{ir} is the distance from firm i to segment r. The set $\mathcal{R}_i$ includes every segment within 250 metres, whether or not any works were carried out on it, so Adapt_i is the average completed spending on the watercourses around the firm, with nearer segments weighing more: a firm with €1 million of Adapt_i faces segments that carry on average €1 million of completed works. Firms with no completed works on the segments within 250 metres have $\text{Adapt}_i = 0$. I construct the analogous index on planned expenditure, W_i, with the same weights. Cost figures are deflated to constant 2020 euros using the Italian CPI from the AMECO database of the European Commission[18].

3.2 Water Data

Flood perimeters. To characterise the spatial footprint of the May 2023 floods, I use the official flood-perimeter dataset produced by the Regional Agency for Territorial Safety and Civil Protection of Emilia-Romagna (Regione Emilia-Romagna, 2025)[19]. The dataset delineates inundated territories associated with the two major breach-and-flood episodes of the month—Event 1 (2–4 May 2023) and Event 2 (16–17 May 2023). Flooded areas are mapped through photo-interpretation of multiple remote-sensing sources and are cross-validated with ground information from local authorities and reclamation consortia.

Event 1 comprises 150 mapped polygons and inundates about 86.3 km^2, corresponding to 0.38 percent of Emilia-Romagna's surface. Event 2 is markedly larger: it includes 2024 polygons and covers roughly 852.8 km^2 (3.80 percent of the regional surface), nearly an order of magnitude greater in spatial extent and substantially more fragmented. Figure 2 visualises these differences, showing the relatively localised footprint of Event 1 against the broad and dispersed inundation pattern of Event 2.

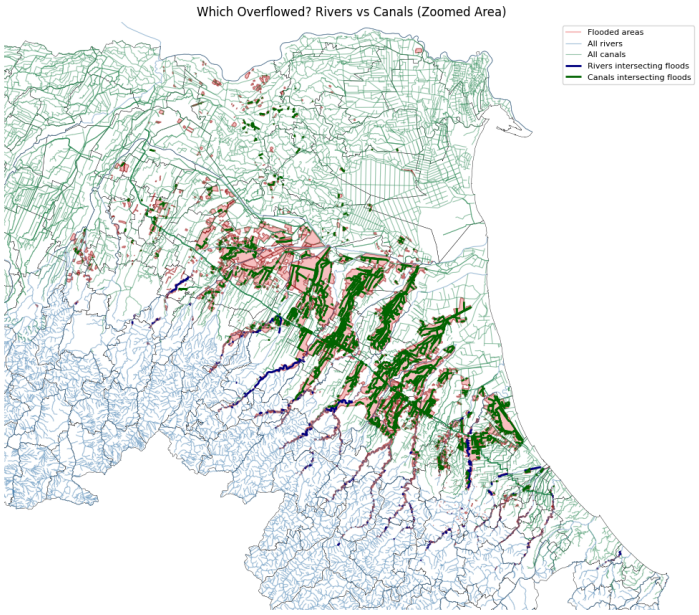
Watercourses. The hydrographic layers used in the analysis draw on two ARPAE Emilia-Romagna products: a harmonised canal shapefile and a river network distributed as individual tiles by the regional cartographic service (*Servizio Cartografico Regionale*)[20], which are merged into a single layer comprising 7,079 simple and 1,327 multi-part river geometries, and 11,619 simple and 465 multi-part canal geometries (Figures A.11; construction details in Appendix 1.6.1). Both rivers and canals are retained. Overlaying the official flood perimeters with the full hydrographic network (Figure 4) shows that the May 2023 inundation was primarily driven by rivers descending from the Apennines, whose

[18] https://dashboard.tech.ec.europa.eu/.

[19] Metadata and documentation: https://metasfera.regione.emilia-romagna.it/ricerca_metadato?uuid=arlst:2025-05-08T141038. The analysis uses version V6 of the perimeters, updated with institutional contributions and adopted under DSG 45/2025. A subsequent version V7, adopted under DSG 88/2025 and incorporating further contributions from territorial authorities with data updated to 19 September 2025, was released after the dataset used here was assembled. The analysis will be reproduced on V7.

[20] Available at https://geoportale.regione.emilia-romagna.it/.

Figure 4: Flood perimeters and hydrographic network, May 2023 (zoomed area).



Notes: The map shows flooded areas (pink polygons), the full river network (light blue), and the full canal network (light green) for the most severely affected sub-region of Emilia-Romagna. Dark blue lines indicate river segments intersecting flood perimeters; dark green lines indicate canal segments intersecting flood perimeters. Primary inundation was driven by Apennine rivers; the downstream canal network intersects a substantial share of the flooded area.

rapid response to intense rainfall caused widespread breaching along their courses into the plain. At the same time, a substantial portion of the flooded area intersects the canal network of the Po Valley. This does not imply that canals themselves overflowed; rather, the maintenance and hydraulic capacity of these canals, responsibilities that fall under the Reclamation Consortia (*Consorzi di Bonifica*), may have been a material factor in determining how floodwaters propagated and accumulated across the territory. Excluding canals would therefore risk understating the adaptation infrastructure relevant to firms in the plain.

Rainfall. This study uses the ERG5 gridded meteorological product released by ARPAE Emilia-Romagna[21]. ERG5 provides hourly and daily observations for the main meteorological and agrometeorological variables from 2001 onwards, obtained by spatially interpolating measurements from the regional station network onto a regular grid covering the whole territory[22]. Each grid cell is approximately 25 km^2 (about 5×5 km), allowing firms to be matched to highly localised weather conditions during the May 2023 flood episode. Table A.17 (Panel A) summarises the structure of the grid used in the analysis, while Panel B lists the raw hourly weather variables available in ERG5.

I exploit the hourly precipitation field to construct grid-cell measures of rainfall intensity and

[21] https://dati.arpae.it/dataset/erg5-interpolazione-su-griglia-di-dati-meteo

[22] Only observations that pass ARPAE's daily quality controls enter the released series, ensuring reliable high-frequency information at fine spatial resolution. The dataset is intended for short- to medium-horizon climatic and event-based analyses, and is not suited to long-run trend estimation because changes in station density may affect temporal comparability; spatial accuracy is highest for precipitation and temperature, moderate for relative humidity, and lower for wind and solar radiation.

accumulation over May 2023—capturing both short-run bursts and multi-day saturation—and assign each firm the exposure corresponding to the grid cell in which it is located. The resulting cell-level precipitation indicators employed in the empirical analysis are reported in Table A.17 (Panel B).

3.3 Firm-Level Data

The firm-level dataset is drawn from Orbis (Bureau van Dijk), the standard source for firm-level balance-sheet data in European economics research[23].

Every variable in the analysis reaches a firm through its location: the grid cell whose rainfall it receives, the flood polygons it falls in, and the watercourse segments, with the works built along them, that lie within 250 metres. A firm therefore needs one address. I locate each firm from its registered postal address, which is geocoded and then validated against the municipality the firm declares in Orbis, and I keep the firms that file unconsolidated accounts (BvD code U1), which report one legal entity at one registered address rather than a parent company and its subsidiaries combined. Consolidated accounts pool activity across sites, so such a firm has no single location to match. Appendix 1.5 documents the download, the geocoding and the validation steps.

The panel comprises the 74,303 firms that were active and reported revenue in 2022, the last year before the flood[24]. The adaptation analysis, which requires firms to lie near planned hydraulic works, uses the 15,540 of them within 250 metres of at least one (Table A.12).

Firm-level flood exposure. Taking all geocoded firms as the reference population, I assign each establishment to a flood exposure category based on its distance to the nearest flooded polygon. Table A.11 summarises the resulting distribution, and Figure A.9 maps it across the region. A total of 2.21 percent of firms lie inside the flooded perimeters, while 17.5 percent are within one kilometre of a flooded area. The remaining 82.5 percent have no flooded polygon within 1 km and form the reference group in the analysis. Most exposed firms are relatively close to the water: among those within 1 km, around three fifths are within 500 metres. The spatial map confirms that exposure is geographically concentrated in the eastern Romagna corridor, consistent with the flood footprint documented in Section 2.3. The construction of the firm-level distance measures, including the geometric setup, the definition of mutually exclusive exposure bands, and the treatment of firms close to multiple flooded polygons, is described in detail in Appendix 1.5.3.

4 Empirical Strategy

The empirical analysis proceeds in two steps, each with a distinct estimating equation and identification argument. The first step estimates the reduced-form effect of the May 2023 floods on firm revenues,

[23] Orbis is not the full population of Italian firms. It is skewed towards larger establishments, older firms, and limited liability companies relative to partnerships and sole traders. The most comprehensive firm-level registry for Italy is the Unioncamere database, which covers the full population of registered firms; this source is used in Loberto and Russo (2024) but is costly to access and not publicly available. Orbis provides only annual financial statements, which precludes tracking within-year or high-frequency effects; the 2023 revenue estimates therefore capture the net annual impact of the flood rather than its precise timing. These limitations are discussed further in Section 5.5.

[24] The firm data come from a single Orbis extract covering 2016–2024, downloaded on 20 March 2026.

using geographic variation in flood exposure and a difference-in-differences design. The second step estimates whether completed ex-ante adaptation attenuates that damage, using rainfall intensity as a continuous measure of meteorological exposure and exploiting variation in the administrative speed of project completion as an instrument for realised protection.

4.1 Flood Impact

To estimate the effect of the May 2023 floods on firm revenues, I compare the evolution of outcomes for firms located inside the flooded area to those located outside, before and after the event. The sample covers the period 2018–2024, with 2022 as the baseline year — the last pre-treatment period before the flood.

Event study. The baseline specification is:

$$Y_{it} = \sum_{\substack{k=2018 \\ k\neq 2022}}^{2024} \beta_k \cdot \mathbb{1}[t = k] \times \text{Flooded}_i + \alpha_i + \gamma_t + \varepsilon_{it} \tag{2}$$

Here $Y_{it} = \log(\text{sales}_{it})$, defined for firm-years with strictly positive sales, so that the coefficients measure the intensive margin; the extensive margin is estimated separately on an indicator for positive sales[25]. Flooded_i is a time-invariant indicator equal to one if firm i is located within the inundated area as mapped by the official flood perimeters, and zero otherwise. The coefficients β_k capture the differential evolution of outcomes between flooded and non-flooded firms in each year k relative to 2022. Firm fixed effects α_i absorb all time-invariant unobserved heterogeneity; year fixed effects γ_t absorb aggregate shocks common to all firms. Standard errors are clustered at the ERG5 5×5 km grid-cell level[26].

The corresponding static specification pools the post-flood period,

$$Y_{it} = \beta \left(\text{Post}_t \times \text{Flooded}_i\right) + \alpha_i + \gamma_t + \varepsilon_{it}, \tag{3}$$

where Post_t marks the post-flood years. Neither Post_t nor Flooded_i appears on its own: the first is common to all firms in a given year and is absorbed by γ_t, the second does not vary over time and is absorbed by α_i. What identifies β is therefore the interaction, that is, the change in revenue of flooded firms after the flood relative to the change over the same years among firms outside the perimeter. Richer versions replace γ_t with sector-by-year and with grid-cell-by-year fixed effects.

The identifying assumption is parallel trends: absent the flood, flooded and non-flooded firms would have continued on the same trajectory. The pre-treatment coefficients $\beta_{2018}, \ldots, \beta_{2021}$ measure

[25] I do not use $\log(\text{sales} + 1)$. Chen and Roth (2024) show that this transformation is not scale-invariant when the outcome contains zeros, so that the estimated percentage effect depends on the units in which sales are measured. Estimated on $\log(\text{sales} + 1)$, the baseline damage specification gives -0.088 with sales in euros, -0.063 in thousands of euros and -0.015 in millions. Separating the two margins avoids the problem and is more informative about which one the flood operates on.

[26] Clustering at the firm allows a firm's residuals to be correlated across years, but treats different firms as independent, including neighbouring ones. Firms a few hundred metres apart are exposed to the same local conditions, so this is unlikely to hold; clustering at the grid cell allows for correlation both within a firm over time and across firms in the same cell.

whether the two groups were already moving apart before the flood. Section 5 takes up the assumption in full before reporting the estimates.

Distance-gradient DiD. Physical proximity to the flood is not binary. Damage travels beyond the inundation perimeter through supplier and customer links (Barrot and Sauvagnat, 2016; Carvalho et al., 2021) and through the contraction of local activity that follows a disaster (Boustan et al., 2020; Cavallo and Noy, 2011). To capture this spatial gradient, I complement the event study with a specification that partitions firms into mutually exclusive distance bands relative to the nearest flooded polygon:

$$Y_{it} = \sum_{b \in \mathcal{B}} \delta_b \cdot \mathbb{1}[\text{Band}_i = b] \times \text{Post}_t + \alpha_i + \gamma_t + \varepsilon_{it} \tag{4}$$

where $\mathcal{B} = \{\text{inside},\ (0, 100\text{m}],\ (100, 250\text{m}],\ (250, 500\text{m}],\ (500, 1{,}000\text{m}]\}$, $\text{Post}_t = \mathbb{1}[t = 2023]$, and the reference group comprises firms with no flooded polygon within 1 km. The coefficients δ_b trace how the revenue impact decays with distance from the inundated area, allowing for non-linear spillover effects. Firm and year fixed effects are included throughout.

4.2 Adaptation and Firm Resilience: Triple Diff-in-Diff

The second step asks whether firms located near completed ex-ante adaptation works experienced smaller revenue losses during the flood. A natural approach would be to interact the flood indicator with a measure of nearby adaptation. This is however problematic for two reasons.

First, whether a firm ends up inside the flooded perimeter is itself a function of whether nearby levees and drainage works were in place and completed: completed protection may have shielded some firms from direct inundation, making realised flood status endogenous to the very treatment variable I wish to study. Second, adaptation works are disproportionately located on the most vulnerable watercourses, so areas with more protection are also areas with higher underlying flood risk — a confound that makes even the comparison of outcomes across adapted and unadapted firms uninformative absent further controls.

To avoid conditioning on an outcome of adaptation, I measure meteorological exposure using local *rainfall intensity* R_i. Rainfall is unaffected by nearby hydraulic works and predicts whether a grid cell contains flooded ground (Table A.25). The assumption is that firms in wetter and drier cells would have followed the same path in 2023 had the flood not occurred. The firm fixed effects take care of what separates those cells permanently, such as differences in size or sector. The pre-flood coefficients in Section 5.3 then show that their revenues were not already diverging in the years before the flood.

I then interact rainfall with the stock of completed adaptation works in the vicinity of the firm and estimate the following triple difference-in-differences specification:

$$\begin{aligned} Y_{it} = \alpha &+ \beta_1 \left(\text{Post}_t \times R_i \times \text{Adapt}_i\right) + \beta_2 \left(\text{Post}_t \times R_i\right) + \beta_3 \left(\text{Post}_t \times \text{Adapt}_i\right) \\ &+ \beta_4 \left(\text{Post}_t \times R_i \times W_i\right) + \beta_5 \left(\text{Post}_t \times W_i\right) + \alpha_i + \gamma_t + \delta_{p(i)t} + \varepsilon_{it} \end{aligned} \tag{5}$$

where Y_{it} is the revenue of firm i in year t and $\text{Post}_t = \mathbb{1}[t \geq 2023]$. R_i is rainfall intensity in the firm's grid cell during May 2023. Adapt_i is the inverse-distance-weighted average of *completed* ex-ante adaptation expenditure on watercourse segments within 250 metres of firm i by the flood date (equation 1), and W_i is the same average computed on *planned* investment, which proxies for the social planner's ex ante assessment of local flood risk.

Since R_i, Adapt_i, and W_i are all time-invariant at the firm level, their levels and all pairwise interactions with each other are absorbed by firm fixed effects α_i; only their interactions with Post_t enter the specification. Year fixed effects γ_t absorb shocks common to all firms and $\delta_{p(i)t}$ denotes province-by-year fixed effects. Standard errors are clustered at the firm level[27].

Revenue enters in levels here, whereas the damage channel of Section 5 is estimated in logs, because the two channels ask different questions. Inundation damages a firm in proportion to its scale: water in the premises interrupts whatever share of production that firm was carrying out, so the natural object is a percentage loss. A hydraulic work is different. A levee or a retention basin holds back a volume of water over a stretch of watercourse, and it does not hold back more water for a larger firm. What it delivers to the firms behind it is an absolute quantity of avoided flooding, and the question asked of it, how many euros of output a euro of public works protects, is a question about levels[28].

The coefficient of interest is β_1: it captures how the post-flood rainfall gradient in revenues varies with the stock of completed protection nearby, so that a positive β_1 means that adaptation attenuates the revenue loss induced by rainfall. It is identified if, absent the works, that gradient would have evolved in the same way for firms with more and with less completed protection nearby (Olden and Møen, 2022); Section 5 examines the assumption.

Two elements of equation (5) carry the identification. The first is the pair of terms in W_i. Works are planned where the planner judged flood risk to be highest, so Adapt_i is not allocated at random with respect to the outcome. Conditioning on $\text{Post}_t \times W_i$ and $\text{Post}_t \times R_i \times W_i$ holds the planned stock fixed, so that β_1 compares firms on watercourses assessed as equally risky that differ in how much of the planned protection had actually been completed by May 2023. Identification then rests on the execution margin rather than on the targeting margin. Execution is itself not random, which is what Section 4.3 addresses.

The second is the province-by-year fixed effects $\delta_{p(i)t}$. The two May events did not cover the region uniformly: they fell on different watercourses two weeks apart, so the timing and reach of the shock, and of the disruption to roads, power and suppliers that followed it, differ across the nine provinces. Provinces also differ in industrial composition and in the administrative structures through which hydraulic works are commissioned. Absorbing any shock common to firms in a province in a given year leaves β_1 identified from firms observed in the same province and year, differing in rainfall and in completed protection nearby.

[27] The grid cell is the alternative, and here the two are not substitutes: the rainfall of the event is constant within the cell, so clustering there credits the data with far less independent variation in R_i than the firm-year sample contains. The estimates are reported under both, together with Conley (1999) spatial standard errors that let the covariance decay with distance rather than stop at a boundary.

[28] Taylor and Druckenmiller (2022) regress flood insurance claims in dollars on wetland area in hectares, with both sides of the equation in levels, and report the value of a wetland as the dollars of flood damage that a hectare of it avoids.

4.3 Identification: The Endogeneity of Completed Protection

To fix ideas, consider the following outcome equation:

$$Y_i = \mu - \theta F_i - \kappa V_i + \beta\,(D_i \times F_i) + \varepsilon_i, \tag{6}$$

where F_i is the rainfall shock faced by firm i, V_i is its latent flood vulnerability (unobserved by the econometrician), D_i is the stock of completed adaptation nearby, and $\beta > 0$ is the parameter of interest, the extent to which completed protection mitigates the effect of the rainfall shock on revenues. Equation (5) is the panel version of this model: F_i is measured by rainfall intensity R_i, D_i by Adapt_i, and the interaction with Post_t exploits the before-after dimension of the panel to difference out time-invariant firm heterogeneity. The coefficient β_1 in equation (5) is the empirical counterpart of β in equation (6).

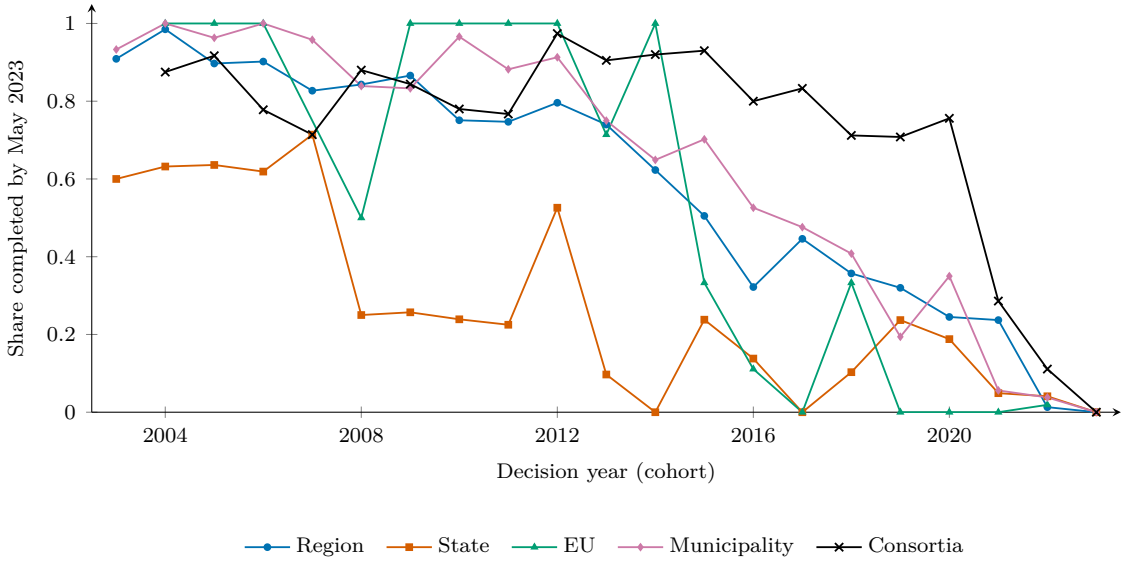
OLS estimation of β_1 is biased because D_i is not randomly distributed. The social planner allocates adaptation projects in response to local vulnerability: $\mathbb{E}[W_i \mid V_i]$ is increasing in V_i, so firms near more-adapted watercourses are precisely those at higher underlying exposure. This generates a positive correlation between D_i and V_i, which, because higher V_i reduces revenues, biases $\hat{\beta}_1$ downward. Conditioning on W_i partially addresses this by holding fixed the risk recognised at the planning stage, but it does not fully resolve the problem because total planned investment is an imperfect proxy for latent vulnerability.

The key to identification is that D_i is determined by two separate margins: a *planning* margin, which is risk-based and endogenous, and an *execution* margin, which determines whether a planned project is completed by the time a shock materialises. Conditional on the menu of projects allocated to a firm's vicinity, variation in how much of that allocation is completed by May 2023 is driven by differences in administrative execution speed across funding pipelines, and not by the firm's residual vulnerability V_i.

The procurement literature predicts that the execution margin carries substantial independent variation. Delay in public works is generated less by the characteristics of the project than by the procedural environment through which it passes: Bandiera et al. (2009) attribute the bulk of measured waste in Italian public purchasing to passive sources, procedural burden, approval layers, administrative effort, rather than to corruption or project selection, and cross-country evidence shows that the same nominal procurement rules produce very different delivery outcomes depending on the capacity of the administration applying them (Bosio et al., 2022; Decarolis et al., 2020; Best et al., 2023). In the Italian setting specifically, completion times respond to features of the administrative environment that are plainly unrelated to the hazard a given work is meant to contain, such as the tenure of the responsible official (Coviello and Gagliarducci, 2017) and the efficiency of the local courts that enforce the contract (Coviello et al., 2018). The funding-pipeline gap documented below is a variable of this kind: it is a property of the approval chain, not of the river.

Motivating evidence. Completion rates by May 2023 vary sharply and persistently across funding sources (Figure 5; Table A.19). Among projects decided between 2003 and 2010, regionally funded

Figure 5: Completion rates by funding source and decision-year cohort.



Notes: Each line plots the share of projects in a given funding-source-by-vintage cohort that reached financial closure by May 2023, as a function of the year in which the project was decided. The downward slope for all sources reflects the mechanical fact that more recent projects have had less time to close. The persistent gap between regional/consortia and state/EU pipelines reflects differences in administrative processing speed.

ones were completed at a rate of 84 percent against 42 percent for state-funded ones, a gap of 42 percentage points. For the 2011–2015 cohorts the gap widens to 47 points (66 against 19 percent), so that a regionally funded project is approximately 3.5 times more likely to have been completed by May 2023. Consortia-funded projects track regional ones closely, while EU-funded projects, fast in the early cohorts, slow to state-like rates from the mid-2010s.

This gap is not driven by differences in project scale or complexity: the cost distributions of regional and state-funded projects overlap substantially across all start years (Appendix 2.1, Figure 2.1). The formal regression evidence, confirming the gap conditional on project size, municipality, and watercourse fixed effects, is reported in Section 5.2 (Table 4, Figure 2.1).

Instrument construction. To recover β consistently, I need variation in completed protection, D_i, that is unrelated to a firm's latent vulnerability, V_i, once planned investment, W_i, is held fixed. The funding-pipeline gap provides exactly this. Once a project is planned, and W_i is therefore fixed, whether it is completed by May 2023 depends on which administrative pipeline it enters. Since pipeline assignment is determined by funding-source rules set at the programme design stage, years before the flood and independently of local firm-level conditions, it shifts D_i without moving V_i. Within the same planned project environment, a firm whose projects fall in faster pipelines ends up with more completed protection not because it is more or less vulnerable, but because of administrative execution speed.

I formalise this by constructing a leave-one-out (LOO) predicted completion index. The leave-one-out construction is important for two reasons. First, it ensures that the instrument is formed from cohort-level execution patterns observed in *other* projects, not from the firm's own project outcomes,

avoiding the mechanical feedback that would arise if a firm's own projects' completion entered both the instrument and the endogenous variable. Second, it means the instrument inherits only the variation in $\pi_{c(p)}$ that is common to the cohort, purging idiosyncratic project-level noise.

For each project p in cohort $c(p)$, the leave-one-out completion rate is:

$$\hat{\pi}_{c(p),-i} = \frac{\sum_{j \in c(p),\, j \notin \mathcal{P}_i} \mathbf{1}\{Completed_j\}}{|\{j \in c(p) : j \notin \mathcal{P}_i\}|} \tag{7}$$

The firm-level instrument is then

$$Z_i^{LOO} \equiv \sum_{r \in \mathcal{R}_i} \omega_{ir} \sum_{p \in \mathcal{P}_r} C_p \, \hat{\pi}_{c(p),-i} = \sum_{c \in \mathcal{C}} w_{ic} \, \hat{\pi}_{c,-i}, \tag{8}$$

where $\mathcal{P}_r$ is the set of projects planned on segment r, ω_{ir} are the weights of equation (1), and w_{ic} collects the weighted planned spending of cohort c around the firm. Z_i^{LOO} is the predicted average of completed protection around firm i, formed by weighting each planned project by the out-of-sample completion rate of its cohort, and is built on the same segments and with the same weights as Adapt_i and W_i. In terms of equation (6), Z_i^{LOO} shifts D_i through $\pi_{c(p)}$ while leaving V_i unchanged, which is precisely what is needed to identify β.

The construction also fixes the estimation sample. For a firm with no planned works within 250 metres every w_{ic} is zero, and so is Z_i^{LOO}: such a firm contributes nothing to the first stage. The adaptation analysis is therefore restricted to firms with $W_i > 0$, that is, to establishments within 250 metres of at least one planned hydraulic intervention. Table A.12 in Appendix 1.5 documents the sample construction and characterises the excluded firms.

First stage and excluded instrument. The endogenous variable of primary interest is the triple interaction $\text{Adapt}_i \times R_i \times \text{Post}_t$, instrumented by $Z_i^{LOO} \times R_i \times \text{Post}_t$. The first stage is:

$$\text{Adapt}_i \times R_i \times \text{Post}_t = \phi + \rho \left(Z_i^{LOO} \times R_i \times \text{Post}_t \right) + \Psi' X_i + \Omega' P_i + \nu_i, \tag{9}$$

where P_i collects controls for the planning margin, at minimum total planned expenditure W_i and its interactions with Post_t and R_i. Identification comes from $\rho \neq 0$: firms whose planned projects fall disproportionately in faster pipelines realise more completed protection by May 2023, conditional on having received the same total planned investment. The double interaction $\text{Adapt}_i \times \text{Post}_t$ is also endogenous and is instrumented by $Z_i^{LOO} \times \text{Post}_t$ in a corresponding first stage; given the construction of the LOO index, this second first stage is mechanically implied by the first[29].

Exclusion restriction. The exclusion restriction requires that, conditional on the planning margin W_i, firm fixed effects α_i, and year fixed effects γ_t, cohort-level pipeline completion rates $\hat{\pi}_{c,-i}$ affect

[29] In practice the 2SLS estimator instruments both endogenous interactions simultaneously. The Kleibergen–Paap F-statistic reported in the results refers to the joint first-stage strength across both instruments.

firm revenues Y_{it} only through realised completed protection Adapt_i. Formally:

$$\text{Cov}\Big(Z_i^{\text{LOO}} \times R_i \times \text{Post}_t,\ \varepsilon_{it} \;\Big|\; W_i, \alpha_i, \gamma_t\Big) = 0. \tag{10}$$

There are at least three ways in which the restriction can fail, and the design of the instrument makes two of them less plausible. The first is timing. If the pipeline through which a project was funded, and hence how fast it was completed, responded to the flood itself or to expectations about how the area would fare, the instrument would be correlated with post-flood revenues for reasons unrelated to the works. For example, if the faster pipelines had been directed towards areas whose economies were expected to grow, firms near the works they funded would have recorded higher revenues after the flood whether or not the works protected them. Pipelines, however, are assigned at the programme design stage, years before the flood, under budget rules written without reference to any particular watercourse, so the instrument does not respond to the flood or to what firms stood to lose from it.

The second is the nearby works themselves. If the completion of the projects around a firm entered the instrument directly, any local factor that speeds those projects up, such as an efficient municipal administration that also serves local firms well, would move the instrument and revenues together. Completion rates are instead computed over the other projects in each cohort, excluding those in the firm's surroundings, so what became of the works that enter a firm's exposure does not feed back into the index that predicts them.

The third is selection on place: the kind of territory a watercourse runs through could determine which programme pays for the works along it. Funding programmes would then come with geographies. Eligibility rules, national priority lists and co-financing requirements could govern where the money goes as much as how fast it moves, so a watercourse financed through an EU programme might lie in a different kind of place from one financed by the region: different in income, in industrial structure, in distance from a city. If the firms in those places also differed in how they recover from a flood, the instrument would carry that difference into the second stage, and β_1 would pick up the territory rather than the protection.

I run three tests against this possibility, reported in Section 5, with the estimates in Appendix 2.1. The first asks whether the pipelines served the places where water is a standing problem: I regress each adaptation measure on long-run May rainfall in the firm's grid cell, and ask in particular whether the instrument is related to it. The second asks whether they served places where firms are larger or better capitalised: I regress the instrument on firm size, employment and assets measured before the flood. The third asks whether the pipelines raised revenues through some channel other than protection in place: I estimate the reduced form year by year, since a pipeline working through anything else would not have waited for the rain to fall[30].

[30] A fourth concern is bundling: if the programmes that complete hydraulic works quickly also deliver other local public investment on the same schedule, the instrument would pick up that investment as well, and the estimate would attribute to protection a return that belongs to public spending more broadly. This one is measurable here. The classification covers the whole OpenCUP registry and not only the hydraulic projects, so the non-hydraulic investment carried by the same funding pipelines is observed alongside it, and work in progress tests whether pipeline speed moves that spending too.

5 Results

Section 5.1 estimates the revenue loss of firms inside the inundated perimeter and how it decays with distance from the water. Section 5.2 reports, at the level of the watercourse, how much of the planned protection had been completed by May 2023. Section 5.3 estimates the effect of completed ex-ante protection on the revenue losses caused by extreme rainfall, instrumenting completed works with the completion record of the funding pipelines they belong to; it also reports how far that protection reaches along the river network, and whether it is specific to works built to prevent rather than to repair. Section 5.4 turns to the channels. Section 5.5 discusses the limitations of the design and the external validity of the estimates.

5.1 Flood Impact on Firm Sales

Identifying assumption. Identification requires that, absent the flood, firms inside and outside the perimeter would have followed the same revenue path. The counterfactual path is not observed, so the assumption cannot be tested directly. What can be examined is how the two groups behaved relative to each other in the years before the flood.

The event-study coefficients on the flooded-by-year indicators for 2018–2021 are small and individually indistinguishable from zero (Figure 6), confirming that no pre-existing divergence between the two groups drives the results.

A placebo with false event dates asks whether the specification would report an effect even where none can exist. I re-estimate equation (3) on the 2018–2022 window alone, in which the flood has not yet occurred, assigning the treatment to a year inside that window. Because no shock takes place, any coefficient the specification returns must reflect pre-existing divergence between the firms that would later be flooded and the rest. None of the three false dates returns one (Table A.27), while the true date returns −0.041.

Baseline flood effect. Table 3 reports the estimate of equation (3). In Panel A the sample runs to 2023, so that Post_t marks the single year of the flood. Firms located inside the perimeter lose approximately 4.0 percent of their revenue in that year relative to firms elsewhere in the region ($p = 0.005$); the estimated coefficient is −0.041 on the log of sales.

The average firm outside the perimeter reported revenue of €2.87 million in 2022, the last pre-flood year. A loss of 4.0 percent corresponds to about €115,000 relative to that benchmark.

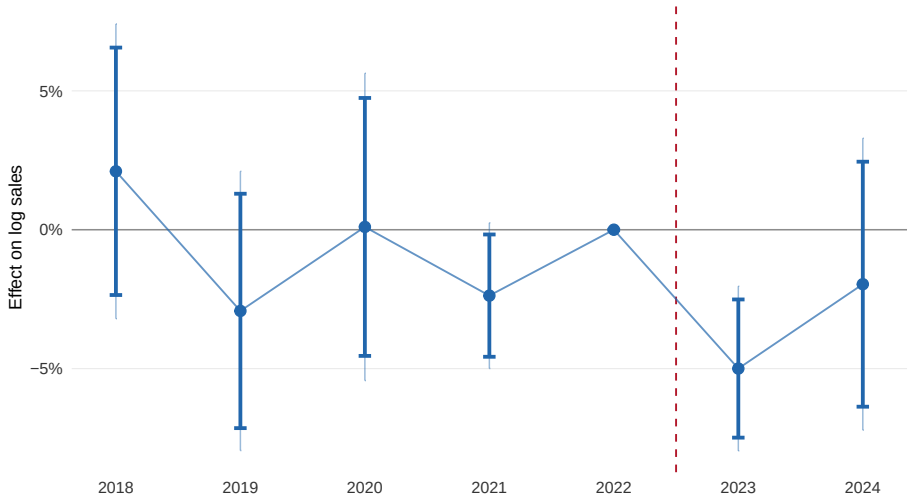
In every column the effect is identified from the change in sales between the pre-flood years and 2023: firm fixed effects mean that each firm is compared with itself over time, and what the columns vary is which unflooded firms supply the counterfactual change against which that within-firm change is measured. In column 1, with firm and year fixed effects, it is the average change of every other firm in the region (−0.041, s.e. 0.015, $p = 0.005$). Column 2 adds sector-by-year fixed effects, so the change of a flooded metalworker is set against the change of non-flooded metalworkers in the same year rather than against the regional average, and industry-specific cycles cannot drive the result (−0.036, s.e.

Table 3: Flood Damage on Firm Sales

	Log sales			
	(1)	(2)	(3)	(4)
Panel A. Sample 2018–2023; Post = 2023				
Post × Inside	-0.0410***	-0.0363**	-0.0409*	-0.0394*
	(0.0145)	(0.0148)	(0.0222)	(0.0214)
Observations	352,707	352,707	352,707	352,707
Panel B. Sample 2018–2024; Post = 2023–2024				
Post × Inside	-0.0290*	-0.0274*	-0.0268	-0.0276
	(0.0157)	(0.0157)	(0.0226)	(0.0226)
Observations	410,003	410,003	410,003	410,003
Firm	✓	✓	✓	✓
Year	✓	✓	✓	✓
Sector × year		✓		✓
Grid cell × year			✓	✓

Notes: Estimates of equation (3). The outcome is log sales for firm-years with strictly positive sales, so the estimates refer to the intensive margin. Inside_i equals one for firms located within the official May 2023 flood perimeter. The two panels differ only in the sample window and in what Post_t therefore denotes: in Panel A the sample ends in 2023 and Post_t marks that year alone, while in Panel B it extends to 2024 and Post_t covers both post-flood years. The fixed-effect ladder across columns is identical in the two panels, and within each panel all four columns are estimated on the same observations. Standard errors, in parentheses, are clustered at the ERG5 5×5 km grid-cell level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure 6: Event Study, Firms Inside the Flood Perimeter



Notes: Coefficients β_k from equation (2), with 2022 as the omitted year. The outcome is log sales for firm-years with strictly positive sales. Flooded_i equals one for firms located inside the official flood perimeter. Bands are 90 and 95 percent confidence intervals; standard errors are clustered at the ERG5 5×5 km grid-cell level. Firm and year fixed effects throughout. The vertical dashed line marks the last pre-flood year.

0.015, $p = 0.014$). Column 3 replaces this with grid-cell-by-year fixed effects, so the counterfactual comes from firms in the same 5×5 km cell — in practice, a flooded firm and an unflooded firm a few streets away. This absorbs every confounder that varies smoothly in space, including the rainfall of the event itself, which is constant within a cell (-0.041, s.e. 0.022, $p = 0.065$). Column 4 imposes both restrictions (-0.039, s.e. 0.021, $p = 0.066$).

The point estimate is therefore stable across comparison groups that differ substantially in how demanding they are. What changes is precision: the standard error rises from 0.015 to 0.022 once the comparison is confined within cells, because that restriction discards most of the variation in the data, and columns 3 and 4 are correspondingly significant only at the ten percent level.

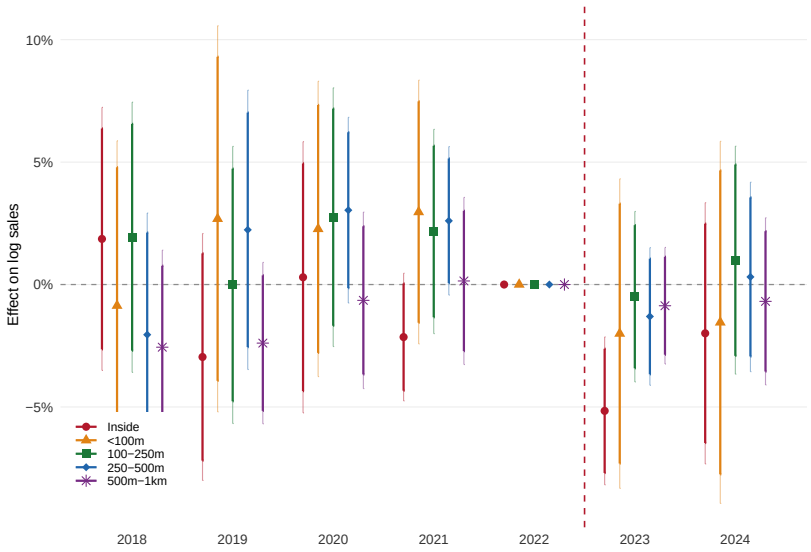
Panel B extends the sample to 2024, so that Post_t now covers both post-flood years. The estimate falls to -0.029 in column 1 ($p = 0.065$) and is no longer distinguishable from zero in the two columns with grid-cell-by-year fixed effects ($p = 0.24$ and $p = 0.22$). This is not a fragile result but an expected one, and the event study explains it: the loss is concentrated in 2023, and 2024 is a year of partial recovery, so averaging the two together attenuates the coefficient towards zero. Which panel is the right one to read depends on the question. Panel A measures the damage the flood inflicted; Panel B measures the average shortfall over the two years that followed, by which time compensation transfers had begun to arrive. I take Panel A as the baseline throughout, since the object of interest is the damage itself rather than the damage net of the policy response.

Because the outcome is the log of sales for firm-years with strictly positive sales, these estimates describe the intensive margin. The extensive margin is negligible: estimating the same equation on an indicator for positive sales gives -0.003 ($p = 0.54$). The flood did not push firms out of production; it reduced the revenue of firms that continued to produce.

Dynamics and recovery. Estimating equation (2) year by year shows how the effect evolves (Figure 6): the coefficients are flat and indistinguishable from zero throughout the pre-flood period, the decline is concentrated sharply in 2023, and it is followed by a partial but incomplete recovery in 2024. The 2023 estimate captures unmitigated economic damage: no public disaster compensation was disbursed that year, so the revenue loss is attributable to the flood shock itself. Public transfers, the so-called *ristori*, only began to reach affected firms in 2024, and even then remained far below the scale of declared losses[31]. The partial rebound in 2024 therefore reflects primarily firms' own adjustments, drawdown of inventories, reallocation of orders, repair of premises, rather than policy-induced recovery. It also explains why the pooled estimate is smaller in absolute value than the 2023 coefficient alone. This recovery is slower than measures of aggregate activity suggest. For the same floods, Gentili (2026) finds from night-time lights that activity in the affected municipalities fell for about a month, so that the effect on regional growth was small; across the world's cities, night-time lights recover within a year of a flood (Kocornik-Mina et al., 2020). At the level of regions, firms may

[31] The OpenCUP classification pipeline identifies all public transfers directed to firms as compensation for flood damage and assigns them an ex-post label, given their explicitly restorative nature. These records can be matched to individual firms through the CUP identifier, making it possible to construct a firm-level panel of received transfers. Analysis currently underway examines whether firms that received *ristori* in 2024 outperformed those that did not, in the spirit of Erda (2026).

Figure 7: Event Study by Distance Band



Notes: Coefficients from equation (4) estimated year by year, with all distance bands entering jointly and 2022 as the omitted year. The outcome is log sales for firm-years with strictly positive sales. Distance is measured to the nearest flooded polygon; the omitted category is firms located more than one kilometre from any such polygon. Bars are 90 percent (thick) and 95 percent (thin) confidence intervals; standard errors are clustered at the ERG5 5×5 km grid-cell level. Firm and year fixed effects throughout. The vertical dashed line marks the last pre-flood year.

even gain: after the 2013 floods in Germany, firms in the affected regions recorded higher sales in the following years (Noth and Rehbein, 2019). The difference is likely to reflect the unit of observation. A municipality or a region contains many firms that the water did not reach, and activity can shift towards them, whereas the estimates here follow the firms inside the perimeter, which bear the loss themselves.

Spatial decay of the damage. Physical proximity to the water, not membership of an administrative polygon, is what should determine the loss. Figure 7 estimates equation (4) dynamically: one series per distance band, all bands entering jointly, with firms located more than one kilometre from any flooded polygon as the omitted category.

The figure does two things at once. It shows that every band is on a flat pre-flood trend, which is a considerably more demanding test of the identifying assumption than a single aggregate test on the treated group; and it shows that the collapse in 2023 is ordered in proximity to the water. Firms inside the perimeter fall by roughly five percent, those within a hundred metres by about two, and the bands beyond two hundred and fifty metres are indistinguishable from zero throughout. The spatial reach of the shock is therefore short, a few hundred metres, which is consistent with damage operating through direct inundation and the immediate disruption of access to premises, rather than through a diffuse contraction of demand in the local economy, which would not attenuate so sharply over such a distance.

Table 4: State Funding and Project Closure

	(1)	(2)
State funding	−0.1552***	−0.2021***
	(0.020)	(0.022)
Log project cost	−0.0482***	−0.0418***
	(0.006)	(0.006)
Municipality fixed effects	✓	
Watercourse fixed effects		✓
Cohort fixed effects	✓	✓
Project weights ($1/n_rows_cup$)		✓
Observations	6,844	6,844

Notes: The dependent variable is an indicator equal to one if the project is closed. Column (1) is a linear probability model with municipality and cohort fixed effects. Column (2) uses watercourse fixed effects and weights each observation by the inverse of the number of rows associated with each project, so that projects linked to multiple watercourses do not receive disproportionate weight. Standard errors clustered at the project level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5.2 Adaptation and the Execution Gap at the Watercourse Level

Before estimating the firm-level effect of adaptation, three patterns at the watercourse level frame the estimates and motivate the identification strategy of Section 4.3.

Endogenous targeting. Planned investment was allocated on risk. Table A.8 compares the segments that breached in May 2023 with those that held. The breached ones had received systematically *more* planned adaptation: €1.75 million and 7.6 interventions on average, against €1.10 million and 3.0, and only 21 percent of them had no planned spending at all against 32 percent of the rest. The planner directed resources towards the reaches most likely to fail, which is what makes a naive comparison of firms near more- and less-adapted segments uninformative, and what the instrument is built to address.

The two margins. What varied was execution. Panel B of the same table reverses the ranking: on the segments that breached, completed expenditure averages €0.19 million against €0.38 million on those that held. Within-river logit regressions state the reversal formally (Table A.24): conditional on the stock of planned interventions, each additional planned intervention is associated with a 9 to 10 percentage point *increase* in breach probability, while each additional closed intervention is associated with a 16 to 22 percentage point *decrease*. Planning tracks risk; completion is what reduces it, and it is completed protection, conditional on planned exposure, that the instrument has to shift.

Despite substantial planned investment, a majority of at-risk watercourse segments lacked completed works when the flood arrived (Table A.7). Among the 1,591 segments with positive planned pre-2023 adaptation spending (Panel A), 965 of them, or 60.7 percent, had an incomplete portfolio at flood time (Panel B). The mean financial shortfall on incomplete segments was €1.84 million, with a median of €192,000 (Panel C). Only 626 segments (39.3 percent) had fully closed their planned works by May 2023.

Funding variation. The gap is not uniform across projects. Table 4 documents one systematic correlate: conditional on municipality and cohort fixed effects, state-funded projects are 15.5 percentage points less likely to be closed by the time of the flood than all other projects in the same cohort. When municipality fixed effects are replaced by watercourse fixed effects, which compare projects on the same river or canal, the estimated gap widens to 20.2 percentage points. These patterns suggest that funding source captures meaningful variation in administrative execution speed, reflecting differences in procurement timelines, audit requirements, and bureaucratic burden across funding pipelines. This is the variation the leave-one-out index exploits. Appendix 2.1 shows that the funding-source gap in closure rates is stable as more recent cohorts are progressively excluded, ruling out a mechanical explanation based on time-to-close.

5.3 Effect of Adaptation on Firm Sales

I first estimate the effect of completed preventive adaptation on firm sales during extreme rainfall, using static and event-study specifications. I then present evidence supporting the exclusion restriction and assess the economic magnitude of the estimates through a back-of-the-envelope calculation. Finally, I extend the analysis to spatial spillovers along the river network (Section 5.3.1) and compare the effects of remedial and preventive adaptation (Section 5.3.2).

Main result. Table 5 reports two-stage least squares estimates of β_1, the coefficient on $\text{Post}_t \times R_i \times \text{Adapt}_i$ in equation (5), with the planned-works terms and the province-by-year fixed effects motivated in Section 4.2. The sample is the 15,540 firms located near a watercourse on which some hydraulic work had been planned by 2023, spanning 522 grid cells and nine provinces[32]. Adaptation is the inverse-distance-weighted stock of completed works within 250 metres of the firm. Standard errors are clustered at the firm level. The two columns differ only in the estimation window. Column 1 restricts the sample to 2018–2023, so that Post_t marks the single year of the flood, as in the damage channel; Column 2 extends it to 2024.

In Column 1, each €1 million of completed works within 250 metres of a firm is associated with €104,338 of revenue preserved per additional 10 mm of extreme rainfall ($p < 0.01$). The rainfall gradient for a firm with no completed adaptation nearby is −129,717 ($p < 0.10$). The separation of the two terms is the object of interest, rainfall reduces revenue, and completed works attenuate that reduction.

Column 2, which adds 2024 to the post period, attenuates the rainfall gradient to −107,186, which loses significance, while the protective interaction rises to €116,768 and remains significant at one percent. The pattern matches the damage channel, and for the same reason: 2024 is a year of partial recovery, so pooling it with 2023 dilutes the shock. Column 1 is the baseline for that reason.

Comparing the estimate with ordinary least squares shows the bias the instrument corrects. Section 4.3 argues that OLS understates the protective effect, because works are placed where expected

[32]The panel is built on firms reporting non-missing sales in 2022, the last pre-treatment year, which leaves 74,303 firms. Of these, 15,540 lie within 250 metres of at least one planned hydraulic work and so contribute variation to the adaptation specifications.

damage is greatest and firms near completed protection are therefore firms whose locations were selected for being vulnerable. On the same sample and with the same controls, OLS puts the protective interaction at €60,322 against €104,338 under two-stage least squares, a gap of €44,016, and the Wu–Hausman test rejects exogeneity of completed works ($p = 0.039$). Appendix 2.2 reports OLS, the reduced form and the two first stages side by side.

Event-study dynamics. Figure 8 traces the year-by-year evolution of both coefficients. The protective interaction is absent before the flood: the coefficients for 2018 to 2021 are −28,914, 1,145, −2,399 and −29,677, none distinguishable from zero. It turns positive in 2023 and 2024, at 91,960 and 114,850. The rainfall series moves in the opposite direction at the same moment, turning sharply negative in 2023. Nothing distinguishes more- from less-protected firms until the water arrives.

Read together, the two post-flood years give 117,125 with a standard error of 59,044 ($p = 0.047$). The dynamic evidence therefore establishes the timing of the effect, it is absent before the flood and present after, more sharply than it establishes the magnitude in any single year.

Support for the exclusion restriction. The instrument is credible only if the speed of a funding pipeline is unrelated to anything else that shaped firm revenues after the flood. Section 4.3 set out three tests (full results in Appendix 2.1).

The first concerns risk. Protection is built where water is a standing problem, so if the fast pipelines served exactly those places, the instrument would be carrying the hazard rather than the administration. Adaptation does follow long-run rainfall, but only on the margin the specification already holds fixed: planned expenditure rises with mean May rainfall in a firm's grid cell ($p < 0.05$), while completed works and the instrument are both unrelated to it (Table A.20). Targeting operates where the planner decides what to build, not where the administration decides how fast to build it, which is the variation β_1 uses.

The second concerns the firms. If the same pipelines served places where firms are larger or better capitalised, the instrument would be picking up the resilience that size confers rather than the protection that works deliver. It does not. Entered one at a time or together, 2022 sales, employment, tangible assets and total assets do not predict it, conditional on the planning margin and the province ($F = 1.87$, $p = 0.11$; Table A.21).

The third concerns timing, and it is the sharpest of the three. Whatever else a fast pipeline might deliver to the firms around it, it would deliver it in the ordinary course of business, not in the one year in which heavy rain fell. Estimating the reduced form year by year, the coefficient on $Z_i \times R_i$ is indistinguishable from zero in every pre-flood year and turns sharply positive in 2023, at 73,496 ($p < 0.001$), and again in 2024, at 89,022 ($p = 0.004$) (Table A.22). The instrument is inert until the water arrives.

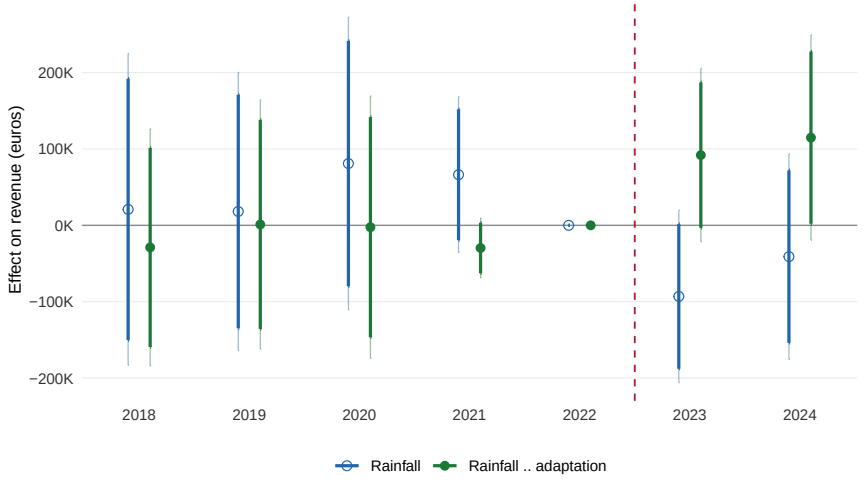
Economic magnitude. The natural point at which to read the estimates is the event that occurred. Rainfall in May 2023 averaged 22.2 mm over three hours across the firms in the sample, and the average

Table 5: Effect of Ex-Ante Adaptation on Firm Sales

	(1) 2018–2023 Post = 2023	(2) 2018–2024 Post = 2023–2024
Post × Rain	-129,717*	-107,186
	(71,574)	(75,198)
Post × Rain × Adaptation	104,338***	116,768***
	(28,366)	(32,278)
First-stage F	309.8	296.8
grid-cell clustering	11.2	11.0
Observations	81,492	94,549
Firms	15,540	15,540

Notes: Two-stage least squares estimates of equation (5); the dependent variable is annual revenue in euros. R_i is maximum precipitation over any three consecutive hours in May 2023 (10 mm units); Adapt_i is the inverse-distance-weighted stock of completed ex-ante adaptation within 250 metres of firm i by May 2023 (€1M units). The columns differ only in the estimation window. Planned-works controls and firm, year and province-by-year fixed effects are included throughout; the completed-works terms are instrumented with the leave-one-out completion index of equation (8). The first-stage F is the Wald statistic on the excluded instruments, reported under firm clustering and, in the row below, under grid-cell clustering. Standard errors clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure 8: Event Study: Rainfall and the Protective Effect of Adaptation



Notes: Year-specific coefficients from the dynamic version of equation (5), estimated by two-stage least squares with 2022 as the omitted year. The blue series is the coefficient on $R_i \times \mathbb{1}[t = k]$, the revenue effect of rainfall for a firm with no completed adaptation nearby; the green series is the coefficient on $R_i \times \text{Adapt}_i \times \mathbb{1}[t = k]$, the additional effect per €1 million of completed works within 250 metres. Both completed-works interactions are instrumented year by year. Bars are 90 percent (thick) and 95 percent (thin) confidence intervals. Firm, year, and province-by-year fixed effects; standard errors clustered at the firm level. The vertical dashed line marks the last pre-flood year.

firm had €0.89 million of completed works within 250 metres. At that rainfall and with no protection nearby, Column 1 implies a revenue loss of €287,780, or 12.9 percent of mean 2022 sales. The average stock offsets €205,743 of it, leaving €82,037, or 3.7 percent of sales. Protection returned roughly seven tenths of what the rain took.

$$\Delta\hat{Y} = \bar{R}\,(\underbrace{-129{,}717}_{\text{rainfall loss}} + \underbrace{104{,}338}_{\text{attenuation}} \times \bar{A}) = -82{,}037 \text{ euros.}$$

Turning this into a return on public spending requires two adjustments. The first concerns the unit. Multiplying the per-firm coefficient by a count of firms would be wrong, because Adapt_i is an inverse-distance weighted average of spending on shared watercourse segments rather than money spent on firm i: summing it over the estimation sample gives €13.8 billion against €465 million actually spent on the segments involved, each counted once, so one euro enters the treatment of roughly thirty firms. The return to a euro is the sum of the effects it produces across all of them. Summing the predicted effects $\beta_1 R_i \text{Adapt}_i$ across the 15,540 firms gives €2.69 billion of revenue preserved in 2023.

The second adjustment concerns the measure. Revenue is not the right object for this comparison. Of the €100,713 of revenue that each €1 million of completed works preserves per 10 mm of rainfall, €67,574 is cost of goods sold (Table A.31): inputs the firm would have bought to produce, and which are not lost when it stops. Revenue preserved therefore overstates the resources the works saved, and summed across firms it counts intermediate goods more than once. I measure the return in value added, the income a firm generates, which is the quantity comparable with the cost of the works and with the effects on GDP estimated by Ficarra and Mari (2025). Value added is reported for about half of the firms, mostly the larger ones; on those firms each euro of revenue preserved carries 37 cents of value added (Table A.33), a ratio that is the same in both samples of that table.

On this basis, an illustrative calculation suggests that each €1 million of completed preventive works preserved about €2.1 million of value added among the firms within 250 metres in the year of the flood, with a 95 percent interval from €1.0 to €3.2 million that reflects the imprecision of β_1[33]. Taken at face value, a single event would thus have returned about twice what the works cost to build. The figure is conditional on an event of this severity, and I do not convert it into an annual return, which would require the future frequency of such events and the response to smaller ones, neither of which this design identifies. Within the event it is a floor: it counts one year of one storm, and it counts only the firms in the estimation sample, while the same works protect households, farmland and public infrastructure whose cost sits in the denominator.

Robustness. I check the protective interaction along three dimensions, reported in full in Appendix 3.2: the treatment of the errors, the fixed effects and the measure of adaptation.

First, inference on the interaction is robust to the assumed correlation of the errors. Clustering

[33]The cost is the completed spending recorded on the segments, each segment counted once. The registry divides the cost of a project across all the municipalities it lists, including some that the watercourse does not cross, so part of the cost of a project is not placed on any segment. Using instead the full cost of the 1,060 projects whose spending reaches these segments, €503 million, gives €1.95 million.

by grid cell rather than by firm raises the standard error on the interaction from €28,366 to €50,670. Conley standard errors with cut-offs of 2, 5 and 10 km are almost identical to the cell-clustered ones, which indicates that the spatial dependence in the errors is local and already captured by the cell. The interaction remains significant at the 5 percent level under every estimator, although under cell clustering the first-stage F falls to about 11. The rainfall term is significant at the 10 percent level under firm clustering only.

Second, the interaction does not depend on the fixed effects. Adding sector-by-year fixed effects lowers it to €77,368, and adding them together with province-by-year effects to €67,513; both estimates remain significant. The most demanding specification replaces these with grid-cell-by-year fixed effects. Because rainfall is constant within a cell, Post × Rain is absorbed, and the interaction is identified only by comparing firms in the same cell and year that differ in the completed protection nearby. Any shock common to the cell, the storm included, is differenced out, and the estimate rises to €108,957. What this specification cannot exclude, and what the instrument addresses, is that within a cell protection was completed near firms that would have fared differently in any case.

Third, the interaction does not depend on how adaptation is measured. Replacing the 250-metre index with the works on the single nearest watercourse, or restricting it to rivers or to canals, gives interactions between €82,547 and €104,338. All are significant, the canal-only estimate at the 10 percent level.

5.3.1 Spatial Spillovers Along the River Network

Rivers move water across space, so the rain that damages a firm need not have fallen on it. Rainfall upstream increases the volume arriving downstream; rainfall downstream, by contrast, cannot reach a firm located above it. This asymmetry is useful, because it supplies a prediction the design must satisfy: upstream rainfall should damage, downstream rainfall should not.

Table 6 holds the specification of Column 1 of Table 5 fixed and changes only the spatial origin of the rainfall shock. The adaptation measure is in every column the firm's own stock of completed works within 250 metres, so the question the table answers is whether local works protect against non-local water. Upstream and downstream rainfall are averages over the municipalities ranked above and below the firm on a shared watercourse, using the flow-direction ordering of Section 1.6.3.

The prediction holds. Rainfall upstream damages firms by €138,447 per 10 mm, an effect indistinguishable in size from the €129,717 of local rainfall; rainfall downstream has no effect at all (−33,606, $p = 0.56$). Water that cannot arrive does not damage.

Protection, however, does not travel with the water. Local works attenuate local rainfall by €104,338 ($p < 0.001$), but their estimated attenuation of upstream rainfall is €41,329 and is not distinguishable from zero ($p = 0.14$). The natural reading is that protection is local in a strong sense: embankments and retention works near a firm hold back the water that falls nearby, and do much less against a volume that arrives already channelled from above.

The asymmetry matters for policy. If local works do not offset water originating upstream, the protection a municipality obtains from its own investment is bounded by what its upstream neighbours

have built. Adaptation decided at the local level will then be undersupplied relative to the basin optimum, since each authority internalises the protection it provides to itself and not the protection it fails to provide downstream.

5.3.2 Remedial versus Preventive Adaptation

The classification distinguishes ex-ante from ex-post works, defined by the purpose that the implementing authority declares in the project description: preventive works are those intended to reduce future flood risk, remedial works those intended to repair the damage of an earlier event (Section 3.1). The estimates so far use preventive works alone, completed before the May 2023 floods. One objection, however, is that they capture the effect of having any hydraulic work nearby, rather than of works designed to prevent. If so, the distinction would be immaterial, and remedial works completed before the flood should protect as much as preventive ones.

A first, descriptive piece of evidence against this hypothesis is that the two categories respond to different things. Across municipalities, remedial registrations rise in the quarter after extreme rainfall, while preventive registrations show no response under any definition of the shock (Appendix 1.3.3): repair follows realised damage, and prevention follows funding rules.

Both categories are recorded in the same registry and placed on the same network, so for remedial works I construct the same index of completed spending within 250 metres and the same leave-one-out instrument, computed on remedial projects alone. As for preventive works, only remedial works completed before May 2023 enter the index, so that both stocks measure what stood when the flood arrived; repairs undertaken after the May 2023 floods are excluded.

The regressions answer the objection (Table 7). Protection is found for preventive works only. When the two types enter jointly (Column 3), the preventive interaction is €91,341 ($p = 0.004$), while the remedial interaction is −€96,234 and indistinguishable from zero ($p = 0.49$). Estimated separately, the preventive interaction is €97,690 ($p = 0.002$) and the remedial one −€198,182 ($p = 0.07$). The first-stage F statistics are 164 and 163 in the joint specification, so neither margin is weakly identified.

Two caveats apply. First, the remedial coefficient is far less precise: its standard error is €139,327, against €31,927 for preventive works, because the remedial stock is smaller (€196,000 on average, against €829,000) and positive for about half of the firms rather than two thirds. Second, the negative remedial coefficient should not be read as damage caused by repair. Remedial works are built where earlier floods struck, as the response of their registrations to extreme rainfall shows, and the instrument shifts the timing of their completion, not the choice of their location.

The protective effect is therefore confined to preventive works, with no comparable effect for remedial ones. Because both enter the same regression on the same firms, this asymmetry cannot be attributed to differences in sample, rainfall measure or index construction.

A further margin is not analysed here: whether prevention reduces the relief that the public sector pays once an event has occurred. The classification also identifies the compensation transfers (*ristori*) paid to individual firms after the May 2023 floods, so it is possible to test whether, for the same rainfall, firms with more completed protection nearby received less relief. This analysis is in progress.

Table 6: Local Protection against Local, Upstream and Downstream Rainfall

	(1) Local rain	(2) Upstream rain	(3) Downstream rain
Post × Rain	-129,717*	-138,447*	-33,606
	(71,574)	(79,892)	(57,309)
Post × Rain × Adaptation	104,338***	41,329	16,501
	(28,366)	(28,222)	(34,633)
First-stage F	310	1172	2028
Observations	81,492	81,492	81,492

Notes: Each column reports the specification of Column 1 of Table 5 and differs only in the spatial origin of the rainfall shock: precipitation in the firm's own ERG5 grid cell, the average across municipalities located upstream of the firm on a shared watercourse, and the average across those located downstream. Adapt_i is in every column the firm's own stock of completed works within 250 metres, instrumented as in Table 5; planned-works controls are included throughout. The sample is 2018–2023. Firm, year and province-by-year fixed effects. Standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 7: Preventive and Remedial Adaptation

	(1) Ex-ante only	(2) Remedial only	(3) Both
Post × Rain	-125,228*	-68,143	-93,138
	(75,727)	(74,059)	(75,424)
Post × Rain × Ex-ante adaptation	97,690***		91,341***
	(31,976)		(31,927)
Post × Rain × Remedial adaptation		-198,182*	-96,234
		(110,496)	(139,327)
First-stage F, ex-ante	317		164
First-stage F, remedial		249	163
Observations	73,059	73,059	73,059

Notes: Estimates of equation (5) by two-stage least squares. The dependent variable is annual revenue in euros. Ex-ante adaptation is the inverse-distance-weighted stock of completed preventive works on watercourse segments within 250 metres of firm i by May 2023, in €1 million units; remedial adaptation is the same index computed on works classified as restorative. Each margin is instrumented with its own leave-one-out predicted completion index, constructed from the projects of that type. Planned-works controls of the corresponding type are included in every column. The sample is 2018–2023. Firm, year and province-by-year fixed effects. The first-stage F is the Wald statistic on the excluded instruments in the first stage for the interaction named in the row. Standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5.4 Channels

The main result establishes that completed ex-ante adaptation attenuates rainfall-driven revenue losses. Two questions follow naturally: what does adaptation protect (the income statement, the asset stock, or both) and why does physical flood protection translate into firm-level revenue protection? This section addresses the first question directly, using a rich set of alternative outcomes estimated under the same 2SLS specification as the main result. On the second question, the data offer suggestive evidence but do not permit a fully identified answer.

Tables A.31 and A.32 in Appendix 4 report the results. The pattern divides cleanly along the accounts. Every item on the income statement moves, and none on the balance sheet does. Each €1 million of completed adaptation, per 10 mm of rainfall, preserves €100,713 of revenue ($p < 0.001$), €67,574 of cost of goods sold ($p = 0.003$), €33,153 of gross profit ($p = 0.013$) and €12,237 of operating profit ($p = 0.013$). Value added, which is reported for only about half of the firms and is therefore estimated on a different sample, rises by €30,586 ($p = 0.004$; Table A.33). Tangible assets, total assets, current assets, the solvency ratio and the current ratio are all indistinguishable from zero[34].

This split is consistent with protection preserving the operation of firms rather than their assets. Firms near completed works sold more, bought and processed more inputs and earned more on the difference, but they did not end the year with more machinery, inventory or liquid assets, or with a stronger balance sheet, than comparable firms without protection. One interpretation is that protection allowed firms to keep operating during the flood. Operation, unlike the asset stock, is a flow that ex-post compensation cannot restore: transfers and subsidised credit can rebuild an asset after the interruption, but not output that was not produced. The annual accounts do not allow this interpretation to be tested further[35].

5.5 Limitations and External Validity

The estimates have a defined scope. They apply to standalone firms, the only firms that can be assigned a single location and therefore a single exposure; firms in multi-establishment groups, which can shift production across plants, may lose less from a flood and gain less from protection. They also apply to a particular hazard and a particular kind of protection. The May 2023 floods were riverine floods triggered by intense rainfall, with a sharply bounded footprint, and the works studied here sit along the channels that carry the water. Hazards with a different footprint, such as storm surges along a coast, wind along the track of a storm or heat across a whole area, relate differently to local protection, and slow-onset floods leave firms time to move inventory and equipment.

Within this scope, several features of the design make the estimates conservative. They measure

[34] The wage bill (€3,502, $p = 0.42$) and employment ($p = 0.62$) are also indistinguishable from zero, but the labour variables are recorded for roughly half the sample in Orbis, so these nulls are uninformative; an extension using InfoCamere records is in progress.

[35] The physical mechanisms are not observed. Annual accounts do not record whether a firm's premises were inundated, for how long it stopped, or whether it lost access to roads, power or suppliers, so the estimates cannot separate protection of the site itself from protection of the infrastructure around it. Linking firms to their flood status, to road closures and to the duration of inundation is left to future work.

the benefit to firms alone, while the same works also protect households, farmland and public infrastructure. They are a return per euro of declared adaptation spending, which includes ordinary capital replaced alongside the protective component of each work (Section 3.1), so the return per euro of adaptive capital is higher. And they are averages: if, as Tarsia (2026) shows for weather shocks, losses fall disproportionately on low-productivity firms with little capacity to self-insure, protection is worth more than the average effect to the firms that depend on it most.

The post-flood period covers two years. It is sufficient to measure the immediate loss and its partial recovery, but not the medium-run consequences for employment, investment or survival, which over longer horizons can include out-migration and lower property values (Boustan et al., 2020).

Finally, transferring the estimates requires care. The instrument rests on the different speeds of Italian funding pipelines, which other settings may not share, whereas the measurement of adaptation needs only project descriptions and can be applied elsewhere. Applying the coefficient to a warming climate requires the future frequency of comparable events, which climate projections rather than this design must supply (Lemoine, 2018; Carleton et al., 2024); it also assumes returns that are linear in the stock of protection, although works with design thresholds may yield more when correctly sized and less once thresholds are exceeded; and the two-stage least squares estimand is local to works whose completion depended on the speed of their funding pipeline. Within these bounds, the design identifies the losses that completed preventive works avoided for firms in a realised event, which is the quantity on which decisions about building and completing such works depend.

6 Conclusion

As floods become more frequent, governments are expected to commit more resources to adaptation, and the relevant question concerns not only the amount of public spending but also its composition. This paper provides firm-level evidence on that question. Using a procedure that classifies public investment projects from their descriptions, it distinguishes preventive from remedial works and compares firms exposed to the same planned protection but to a different share of completed works at the time of the May 2023 floods.

Firms located within the flood perimeter experienced a revenue decline of approximately 4 percent in the year of the event, with losses confined to a few hundred metres from the inundated area. Completed preventive works substantially attenuated these losses: each €1 million of works within 250 metres of a firm preserved approximately €104,000 of revenue per additional 10 mm of extreme rainfall. The effect is concentrated on the income statement rather than on the balance sheet, which is consistent with protection preserving the continuity of production. It is specific to preventive works, as remedial works show no comparable effect, and it does not extend to rainfall originating upstream.

These findings are relevant to the allocation of public resources in the aftermath of a disaster. In Emilia-Romagna, registrations of remedial projects increased from approximately 1,100 in 2022 to more than 5,000 in 2023 and 6,000 in 2024 (Figure 3), as the public response necessarily followed the damage. The evidence presented here indicates, however, that protection derives from preventive

works and, in particular, from the completion of works already planned. Prior to May 2023, planned investment had been directed towards the most exposed watercourses, yet on most of them the works remained incomplete at the time of the flood. Moreover, since local works do not attenuate the effect of upstream rainfall, the benefits of completion are likely to be greater when investment decisions are coordinated at the level of the river basin rather than at the municipal level.

Several questions remain open. Prevention may reduce not only the losses incurred by firms but also the relief disbursed after an event; the compensation transfers recorded in the registry allow a test of whether firms with greater completed protection received less relief. Linking firms to their inundation status, to road closures and to the duration of flooding would clarify the mechanisms through which protection operates. Finally, because the estimates hold the location of firms fixed, they are silent on whether protection attracts economic activity to protected areas, as Hsiao (2023) shows for a planned sea wall in Jakarta, and on whether the resulting concentration of exposure increases the damage associated with an eventual failure (Kousky et al., 2006; Di Baldassarre et al., 2018). Pending these extensions, the estimates should be interpreted as returns to a form of adaptation that protects existing locations of economic activity, rather than to forms that would relocate or transform it.

More broadly, as floods become more frequent and adaptation spending grows, evidence on what public prevention returns to economic activity, and to the firms that carry it out, becomes increasingly necessary. Such returns are difficult to observe, because they materialise only when a hazard occurs and only as losses that do not take place, which may help explain why prevention attracts less political reward than relief (Healy and Malhotra, 2009). A further question is how public protection interacts with the adaptation that firms undertake themselves. If protection weakens the incentive to invest in private resilience, as Fried (2022) shows for disaster aid and Zhao et al. (2026) find for public flood protection in China, part of its return may be offset over time by lower private adaptation. Measuring this interaction, which requires observing both public works and the investments of the firms they protect, is a natural next step.

References

ANIA. 2024. "Rischi Catastrofali: Esposizione e Coperture Assicurative in Italia."Technical report, Associazione Nazionale fra le Imprese Assicuratrici, Rome, Estimates of catastrophe exposure and insurance penetration for Italian households and firms, presented 2 July 2024.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131 (4): 1593–1636, https://doi.org/10.1093/qje/qjw024.

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review* 99 (4): 1278–1308. 10.1257/aer.99.4.1278.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro. 2016. "Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship over the Twentieth Century." *Journal of Political Economy* 124 (1): 105–159, https://doi.org/10.1086/684582.

Barrot, Jean-Noël, and Julien Sauvagnat. 2016. "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks." *Quarterly Journal of Economics* 131 (3): 1543–1592. 10.1093/qje/qjw018.

Basker, Emek, and Javier Miranda. 2018. "Taken by Storm: Business Financing and Survival in the Aftermath of Hurricane Katrina." *Journal of Economic Geography* 18 (6): 1285–1313. 10.1093/jeg/lbx023.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi. 2023. "Individuals and Organizations as Sources of State Effectiveness." *American Economic Review* 113 (8): 2121–2167. 10.1257/aer.20191598.

Bosio, Erica, Simeon Djankov, Edward Glaeser, and Andrei Shleifer. 2022. "Public Procurement in Law and Practice." *American Economic Review* 112 (4): 1091–1117. 10.1257/aer.20200738.

Boustan, Leah Platt, Matthew E. Kahn, Paul W. Rhode, and Maria Lucia Yanguas. 2020. "The Effect of Natural Disasters on Economic Activity in US Counties: A Century of Data." *Journal of Urban Economics* 118 103257, https://doi.org/10.1016/j.jue.2020.103257.

Bradt, Jacob T., and Joseph E. Aldy. 2026. "Private Benefits from Public Investment in Climate Adaptation and Resilience." January, https://jacobbradt.com/assets/pdf/papers/BradtAldy_ClimateAdaptation.pdf, Revise and resubmit, *Review of Economic Studies*. Earlier version: NBER Working Paper 33633.

Brath, Armando, Nicola Casagli, Marco Marani, Paola Mercogliano, and Renzo Motta. 2023. "Rapporto della Commissione tecnico-scientifica sugli eventi meteorologici estremi del maggio 2023 in Emilia-Romagna."Technical report, Regione Emilia-Romagna, Bologna, Commissione istituita con delibera di Giunta regionale n. 984/2023. https://www.regione.emilia-romagna.it/alluvione/rapporto-della-commissione-tecnico-scientifica.

Burke, Marshall, and Kyle Emerick. 2016. "Adaptation to Climate Change: Evidence from US Agriculture." *American Economic Journal: Economic Policy* 8 (3): 106–140, https://doi.org/10.1257/pol.20130025.

Carleton, Tamma, Esther Duflo, Kelsey Jack, and Guglielmo Zappalà. 2024. "Adaptation to Climate Change." Working Paper 33264, National Bureau of Economic Research, Also forthcoming as a chapter in the *Handbook of the Economics of Climate Change*, Vol. 1 (Elsevier, 2024). https://www.nber.org/papers/w33264.

Carleton, Tamma, Amir Jina, Michael Delgado et al. 2022. "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits." *Quarterly Journal of Economics* 137 (4): 2037–2105, https://doi.org/10.1093/qje/qjac020.

Carvalho, Vasco M., Makoto Nirei, Yukiko U. Saito, and Alireza Tahbaz-Salehi. 2021. "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake." *Quarterly Journal of Economics* 136 (2): 1255–1321. 10.1093/qje/qjaa044.

Catalano, Michele, Lorenzo Forni, and Emilia Pezzolla. 2020. "Climate-Change Adaptation: The Role of Fiscal Policy." *Resource and Energy Economics* 59 101111. 10.1016/j.reseneeco.2019.07.005.

Cavallo, Eduardo, and Ilan Noy. 2011. "Natural Disasters and the Economy — A Survey." *International Review of Environmental and Resource Economics* 5 (1): 63–102. 10.1561/101.00000039.

Chen, Jiafeng, and Jonathan Roth. 2024. "Logs with Zeros? Some Problems and Solutions." *Quarterly Journal of Economics* 139 (2): 891–936. 10.1093/qje/qjad054.

Collier, Benjamin L., Sabrina T. Howell, and Lea Rendell. 2024. "After the Storm: How Emergency Liquidity Helps Small Businesses Following Natural Disasters." Working Paper 32326, National Bureau of Economic Research, https://www.nber.org/papers/w32326.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45. 10.1016/S0304-4076(98)00084-0.

Coviello, Decio, and Stefano Gagliarducci. 2017. "Tenure in Office and Public Procurement." *American Economic Journal: Economic Policy* 9 (3): 59–105. 10.1257/pol.20150426.

Coviello, Decio, Luigi Moretti, Giancarlo Spagnolo, and Paola Valbonesi. 2018. "Court Efficiency and Procurement Performance." *Scandinavian Journal of Economics* 120 (3): 826–858. 10.1111/sjoe.12225.

Decarolis, Francesco, Leonardo M. Giuffrida, Elisabetta Iossa, Vincenzo Mollisi, and Giancarlo Spagnolo. 2020. "Bureaucratic Competence and Procurement Outcomes." *Journal of Law, Economics, and Organization* 36 (3): 537–597. 10.1093/jleo/ewaa004.

Deryugina, Tatyana. 2017. "The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." *American Economic Journal: Economic Policy* 9 (3): 168–198, https://doi.org/10.1257/pol.20140296.

Di Baldassarre, Giuliano, Heidi Kreibich, Sergiy Vorogushyn et al. 2018. "Hess Opinions: An Interdisciplinary Research Agenda to Explore the Unintended Consequences of Structural Flood Protection." *Hydrology and Earth System Sciences* 22 (11): 5629–5637. 10.5194/hess-22-5629-2018.

Dottori, Francesco, Lorenzo Mentaschi, Alessandra Bianchi, Lorenzo Alfieri, and Luc Feyen. 2020. "Adapting to Rising River Flood Risk in the EU under Climate Change." JRC PESETA IV Project – Task 5 EUR 29955 EN,

Publications Office of the European Union, Luxembourg. 10.2760/14505, https://publications.jrc.ec.europa.eu/repository/handle/JRC118425.

Egami, Naoki, Musashi Hinck, Brandon M. Stewart, and Hanying Wei. 2023. "Using Imperfect Surrogates for Downstream Inference: Design-Based Supervised Learning for Social Science Applications of Large Language Models." In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, https://arxiv.org/abs/2306.04746.

Erda, Tarikua. 2026. "Cleansing Floods? Creative Destruction and Government Spending after Disasters." Working Paper CES-26-52, U.S. Census Bureau, Center for Economic Studies, Earlier versions circulated as "Disasters, Capital, and Productivity". https://www.census.gov/library/working-papers/2026/adrm/CES-WP-26-52.html.

European Environment Agency. 2025. "Economic Losses from Weather- and Climate-Related Extremes in Europe."Technical report, European Environment Agency, Copenhagen, Indicator CLIM039, published 14 October 2025; data cover 1980–2024. https://www.eea.europa.eu/en/analysis/indicators/economic-losses-from-climate-related.

European Parliament. 2024. "Flooding in Emilia-Romagna in May 2023 and Management of Hydrogeological Instability." Parliamentary question E-001407/2024, States direct damages "totalling over EUR 8.5 billion". https://www.europarl.europa.eu/doceo/document/E-10-2024-001407_EN.html.

Ficarra, Matteo, and Rebecca Mari. 2025. "Weathering the Storm: Sectoral Economic and Inflationary Effects of Floods and the Role of Adaptation." Staff Working Paper 1120, Bank of England.

Fried, Stephie. 2022. "Seawalls and Stilts: A Quantitative Macro Study of Climate Adaptation." *Review of Economic Studies* 89 (6): 3303–3344, https://doi.org/10.1093/restud/rdab099.

Gallagher, Justin. 2014. "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States." *American Economic Journal: Applied Economics* 6 (3): 206–233, https://doi.org/10.1257/app.6.3.206.

Garrett, Thomas A., and Russell S. Sobel. 2003. "The Political Economy of FEMA Disaster Payments." *Economic Inquiry* 41 (3): 496–509, https://doi.org/10.1093/ei/cbg023.

Gentili, Elena. 2026. "How to Rapidly Assess the Economic Consequences of Floods with Public Data Sources? Evidence from Northern Italy." *Journal of Environmental Economics and Management* 140 103411. 10.1016/j.jeem.2026.103411, https://doi.org/10.1016/j.jeem.2026.103411.

Geyiktepe, Mert, and Dani Rodrik. 2026. "The China Backlash: Quantifying the Narrative Discourse on China." NBER Working Paper 35539, National Bureau of Economic Research, http://www.nber.org/papers/w35539.

Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli. 2023. "ChatGPT Outperforms Crowd Workers for Text-Annotation Tasks." *Proceedings of the National Academy of Sciences* 120 (30): e2305016120, https://doi.org/10.1073/pnas.2305016120.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun. 2019. "Firm-Level Political Risk: Measurement and Effects." *Quarterly Journal of Economics* 134 (4): 2135–2202, https://doi.org/10.1093/qje/qjz021.

Healy, Andrew, and Neil Malhotra. 2009. "Myopic Voters and Natural Disaster Policy." *American Political Science Review* 103 (3): 387–406, https://doi.org/10.1017/S0003055409990104.

Hosono, Kaoru, Daisuke Miyakawa, Taisuke Uchino, Makoto Hazama, Arito Ono, Hirofumi Uchida, and Iichiro Uesugi. 2016. "Natural Disasters, Damage to Banks, and Firm Investment." *International Economic Review* 57 (4): 1335–1370. 10.1111/iere.12200.

Hsiao, Allan. 2023. "Sea Level Rise and Urban Adaptation in Jakarta." working paper, Stanford University, https://allanhsiao.com/files/Hsiao_jakarta.pdf.

IPCC. 2021. "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change."Technical report, Cambridge University Press, Cambridge, UK and New York, NY, USA. 10.1017/9781009157896.

ISPRA. 2024. "Hydrogeological Instability in Italy: Hazards and Risk Indicators, 2024 Edition."Technical report, Istituto Superiore per la Protezione e la Ricerca Ambientale, Rome, https://www.isprambiente.gov.it/en/publications/reports/hydrogeological-instability-in-italy-hazards-and-risk-indicators-2024-edition.

ISTAT. 2024. "Il Benessere Equo e Sostenibile dei Territori: Emilia-Romagna."Technical report, Istituto Nazionale di Statistica, Rome, Resident population 4,451,938 at 1 January 2024; GDP per capita €40,030 (2022). https://www.istat.it/wp-content/uploads/2024/12/Emilia-Romagna_BesT_2024.pdf.

ISTAT. 2025. "Conti economici territoriali — Anni 2021–2023."Technical report, Istituto Nazionale di Statistica, https://www.istat.it/comunicato-stampa/conti-economici-territoriali-anni-2021-2023/.

IVASS. 2026. "Monitoraggio delle Polizze contro i Rischi Catastrofali."Technical report, Istituto per la Vigilanza sulle Assicurazioni, Rome, Monitoring of catastrophe policies taken out by firms following the obligation introduced by the 2024 Budget Law.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–182, https://doi.org/10.1257/0002828053828518.

Killick, Rebecca, Paul Fearnhead, and Idris A. Eckley. 2012. "Optimal Detection of Changepoints with a Linear Computational Cost." *Journal of the American Statistical Association* 107 (500): 1590–1598, https://doi.org/10.1080/01621459.2012.737745.

Kocornik-Mina, Adriana, Thomas K. J. McDermott, Guy Michaels, and Ferdinand Rauch. 2020. "Flooded Cities." *American Economic Journal: Applied Economics* 12 (2): 35–66, https://doi.org/10.1257/app.20170066.

Kousky, Carolyn, Erzo F. P. Luttmer, and Richard J. Zeckhauser. 2006. "Private Investment and Government Protection." *Journal of Risk and Uncertainty* 33 (1–2): 73–100. 10.1007/s11166-006-0172-y.

Lehner, Bernhard, and Günther Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems." *Hydrological Processes* 27 (15): 2171–2186, HydroRIVERS database. https://doi.org/10.1002/hyp.9740.

Lemoine, Derek. 2018. "Estimating the Consequences of Climate Change from Variation in Weather." NBER Working Paper 25008, National Bureau of Economic Research.

Loberto, Michele, and Riccardo Russo. 2024. "The Exposure of Italian Manufacturing Firms to Hydrogeological Risk."Technical report, Bank of Italy, Economic Research and International Relations Area.

Lodi, Chiara, Giovanni Marin, and Marco Modica. 2023. "The Public Finance Response to Floods of Local Governments in Italy." *Journal of Environmental Management* 332 117352, https://doi.org/10.1016/j.jenvman.2023.117352.

Loughran, Tim, and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66 (1): 35–65. 10.1111/j.1540-6261.2010.01622.x.

Ludwig, Jens, Sendhil Mullainathan, and Ashesh Rambachan. 2024. "Large Language Models: An Applied Econometric Framework." working paper, National Bureau of Economic Research, https://arxiv.org/abs/2412.07031. Published as NBER Working Paper No. 33344 (2025).

Masiero, Giuliano, and Michael Santarossa. 2020. "Earthquakes, Grants, and Public Expenditure: How Municipalities Respond to Natural Disasters." *Journal of Regional Science* 60 (3): 481–516, https://doi.org/10.1111/jors.12462.

Michaelowa, Axel, and Katharina Michaelowa. 2011. "Coding Error or Statistical Embellishment? The Political Economy of Reporting Climate Aid." *World Development* 39 (11): 2010–2020. 10.1016/j.worlddev.2011.07.020.

Ministero dell'Ambiente e della Sicurezza Energetica. 2023. "Piano Nazionale di Adattamento ai Cambiamenti Climatici."Technical report, Ministero dell'Ambiente e della Sicurezza Energetica, Rome, Approved by ministerial decree n. 434 of 21 December 2023. https://www.mase.gov.it/portale/documents/d/guest/pnacc_documento_di_piano-pdf.

Noth, Felix, and Oliver Rehbein. 2019. "Badly Hurt? Natural Disasters and Direct Firm Effects." *Finance Research Letters* 28. 10.1016/j.frl.2018.05.009, https://doi.org/10.1016/j.frl.2018.05.009.

Olden, Andreas, and Jarle Møen. 2022. "The Triple Difference Estimator." *Econometrics Journal* 25 (3): 531–553. 10.1093/ectj/utac010.

OpenAI. 2024. "GPT-4o Mini: Advancing Cost-Efficient Intelligence."Technical report, OpenAI, https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal* 128 (608): 413–446. 10.1111/ecoj.12418.

Regione Emilia-Romagna. 2025. "Perimetrazione delle aree allagate, eventi del 2–4 e del 16–17 maggio 2023 (versione 6)."Technical report, Geoportale Emilia-Romagna, Pubblicata con Decreto del Segretario Generale (D.S.G.) n. 45/2025. https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/informazioni-geoscientifiche/zone-a-rischio-naturale/aree-allagate.

Rentschler, Jun, Paolo Avner, Mattia Marconcini et al. 2023. "Global Evidence of Rapid Urban Growth in Flood Zones Since 1985." *Nature* 622 87–92. 10.1038/s41586-023-06468-9.

Ricciotti, Ludovica, and Alessio Pollice. 2024. "An Integrated Georeferenced Dataset of Public Investments for Soil Defence in Italy." *Data in Brief* 57 110909, https://doi.org/10.1016/j.dib.2024.110909.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. 2023. "Firm-Level Climate Change Exposure." *The Journal of Finance* 78 (3): 1449–1498. 10.1111/jofi.13219.

Van der Straten, Yasmine. 2023. "Flooded House or Underwater Mortgage? The Implications of Climate Change and Adaptation on Housing, Income, and Wealth." *Tinbergen Institute Discussion Paper TI 2023-014/IV*.

Tarsia, Romano. 2026. "Heterogeneous Effects of Weather Shocks on Firm Economic Performance." *Journal of Environmental Economics and Management* 140 103393. 10.1016/j.jeem.2026.103393.

Taylor, Charles A., and Hannah Druckenmiller. 2022. "Wetlands, Flooding, and the Clean Water Act." *American Economic Review* 112 (4): 1334–1363.

Toetzke, Malte, Anna Stünzi, and Florian Egli. 2022. "Consistent and Replicable Estimation of Bilateral Climate Finance." *Nature Climate Change* 12 (10): 897–900. 10.1038/s41558-022-01482-7.

Zhao, Yu, Yichen Yang, and Ning Zhang. 2026. "Trading Resilience for Growth: Evidence from Public Flood Protection in China." Working Paper 6996377, SSRN. 10.2139/ssrn.6996377, https://ssrn.com/abstract=6996377.

Appendix

1 Data

This appendix documents how each input to the analysis is built. Most of it concerns the measure of public adaptation, which has to be constructed rather than collected.

Public adaptation is not recorded anywhere as such. What exists is an administrative census of public investment decisions, in which hydraulic works sit undifferentiated among schools, roads and farm subsidies, and whose free-text descriptions state what each project was meant to do. Sections 1.1 to 1.4 follow the measure from that source to its output: the registry and the projects that enter the classification, the pipeline that reads the descriptions, the exercises that establish whether its labels can be trusted, and what the classified series say about adaptation in Emilia-Romagna. The remaining sections cover the firm panel, the hydrographic network and the rainfall data.

1.1 The Registry and the Classification Sample

Evaluating public adaptation requires observing, for each intervention, what was decided, where, when, at what cost, and with what intent. No statistical source is built for this purpose. Italy, however, possesses an administrative one: OpenCUP, the national registry of public investment projects. This section describes the registry, argues that its design makes it unusually well suited to the measurement problem, and details the construction of the analysis sample: every project registered in Emilia-Romagna from the registry's inception in January 2003 through July 2025.

1.1.1 OpenCUP: A Census of Public Investment Decisions

Since 2003, every project involving public investment in Italy, whatever its size, its sector, or the level of government undertaking it, must obtain a *Codice Unico di Progetto* (CUP), a 15-character alphanumeric code that identifies the project uniquely across every administration involved and over its entire life cycle (Law 3/2003). The obligation has teeth: since 2020, an administrative act that finances or authorises a public investment project without reporting its CUP is null (Decree-Law 76/2020, Art. 41). Coverage of the resulting registry is therefore not a matter of reporting diligence but of legal validity, a project absent from the registry is, in the relevant sense, a project that cannot be paid for. The registry is maintained by the Department for Programming and Coordination of Economic Policy (DIPE) at the Presidency of the Council of Ministers and published in full as open data. It covers more than 10.8 million projects registered between 2003 and 2025 by over 36,000 public entities, with a total projected cost of approximately €3.96 trillion[36]. To my knowledge, this paper is

[36] I use the September 2025 release. The data are updated monthly and can be downloaded in bulk from the OpenCUP portal, where individual projects can also be looked up by CUP.

Figure A.1: A project record as displayed in the OpenCUP portal.

Notes: Embankment-stability works on the Panaro river (OpenCUP project link), classified by the pipeline as preparedness (ex-ante): although labelled "urgent," the works reinforce the embankment against seepage in anticipation of future flood events, rather than repairing damage from a realised one. Text translated from Italian.

the first to use OpenCUP as a primary data source in economic research[37].

Two features of the record matter for the exercise here. The first is geography: each record reports the municipalities in which the project is implemented, which is the resolution at which flood exposure varies. The second, central to this paper, is language. Alongside structured fields, expected cost and public funding, implementing body, status, and a sector–sub-sector code assigned by the proposing administration at registration, each record carries free-text fields in which that same administration describes, in Italian, what is being built or repaired, on which structure, and to what end. It is this free-text layer, not the structured fields, that makes it possible to pick adaptation-relevant investment out of the far larger universe of environmental and infrastructure policy the registry otherwise records.

Figure A.1 shows an example, an embankment-stability intervention on the Panaro river, and illustrates why the textual layer of the record, not the structured one, carries the information this paper needs. The sector code places the project, at best, in a broad administrative domain. The description reveals its object, the stability of a river embankment, and its logic, reinforcement undertaken to prevent a future failure, rather than to repair a realised one. The preparedness/remediation distinction at the core of HYDROADAPT lives in this text. At the registry's scale, however, text is precisely what administrative data systems cannot read: with millions of entries of heterogeneous administrative prose, manual classification is prohibitively costly, and keyword rules founder on the distinction that matters most here: reinforcement and restoration may describe the same levee, before and after the flood. Turning this text into data is the task of the classification pipeline described in the next section.

[37]The CUP has so far entered empirical work only as a record-linkage key: Ricciotti and Pollice (2024) use it to merge OpenBDAP, OpenCoesione, and ReNDiS records on soil-defence investments. The literature on Italian public works relies instead on procurement data from the national contracts database maintained by the anticorruption authority (ANAC), which records contracts and their execution rather than investment decisions.

Table A.1: Sample Construction

Step	Projects	Δ
Emilia-Romagna extract	1,055,466	
Valid registration dates, 2003–2025	1,052,460	−3,006
– crop and livestock insurance premia	771,875	−280,585
– photovoltaic installations	734,896	−36,979
– mountain-area compensatory payments	660,600	−74,296
– other constrained-area payments	632,234	−28,366
– internship allowances	618,600	−13,634
– other codified policy lines	600,823	−17,777
Final classification sample	**600,823**	

Notes: Counts refer to unique projects (CUPs); in the raw extract a CUP appears once for every municipality in which it is implemented, with an identical description, so classification operates on unique CUPs. "Other codified policy lines" comprises standardised entries such as post-earthquake residential reconstruction and tax-relief programmes.

1.1.2 The Classification Sample

The classification covers every project registered in Emilia-Romagna over the registry's full span, from January 2003 to July 2025. The window matters: the May 2023 floods fall late in it, so the classification can be tested against two decades of pre-event spending, including years in which climate adaptation had not yet entered public discourse and anticipatory behaviour is unlikely to confound the record. One caveat applies. Coverage is thinner in the registry's earliest years, when reporting practices were still consolidating[38].

Restricting the registry to Emilia-Romagna yields 1,055,466 projects, of which 1,052,460 have a valid CUP generation date between January 2003 and July 2025. Inspection of the project descriptions reveals several large groups of mechanically generated records with standardised text. These records refer to programmes outside the scope of public hydrogeological adaptation examined in this paper and contain no project-specific information requiring textual classification[39]. I remove these series before classification, reducing the final sample to 600,823 projects (Table A.1).

1.2 The Classification Pipeline

The pipeline maps the textual description of each project into one of three mutually exclusive categories: *unrelated*, *preparedness* (ex-ante), or *remediation* (ex-post). It is implemented through a Python script that calls OpenAI's GPT-4o-mini (OpenAI, 2024). The classification follows the taxonomy of adaptation formalised by Carleton et al. (2024): preparedness comprises forward-looking

[38] Enforcement tightened discontinuously with Law 136/2010 on the traceability of financial flows, in force from September 2010, which made the CUP mandatory on every payment related to public works. Registered volumes jump around that date across all project categories, consistent with a coverage effect rather than an investment shift; time-series comparisons that span 2010 should be interpreted accordingly.

[39] The excluded series include crop and livestock insurance premia, serial photovoltaic installations, compensatory payments in mountain and other constrained areas, internship allowances, post-earthquake residential reconstruction, and standard tax-relief measures. Their descriptions repeat fixed programme templates across large numbers of records rather than describing project-specific hydrogeological interventions.

Table A.2: Examples of Classification Output

CUP	Project description	Model response	Hydro	Ex-ante	Ex-post
F39D20000020002	Support for the emergence of vulnerabilities through strengthening of the asylum reception network. Training of public operators.	No	0	0	0
B13H20000300001	Works to improve the stability of embankments against seepage phenomena along the Panaro River.	Yes, Ex-Ante	1	1	0
B23H20000170001	Urgent interventions following the flood event of the Po River starting on 24 November 2019 (Western Emilia-Romagna).	Yes, Ex-Post	1	0	1

Notes: Descriptions translated from Italian for illustrative purposes. *Hydro*, *Ex-ante*, and *Ex-post* are binary indicators parsed from the model response.

measures undertaken to reduce future risk, such as flood defences, slope stabilisation, and drainage upgrades; remediation comprises reactive measures that repair damage and restore functionality after an event.

For each project, the input combines four descriptive fields from the OpenCUP record[40]. The concatenated text and the CUP identifier are inserted into a fixed prompt. The prompt jointly determines whether the project concerns hydrogeological risk and, conditional on relevance, whether the intervention is ex ante or ex post. Translated from Italian, it reads:

> *The intervention described as: '{description}' with the identifier (CUP): {cup}, is it an infrastructural investment aimed at addressing hydrogeological risk (e.g., landslides and floods)? If yes, specify whether it concerns 'Ex-Post' adaptation (reactive measures to repair hydrogeological damage caused by extreme events) or 'Ex-Ante' adaptation (proactive measures to mitigate future risks). Respond with 'No' if it is not related to hydrogeological risk; otherwise, reply with 'Yes, Ex-Post' or 'Yes, Ex-Ante'.*

The prompt instructs the model to return one of three responses: *No*, *Yes, Ex-Ante*, or *Yes, Ex-Post*. These responses are converted into three binary indicators using a deterministic text rule. A *No* response assigns zero to all three indicators; *Yes, Ex-Ante* assigns one to hydrogeological relevance and preparedness; and *Yes, Ex-Post* assigns one to hydrogeological relevance and remediation. The indicators are then merged with the project's structured metadata. Table A.2 reports three representative examples.

The implementation is designed around four choices: zero-shot classification, one project per call, a single-stage prompt, and explicit handling of repeated inputs and technical failures.

First, the model is queried zero-shot[41]. It receives no labelled examples and is neither fine-tuned nor

[40]The four fields are `natura_intervento` (nature of the intervention), `tipologia_intervento` (type of intervention), `categoria_intervento` (category of intervention), and `descrizione_sintetica_cup` (brief project description). The first three are standardised administrative fields; the fourth contains the project-specific narrative description.

[41]In zero-shot classification, the model performs a task from a natural-language instruction without receiving examples

trained on a task-specific corpus. This choice removes the need to construct a training dataset before classification and makes it possible to apply the procedure to the full administrative registry. Evidence from comparable text-classification tasks shows that zero-shot large language models can match or exceed the accuracy of paid human annotators (Gilardi et al., 2023). This evidence motivates the design, while the audit in Section 1.3.1 tests its performance on the specific language, administrative records, and classification margins considered here.

Second, the unit of classification is a single project. Each project whose input is not already stored in the cache is submitted in a separate API call rather than combined with other records in a batch[42]. This preserves a one-to-one correspondence between the project input, the model response, and the resulting indicators. It also prevents the classification of one project from depending on the content or ordering of other projects included in the same model context. Project-level calls make outputs directly traceable to their CUP and allow failed or ambiguous responses to be isolated without resubmitting a group of otherwise valid records.

Third, hydrogeological relevance and temporal orientation are classified in the same call. The procedure does not use a multi-step chain in which one model output becomes the input to a subsequent classification stage[43]. The single-stage design avoids propagating errors across intermediate model outputs and reduces the number of requests required per project. It also leaves the complete classification rule observable in one prompt, simplifying replication and inspection. The corresponding cost is that the model must resolve the complete classification task from a single instruction; the manual audit assesses whether it does so with sufficient accuracy.

Fourth, the implementation limits sampling variation and separates substantive classifications from operational failures. The model is queried at temperature zero[44]. A cache stores the response associated with each unique project description, and a retry routine resubmits requests affected by transient technical failures[45]. Caching ensures that identical descriptions receive the same stored classification and avoids repeated charges. The retry routine prevents a technical failure from being recorded as a substantive *unrelated* response and ensures that every observation in the final sample receives a valid model output.

Together, these choices make the procedure project-specific, auditable, and computationally in-

of previously labelled observations. This differs from few-shot prompting, in which examples are included in the prompt, and from fine-tuning, in which the model's parameters are updated using a task-specific training dataset.

[42] An application programming interface (API) call is a request sent by the Python script to the hosted model, containing an input prompt and returning a model response. In this context, batching would place several project descriptions in the same request and ask the model to classify them jointly.

[43] A multi-step chain divides a task into a sequence of model calls. For example, one call could first determine hydrogeological relevance, and a second call could classify relevant projects as ex ante or ex post. An error in the first stage would then determine which observations reach the second stage.

[44] Temperature is a parameter that controls the dispersion of the model's output distribution. Lower values concentrate selection on higher-probability responses, while higher values allow more variation. Setting temperature to zero minimises sampling variation but does not guarantee identical outputs across changes in the model version or the provider's infrastructure.

[45] A cache is a local record that associates an input with its previously obtained response. If the same description occurs again, the stored response is reused instead of issuing a new billed request. A retry routine repeats a request when no valid response is received because of a temporary connection failure, service interruption, or request-capacity constraint.

expensive. Classifying the full sample of 600,823 projects cost approximately €55 in API fees, less than one cent per project. The zero-shot design also means that application to another setting does not require model retraining or a pre-existing labelled corpus. It does, however, require adapting the prompt to the relevant hazard and administrative context and validating the resulting labels on records drawn from that setting.

1.3 Validation

The classification is validated twice, against different standards. The first check is internal: project-level labels are compared with hand-coded judgment on a random audit sample, which measures how accurately the pipeline reproduces a human reading of the same text. The second is external: it asks whether the series built from those labels behave around realised hazards as preparedness and remediation should, which no amount of label accuracy can guarantee on its own.

1.3.1 Manual Audit

Classification performance cannot be adequately summarised by a single accuracy statistic. Overall exact-match accuracy measures the proportion of projects for which all three indicators are correctly classified, but does not reveal which labels account for the errors, whether those errors arise from false positives or false negatives, or their implications for subsequent inference. Ludwig et al. (2024) show that validation data are required to characterise classification error when language-model outputs are used as measured variables in econometric analysis. Egami et al. (2023) further show that the direct use of imperfect generated labels in downstream regressions can produce biased estimates and invalid confidence intervals even when predictive accuracy is high. They propose an estimator that combines labels generated at scale with a smaller sample of manually validated observations.

The internal validation therefore reports overall exact-match accuracy together with label-specific measures of precision, recall, and F1[46]. Exact-match accuracy assesses the correctness of the complete classification assigned to each project. Precision and recall distinguish over-inclusion from omission and show how performance varies across hydrogeological relevance, preparedness, and remediation.

For the internal validation, I select a sample of 459 projects and manually assess the three indicators assigned to each observation: hydrogeological relevance, preparedness, and remediation. In cases of disagreement, the manual assessment is treated as the reference classification. Twenty-five projects contain at least one incorrectly classified indicator, yielding an overall exact-match accuracy of 94.6 percent.

Table A.3 reports classification performance separately for each indicator. For hydrogeological relevance, precision is 0.918 and recall is 1.000: the audit identifies 18 false positives and no false negatives. Preparedness has precision of 0.893 and recall of 1.000, with 13 false positives and no false

[46]Precision is the proportion of predicted positive observations that are manually confirmed as positive, $TP/(TP + FP)$. Recall is the proportion of manually confirmed positive observations that the classifier identifies as positive, $TP/(TP + FN)$. The F1 score is the harmonic mean of precision and recall, $2PR/(P + R)$, where P denotes precision and R denotes recall. TP, FP, and FN denote true positives, false positives, and false negatives.

Table A.3: Baseline comparisons on the manual audit sample

Method	Label	Accuracy	Precision	Recall	F1	TP	TN	FP	FN
HYDROADAPT (LLM)	Hydro-related	0.961	0.918	1.000	0.957	201	240	18	0
	Ex-ante	0.972	0.893	1.000	0.943	108	338	13	0
	Ex-post	0.965	0.888	0.946	0.916	87	356	11	5
Keyword dictionary	Hydro-related	0.915	0.879	0.935	0.906	188	232	26	13
	Ex-ante	0.867	0.662	0.889	0.759	96	302	49	12
	Ex-post	0.924	0.913	0.685	0.783	63	361	6	29
Sector code (*difesa del suolo*)	Hydro-related	0.878	0.993	0.726	0.839	146	257	1	55
	Ex-ante / Ex-post	not available, sector codes do not record direction							

Notes: Each method is scored against the manually validated labels of the 459-project audit sample. The keyword dictionary flags a project as hydro-related if the concatenated project text matches at least one of 28 Italian hydro-geological terms, and as ex-post if it additionally matches one of 12 past-event or repair cues (ex-ante otherwise). The sector-code filter selects OpenCUP sub-sector 05 *difesa del suolo*, the selection rule of Ricciotti and Pollice (2024); 11 of 459 audit projects lack a sector code and are treated as not flagged. Appendix 1.3.2 documents both classifiers, the full dictionary, and their characteristic errors. TP = true positives; TN = true negatives; FP = false positives; FN = false negatives.

negatives. Remediation has precision of 0.888 and recall of 0.946, with 11 false positives and five false negatives.

The label-specific results qualify the overall exact-match accuracy. No false negatives are observed for hydrogeological relevance or preparedness in the validation sample, while their lower precision reflects the inclusion of projects not confirmed by manual review. Remediation is the least accurately classified margin and is the only category for which the audit identifies both false positives and false negatives. Classification errors therefore arise predominantly from over-inclusion, while the distinction between ex-ante and ex-post interventions remains the principal source of misclassification.

This distribution of errors has practical implications for additional validation. False positives remain within the predicted positive class and can be sought through further manual review of that class. False negatives remain within the substantially larger set of predicted negative observations and cannot be identified without reviewing or sampling from that set. Additional review can therefore address over-inclusion by concentrating on predicted positive observations, whereas estimating omissions requires validation within the negative class.

To assess the contribution of the language model, I compare HYDROADAPT with two alternative classification methods applied to the same validation sample. The first is a keyword dictionary. It classifies a project as hydrogeologically relevant when the concatenated project text contains at least one of 28 Italian hydrogeological terms. Among the projects identified as relevant, a second set of lexical cues referring to past events or repairs determines the ex-post label; the remaining projects are classified as ex ante. The second alternative is the administrative soil-defence code (*difesa del suolo*) used by Ricciotti and Pollice (2024). This method identifies projects from the sub-sector assigned by the proposing administration and does not use the project description. It can identify hydrogeological relevance but cannot distinguish ex-ante from ex-post interventions.

The keyword dictionary performs relatively well on hydrogeological relevance, with an F1 score of

0.906, compared with 0.957 for HYDROADAPT. Its performance declines when the same vocabulary is used to distinguish preparedness from remediation. The F1 score is 0.759 for the ex-ante label and 0.783 for the ex-post label, compared with 0.943 and 0.916 for HYDROADAPT. This difference reflects the overlap in the language used to describe prevention and repair. For example, similar terms may describe an embankment reinforced before a flood and the same embankment reconstructed after one. Lexical cues alone do not consistently identify the temporal orientation of the intervention.

The administrative soil-defence code has precision of 0.993 for hydrogeological relevance but recall of 0.726. It identifies a narrow set of projects explicitly registered under soil defence while omitting relevant interventions recorded under the assets they construct or restore. Post-disaster road reconstruction is the principal example. Because the administrative code contains no information on temporal orientation, it cannot produce separate preparedness and remediation indicators.

The differences in measured coverage are larger in the full classification sample. HYDROADAPT identifies 76,902 hydrogeological projects, corresponding to 12.8 percent of the 600,823 classified records. The keyword dictionary identifies 40,688 projects, or 6.8 percent, while the soil-defence code identifies 12,760 projects, or 2.1 percent. These differences do not by themselves determine which classification is correct, but they show that the choice of method substantially affects the measured coverage of public adaptation. Appendix 1.3.2 reports the complete keyword lists, the construction of both comparison methods, and their principal error types.

Taken together, the internal validation shows that HYDROADAPT achieves high project-level accuracy and that its principal advantage over the alternative methods lies in distinguishing ex-ante from ex-post interventions. The external validation in Section 1.3.3 addresses the separate question of whether the series constructed from these labels exhibit the responses expected around realised hydrogeological events.

1.3.2 Baseline Classifiers

The two baselines in Table A.3 are the two standard technologies for this measurement problem, and each is built here as a strong version of its kind, so that the comparison speaks to the approach rather than to a careless implementation. Curated word lists are the canonical tool of economic text measurement, the policy uncertainty index of Baker et al. (2016) and the firm-level political risk of Hassan et al. (2019) are dictionary products, and Sautner et al. (2023) apply the same technology to climate exposure in earnings calls. Their known weakness is that word lists misfire on polysemous terms and do not transfer across domains (Loughran and McDonald, 2011). Self-declared administrative markers, for their part, are how adaptation finance is tracked in official statistics; audits of the OECD Rio markers document systematic misreporting (Michaelowa and Michaelowa, 2011), and machine classification of project descriptions has been proposed precisely to replace them (Toetzke et al., 2022). Scoring both technologies against the audit sample turns these two critiques into project-level numbers.

The keyword classifier reads the same concatenated text fed to the LLM, lowercased, and matches the 28 hydro-geological terms and 12 ex-post cues of Table A.4 as regular-expression stems, so that

inflected forms count: levee (*argine*), embankment works (*arginatura*), and the adjective (*arginale*) all match the stem *argin-*. A project is hydro-related if any relevance term matches; a hydro-related project is ex-post if any cue matches on top, and ex-ante otherwise. The list was built from domain knowledge and refined in a single pass of reading corpus texts, which added, for instance, watercourse (*corso d'acqua*); it was not tuned against the audit labels. Some exclusions are deliberate and generous to the baseline: consolidation (*consolidamento*), safety works (*messa in sicurezza*), and land reclamation (*bonifica*) describe buildings, schools, and asbestos removal far more often than levees, and admitting them would flood the classifier with false positives, the audit sample itself contains the consolidation of a social taxi service (*consolidamento servizio taxi sociale*).

The sector-code filter reproduces the selection rule of Ricciotti and Pollice (2024): a project is hydro-related when it is filed under soil defence (*difesa del suolo*), sub-sector 05 of sector 2 (environmental infrastructure and water resources) of the CUP classification. The code is chosen by the applicant at registration; it records what a project is filed under, not what it does. It therefore carries no information on whether an intervention precedes or follows a disaster, and the ex-ante/ex-post margin is undefined for this baseline by construction. The 2.4 percent of audited projects without a sector code in the registry count as not flagged.

On relevance, the keyword dictionary over-reads geography. Its precision of 0.88 hides a single dominant failure: 92 percent of its false positives are the word for river (*fiume*) occurring in street addresses and firm names rather than in descriptions of works, furniture for a kindergarten on *Via Fiume Vecchio*, tourism promotion by *Po Delta Tourism S.r.l.*, a web platform for an environmental culture centre named after the Po. The mirror error is quantitatively smaller but harder to fix: the dictionary misses 6 percent of true hydro projects, the damage repairs that name what is rebuilt rather than what destroyed it, the restoration of a severed local road (*ripristino viabilità . . . rifacimento strada interrotta*) contains no hydro-geological term to match, and no reasonable extension of the list would catch it without dragging in every road maintenance project in the registry.

Where the dictionary founders, however, is direction. Ex-ante precision drops to 0.66 and ex-post recall to 0.69, and the reason is structural rather than a matter of list quality: before and after the disaster share a vocabulary. Completing the consolidation of a slope already in landslide (*completamento del consolidamento, versante in frana*) is remediation phrased preventively; the hydraulic regulation of a river to restore its flow sections (*sistemazione idraulica . . . per ripristino sezioni di deflusso*) is preventive maintenance phrased as repair. Urgency markers fail the same way: urgent interventions (*interventi urgenti*) and emergency procedures announce that something must happen quickly, not whether the flood has already occurred. Consolidation (*consolidamento*) and restoration (*ripristino*) can describe the same levee, on either side of the same flood; a word list has no way to tell which side it is on.

The sector code fails in the mirror image. Its precision is nearly perfect (a single false positive, beach nourishment filed under soil defence), but it misses 27 percent of true hydro projects, and the misses are not random. More than half (58 percent) are filed under road infrastructure (*stradali*), and over two thirds of those are post-flood or post-landslide road and bridge repairs, such as the restoration of a bridge sill on the Trebbia river (*ripristino della soglia del ponte sul fiume Trebbia*). The pattern

Table A.4: The keyword dictionary

Panel A: hydro-relevance terms (28)	
Hazard phenomena	flood (*alluvion-*), river overflow (*esondazion-/esondat-*), landslide (*frana/frane/franoso/franamento*), earth slip (*smottament-*), flooding (*allagament-*), erosion (*erosion-*), hydro-geological instability (*dissest-*), cloudburst (*nubifrag-*), sea storm (*mareggiat-*), flood wave (*onda di piena*, *piena del*, *di piena*)
Watercourses and morphology	river (*fiume*), stream (*torrente/i*), watercourse (*corso/i d'acqua*), riverbed (*alveo*), bank (*sponda/spondale*), floodplain (*golena/golenale*), slope (*versante*), ditch (*fosso/fossi*)
Defence structures and works	levee (*argin-*), check dam (*briglia*), dam (*diga/dighe*), flood-control basins (*cassa/casse di espansione*, *laminazion-*, *scolmator-*), drainage (*drenagg-*), flow regulation (*regimazion-/regimentazion-*), soil defence (*difesa del suolo*)
Technical qualifiers	hydraulic (*idraulic-*), hydro-geological (*idrogeologic-*), water-related (*idric-*)
Panel B: ex-post cues (12)	
Repair vocabulary	restoration (*ripristin-*), reconstruction (*ricostruzion-*), damage/damaged (*dann-/danneggiat-*), emergency procurement procedure (*somma urgenza*)
Past-event references	following the… (*a seguito de-*), event of… (*evento/i … del*), the flood of… (*alluvione del*), post-flood (*post-alluvione*), calamity (*calamit-*), emergency (*emergenz-*), exceptional weather events (*eventi eccezionali*, *avversità atmosferiche*), May 2023 (*maggio 2023*)

Notes: Terms are matched as case-insensitive regular-expression stems on the concatenated project text; the exact expressions are in the replication code. A project is hydro-related if at least one Panel A term matches; hydro-related projects are ex-post if at least one Panel B cue matches, and ex-ante otherwise.

has a simple fiscal logic: remediation arrives through the budget lines of the assets being repaired, roads, bridges, sewers, sports facilities, not through the soil-defence line that finances purpose-built protection works. A filing code therefore undercounts precisely the component of adaptation spending that responds to disasters, and any disaster event study built on it would be biased towards finding no response.

The full sample confirms the ordering. HYDROADAPT flags 12.8 percent of the 600,823 projects of the classification sample, the dictionary 6.8 percent, the sector code 2.1 percent; the coded projects are almost a strict subset of the pipeline's selections (98 percent overlap) but one sixth of their number. The two baselines thus bracket the measurement problem from opposite sides: one is noisy on relevance and blind to direction, the other is precise on relevance but silent on direction and on most of remediation. The margin on which this paper is built, separating preparedness from remediation in the same administrative registry, is exactly where both break.

1.3.3 Validation Against Real Shocks

The audit establishes that the labels agree with human judgment on the projects it covers. It cannot establish that the resulting series behave like what they claim to measure. The second check asks whether the two margins respond to realised hydrogeological hazards as the economics literature predicts they should. It rests on two predictions.

The first prediction concerns remediation. It is a response to realisations: disasters trigger prompt and sustained increases in public outlays where they strike (Deryugina, 2017; Masiero and Santarossa, 2020), a response reinforced by electoral incentives that reward visible post-disaster relief (Healy and Malhotra, 2009; Garrett and Sobel, 2003). If the label is correct, remediation should rise sharply in the quarters immediately following the event, and specifically where the water was.

The second prediction concerns preparedness on impact. It is a function of expectations rather than realisations (Carleton et al., 2024), so a single event should not move it mechanically the moment it strikes; any response should be visible only with a lag, as beliefs update.

Local projections on rainfall (Section 1.3.3) trace the dynamic responses of preparedness and remediation around realised hazards.

This design tests the first two predictions: that remediation rises sharply and promptly where it rains hardest, and that preparedness does not move mechanically on impact. I estimate local projections (Jordà, 2005) of newly registered adaptation projects on rainfall shocks at the municipality level:

$$Y_{m,t+h} = \alpha_m^{(h)} + \delta_t^{(h)} + \beta^{(h)} X_{m,t} + \varepsilon_{m,t+h}, \qquad h = 0, \ldots, 12,$$

where $Y_{m,t+h}$ is the number of new projects registered in municipality m in quarter $t+h$, by category; $X_{m,t}$ is a rainfall shock; α_m and δ_t are municipality and quarter fixed effects; and standard errors are clustered by municipality. The panel covers 2010Q1–2024Q4, the period over which registry coverage is mature (Section 1.1). The preferred shock is the number of days in the quarter with rainfall above the 99th percentile of the regional distribution (approximately 32.5 mm); a continuous alternative, the maximum daily rainfall of the quarter, delivers the same picture, and two further definitions serve as robustness checks[47].

A municipality's own rainfall is not the only rainfall that should matter to it. Floods and landslides are one coupled system: rain falling upstream carries risk downstream once the water has had time to travel, so local rain alone would miss part of the mechanism. I therefore estimate the same local-projection design against upstream extreme rainfall, exploiting the river network, and compare it to the local specification. Local rain and upstream rain turn out to produce two different dynamic signatures, not one scaled down from the other, and I examine them in turn.

Local rainfall. The responses validate the first two predictions directly. Remedial projects rise where extreme rain falls: the impact response is small and imprecisely estimated, but the response peaks one quarter after the shock, at 0.229 additional projects per extreme-rain day ($p < 0.001$; Figure A.2a) and

[47] Appendix 1.3.4 reports the complete set of estimates under all four rainfall-shock definitions, separately for local and upstream rainfall and for total hydrogeological projects, preparedness, and remediation at each horizon.

remains positive and statistically significant through the medium run. That a shock produces an increase in project registrations within one quarter, rather than only after a multi-year lag, is consistent with evidence from Italian municipalities. Municipalities affected by floods exhibit a comparable short-run rise in capital expenditure and transfer revenue, concentrated in the years immediately following the event and varying with local vulnerability and resilience (Lodi et al., 2023), while municipalities affected by earthquakes display a similarly prompt and, in some regions, persistent expenditure response (Masiero and Santarossa, 2020).

Preparedness moves in the opposite direction on impact, declining by 0.061 projects per extreme-rain day ($p < 0.01$), and shows no positive short-run response under any of the alternative shock definitions (Appendix 1.3.4). This confirms the second prediction: because preparedness reflects expectations rather than realisations, an event must be translated into revised plans and funding decisions before it generates new project registrations. The aggregate fiscal response documented in the literature therefore has a clear counterpart in HYDROADAPT: within adaptation-project registrations, the short-run increase is concentrated in remediation.

Upstream rainfall. I use the river network to construct upstream connections between municipalities. For each municipality m and quarter t, I compute upstream extreme-rain exposure as the number of extreme-rain days observed in municipalities located along the same river within 30 kilometers upstream[48]. I then estimate the same local-projection specification using upstream rainfall as the shock and projects registered in municipality m as the outcome.

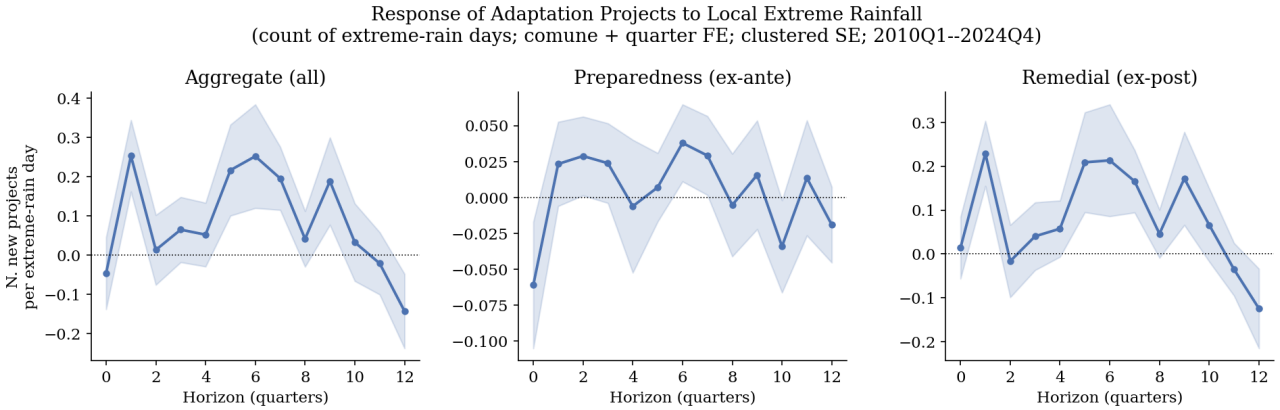
The registration response is not merely delayed but reversed on impact. In the same quarter as an upstream extreme-rain event, municipality m registers 0.247 fewer remedial projects per upstream extreme-rain day ($p < 0.01$; Figure A.2b), while the corresponding coefficient for all hydrogeological projects is -0.312 ($p < 0.01$). The response then rises through the medium run, turns positive by $h = 5$, and peaks at $h = 6$, both for remediation (0.492 additional projects per upstream extreme-rain day, $p < 0.001$) and for all hydrogeological projects (0.548, $p < 0.001$). Both responses remain statistically significant throughout the $h = 5$–10 window before fading. Local and upstream rainfall therefore generate different registration paths during the first two quarters: remediation is already rising by $h = 1$ following local rainfall, whereas the response to upstream rainfall remains below baseline before turning positive at longer horizons.

The timing is not hydrological. Water travels thirty kilometers in hours, whereas the registration response to upstream rainfall peaks six quarters later. The delay is therefore more plausibly interpreted as the administrative interval between damage being recognised and a corresponding project being registered. Damage associated with local rainfall is immediate, visible, and attributable to the municipality in which the rain occurs, and remedial registrations rise within the following quarter. Upstream rainfall instead links project registrations in a municipality to conditions elsewhere in its catchment. The resulting downstream damage is less certain, and its attribution to the upstream

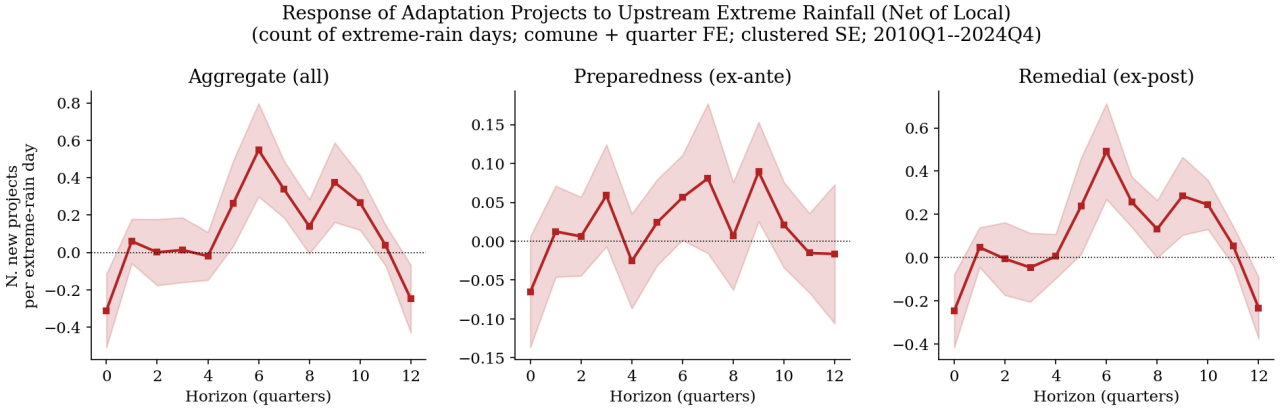
[48] HydroRIVERS provides the mapped course of each river and the direction in which its water flows (Lehner and Grill, 2013). Upstream connections follow the mapped flow direction, and distance is measured along the river rather than by straight-line geographical proximity. The mapped network covers 329 of the 330 municipalities; 318 municipalities (96 percent) have at least one qualifying upstream neighbour and enter the upstream-rainfall regressions.

Figure A.2: Response of adaptation projects to extreme rainfall.

(a) Local extreme rainfall



(b) Upstream extreme rainfall, net of local



Notes: Local-projection responses of newly registered projects at horizons $h = 0, \ldots, 12$ quarters. Panel (a) uses the number of local extreme-rain days in the quarter (daily rainfall above the regional 99th percentile, ≈32.5 mm); panel (b) uses the number of extreme-rain days falling upstream along the same watercourse, within 30 km and net of local rainfall. Within each panel: all hydrogeological-adaptation projects, preparedness, and remediation. Municipality and quarter fixed effects; standard errors clustered by municipality; sample 2010Q1–2024Q4. Shaded bands are 95 percent confidence intervals. Note in panel (b) the significant dip below zero at $h = 0$, absent from the local response, before the delayed rise to a peak at $h = 6$.

event may require additional information and coordination. A longer and more dispersed registration response is consistent with this process.

1.3.4 Local-Projection Robustness: Rainfall Threshold and Upstream Distance

The rainfall local projections of Section 1.3.3 rest on two choices: the percentile threshold that defines an extreme rain day (p99 in the main text) and the along-river distance that defines an upstream neighbour (30km in the main text). I vary each in turn, holding everything else at the main-text specification (Table A.5; Figures A.3 and A.4).

The upstream-distance result is not sensitive to this choice at all. At 15km, 30km, and 50km, the upstream response is significantly negative on impact, turns positive by the medium run, and

Table A.5: Local-projection robustness: rainfall threshold and upstream distance

	Impact ($h = 0$)			Peak		
	Aggregate	Preparedness	Remedial	Aggregate	Preparedness	Remedial
A. Local rainfall shock, by percentile threshold						
p90	-0.117***	-0.026***	-0.091***	+0.118*** ($h7$)	-0.026*** ($h0$)	+0.115*** ($h7$)
p95	-0.060***	-0.032***	-0.028**	-0.147*** ($h12$)	-0.032*** ($h0$)	-0.137*** ($h12$)
p99	-0.047	-0.061***	+0.014	+0.252*** ($h1$)	-0.061*** ($h0$)	+0.229*** ($h1$)
B. Upstream rainfall shock (net of local), by distance threshold						
15 km	-0.264***	-0.057*	-0.207**	+0.431*** ($h6$)	+0.069*** ($h9$)	+0.385*** ($h6$)
30 km	-0.264***	-0.062*	-0.202***	+0.522*** ($h6$)	+0.089*** ($h9$)	+0.469*** ($h6$)
50 km	-0.273***	-0.054*	-0.219***	+0.661*** ($h6$)	+0.095*** ($h9$)	+0.598*** ($h6$)

Notes: Local projections of newly registered adaptation projects (count of extreme-rain days regressor), municipality and quarter fixed effects, clustered standard errors, 2010Q1–2024Q4; same estimator as Figure A.2. Panel A varies the percentile threshold defining an "extreme" rain day (main text: p99, ≈32.5mm) for the local shock. Panel B varies the along-river distance defining an upstream neighbour (main text: 30km) for the upstream shock, net of local rain, at the p99 threshold. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

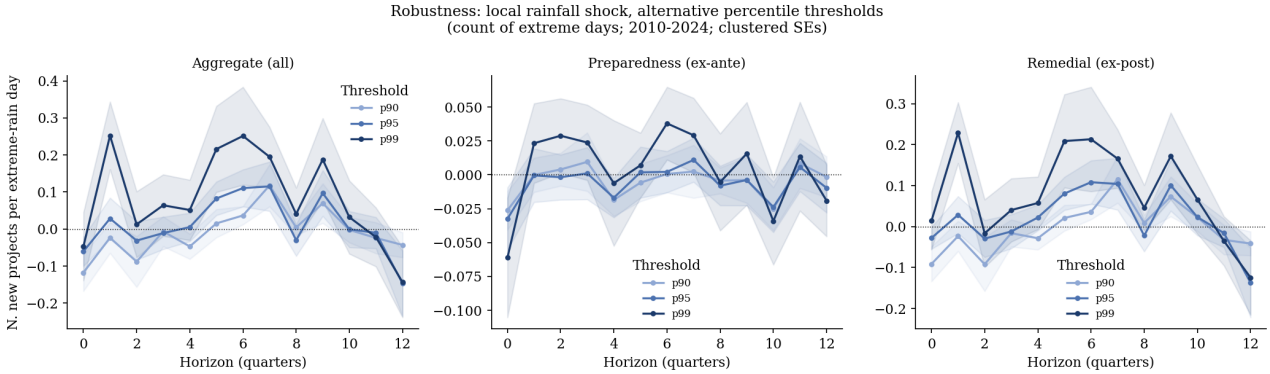
peaks at $h = 6$, for the aggregate series and for remediation alike; the three distance thresholds trace nearly identical paths (Figure A.4). The reversal-then-delayed-peak signature is a feature of the river network, not an artefact of where the 30km cutoff happens to fall.

The rainfall-threshold result is robust on the margin that matters most for the paper's argument, and more nuanced on the margin that matters less. What holds at every threshold I test, p90 through p99, is the central asymmetry: preparedness never shows a positive impact response, and is negative and significant at every threshold; remediation's peak response is positive and significant at every threshold. What shifts with the threshold is the *timing* of that peak and the sign of remediation's impact response. At p99 (the main-text choice, ≈32.5mm, averaging under one extreme day per municipality-quarter), remediation's impact response is small and insignificant and its peak arrives almost immediately, at $h = 1$. At the looser p90 and p95 thresholds (≈7mm and ≈14mm, averaging several "extreme" days per quarter), remediation's impact response turns significantly *negative*, and its peak is pushed out to $h = 7$–12. A plausible reading is that looser thresholds mix genuinely extreme days together with a larger number of merely wet ones, diluting and delaying the signal relative to a threshold that isolates the rarer, more damaging events; I report this pattern rather than resolve it, since distinguishing it from a real timing difference in the underlying response is beyond what this design can settle. The qualitative conclusion the main text draws from the panel, that preparedness does not react mechanically to rain while remediation eventually and robustly does, holds regardless.

1.4 Public Adaptation Investment in Emilia-Romagna

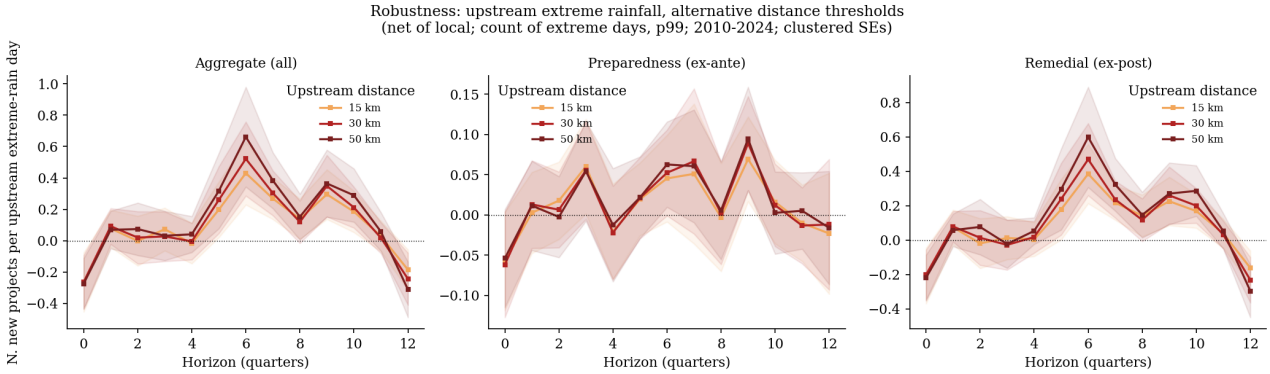
Following the classification validation in Section 1.3.1, I use the HYDROADAPT labels to characterise public investment in hydrogeological adaptation in Emilia-Romagna. The analysis considers three dimensions: scale and composition, temporal evolution, and spatial allocation.

Figure A.3: Robustness: local rainfall shock, alternative percentile thresholds



Notes: Local-projection responses of newly registered projects to the count of local extreme-rain days in the quarter, at horizons $h = 0, \ldots, 12$, for three definitions of "extreme": daily rainfall above the 90th, 95th, and 99th percentile of the regional distribution (≈7, 14, and 32.5mm respectively). Municipality and quarter fixed effects; standard errors clustered by municipality; sample 2010Q1–2024Q4; shaded bands are 95 percent confidence intervals.

Figure A.4: Robustness: upstream extreme rainfall, alternative distance thresholds



Notes: Local-projection responses of newly registered projects to the count of extreme-rain days (p99 threshold) falling upstream along the same watercourse, net of local rainfall, at horizons $h = 0, \ldots, 12$, for three definitions of "upstream": within 15km, 30km (main text), and 50km along the river network. Municipality and quarter fixed effects; standard errors clustered by municipality; sample 2010Q1–2024Q4; shaded bands are 95 percent confidence intervals.

The first concerns scale and composition: how large is registered investment in hydrogeological adaptation, and how is it divided between anticipatory preparedness and reactive remediation? This distinction corresponds to the ex-ante and ex-post margins of adaptation and has not previously been measured across a comprehensive project-level registry. The second concerns temporal evolution: how do preparedness and remediation change over time and around changes in funding arrangements and realised hydrogeological events? The third concerns spatial allocation: how closely is preparedness investment aligned with the corresponding measures of physical hazard and population exposure? Because the benefits of adaptation infrastructure are location-specific, its spatial allocation is relevant to the evaluation of public investment. The three dimensions are examined in turn.

1.4.1 Scale and Composition

I first quantify the scale of registered public investment in hydrogeological adaptation. Of the 1,052,460 projects registered in Emilia-Romagna with a valid CUP date between 2003 and 2025, 600,823 remain after the exclusion of standardised, mechanically generated records described in Section 1.1. HYDROADAPT classifies 76,902 of these projects as related to hydrogeological risk. They represent 12.8 percent of the classification sample and 7.3 percent of all projects with a valid CUP date. Yet these projects account for 16.7 percent of total public works expenditure, representing €47.4 billion in 2020 equivalent. Adaptation is thus systematically more capital-intensive than average public investment, with a mean project cost of €616,000 against €452,000 for unrelated works.

The central contribution of the measure, however, lies in splitting this aggregate burden into its anticipatory and reactive components. Here, the data reveal a sharp divergence between the frequency of adaptation events and their financial weight.

In terms of project counts, the public sector splits its efforts almost evenly: 54.8 percent of hydrogeological interventions are preparedness, while 45.2 percent are remediation. Local governments, in other words, initiate repairs almost as frequently as they initiate prevention. But the fiscal reality is fundamentally different. When measured by expenditure, the ratio is close to three to one: anticipatory preparedness commands 73.4 percent of the total adaptation budget (€34.8 billion), leaving 26.6 percent (€12.6 billion) to post-disaster remediation.

This divergence reflects the distinct underlying economics of the two adaptation channels. Preparedness is heavily capital-intensive: it requires large, forward-looking engineering works, such as levees, retention basins, and major structural reinforcements, averaging €824,000 per project. Remediation, by contrast, is inherently fragmented. Reacting to a disaster involves a high volume of smaller, localised, and cheaper interventions, such as restoring a severed road or clearing debris from a riverbed, with a mean cost of roughly €363,000. The gap is a right-tail phenomenon: the two channels' median project costs are far closer (€71,000 for preparedness against €50,000 for remediation), so the spending asymmetry is driven by a thin tail of very large preventive works rather than by the typical project. Public adaptation, in short, does not lack anticipatory investment; it is structurally divided between massive, capital-intensive preventive infrastructure and a constant, high-frequency stream of reactive repairs.

Hydrogeological adaptation comprises a heterogeneous portfolio of infrastructure addressing distinct hazards. Roads are the largest sub-sector in both categories, accounting for 30.8 percent of registered preparedness costs (€10.7 billion) and 41.6 percent of remediation costs (€5.3 billion). Their prominence reflects the importance of slope stabilisation and road reconstruction in the Apennines, where landslides repeatedly damage and interrupt the transport network.

Water supply and sewerage is the second-largest component of preparedness, accounting for 19.8 percent of registered costs (€6.9 billion), but represents a substantially smaller share of remediation. Soil defence and river, lake, and coastal works account for a further 12.3 and 9.5 percent of preparedness costs, respectively. The three largest sub-sectors together represent 63 percent of preparedness costs and approximately three-quarters of remediation costs.

Table A.6: Composition of public investment in Emilia-Romagna, 2003–2025

	Projects		Spending (2020 EUR)		Cost per project	
	N	%	bn	%	Mean	Median
All classified projects	600,823	100.0	284.2	100.0	473k	26k
Hydro-related	76,902	12.8	47.4	16.7	616k	60k
Preparedness (ex-ante)	42,170	54.8†	34.8	73.4†	824k	71k
Remediation (ex-post)	34,732	45.2†	12.6	26.6†	363k	50k
Unrelated	523,921	87.2	236.8	83.3	452k	22k

Notes: Frozen HYDROADAPT run (GPT-4o-mini, OpenCUP release of September 2025): 600,823 unique CUPs registered in Emilia-Romagna, January 2003–July 2025, after cleaning: excludes several large classes of standardised, mechanically generated entries (crop and livestock insurance, serial photovoltaic installations, mountain compensatory payments, and similar codified policy lines, Section 1.1) from the 1,052,460 Emilia-Romagna projects with a valid CUP generation date. Costs are total project costs at CUP registration, deflated to 2020 euros with the Italian CPI (AMECO). †Shares of the hydro-related subset; all other shares are of the full classified universe.

This composition shows that hydrogeological adaptation is implemented through general transport and water infrastructure as well as through projects explicitly registered under soil and flood defence.

1.4.2 The Time Profile

Disaggregating these two decades of investment, year by year, reveals how public adaptation responds to both gradual institutional learning and sudden extreme-weather shocks. The resulting time series (Figure A.5) exposes a system characterised by steady, long-run growth in anticipatory adaptation, punctuated by significant capital reallocations in the wake of disasters.

On the anticipatory margin, preparedness activity has expanded substantially. The frequency of preparedness interventions exhibits a sustained upward trend over the last decade, with average annual preparedness projects rising by 40 percent between 2015–2017 (1,784 per year) and 2018–2020 (2,497 per year). The most pronounced shift in this trend is administrative rather than climatic. Preparedness registrations jump in the first quarter of 2018 (first-quarter counts rise from 289 in 2016 and 334 in 2017 to 546 in 2018), exactly when the 2018 Budget Law (Law 205/2017, art. 1, paragraphs 853–861) opened a multi-year fund for municipal works to secure hydro-geologically fragile land and buildings, growing from €150 million in 2018 to €400 million by 2020[49]. The break is thus a fingerprint of a funding rule, not of a flood.

Changepoint detection using PELT[50], applied to the quarterly series by category (the algorithm

[49]Applications to the Ministry of the Interior required a valid CUP and were due by 20 February 2018, directly driving the spike in registrations. Alternative programmes do not match this timeline: the *Italiasicura* unit was closed in mid-2018 (which would reduce, not increase, registrations), while the February 2019 *ProteggItalia* decree only affects post-2018 trends.

[50]PELT (Pruned Exact Linear Time; Killick et al., 2012) jointly selects the number and location of changepoints by minimising within-segment variation subject to a penalty for each additional break. I use a radial-basis-function (RBF) cost, which can detect changes in the overall distribution of the series rather than only in its mean. STL (Seasonal-Trend decomposition using Loess) separates a time series into trend, seasonal, and residual components; the decomposition finds only weak seasonality, with a strength of 0.19, and residual spikes at the selected dates. The Augmented Dickey-Fuller (ADF) test evaluates the null hypothesis that the series contains a unit root, whereas the Kwiatkowski-Phillips-Schmidt-

is told neither where to look nor whether a break exists), locates three level shifts (Figure A.7). One, in 2010Q3, appears in every category at once and reflects the registry rather than investment: Law 136/2010, in force that same quarter, made the CUP mandatory on every public-works payment, discontinuously tightening coverage (Section 1.1). A second is the 2018Q1 preparedness break just described. The third, detected in 2023Q2, the quarter of the May 2023 flood, falls *only* on the remedial series; preparedness registers no shift around the event. The two categories keep different time: remediation moves with a realisation of the hazard while preparedness tracks funding rules rather than the event itself.

The magnitude of that reactive response is the defining feature of the series. In 2023, remediation projects surged to 5,112, and climbed further to a record 6,109 in 2024 (for context, remediation counts rarely exceeded 1,500 in any single year over the previous two decades). This post-disaster surge forces an abrupt, temporary reallocation of the adaptation budget (Figure A.6). Historically, preparedness commands the vast majority of adaptation expenditure, averaging 83 percent of the budget between 2003 and 2012, and 73 percent between 2013 and 2022. But when disaster strikes, the budgetary focus shifts overwhelmingly to emergency recovery. In 2023, the preparedness share dropped to 49.1 percent, as €1.75 billion were rapidly deployed for remediation.

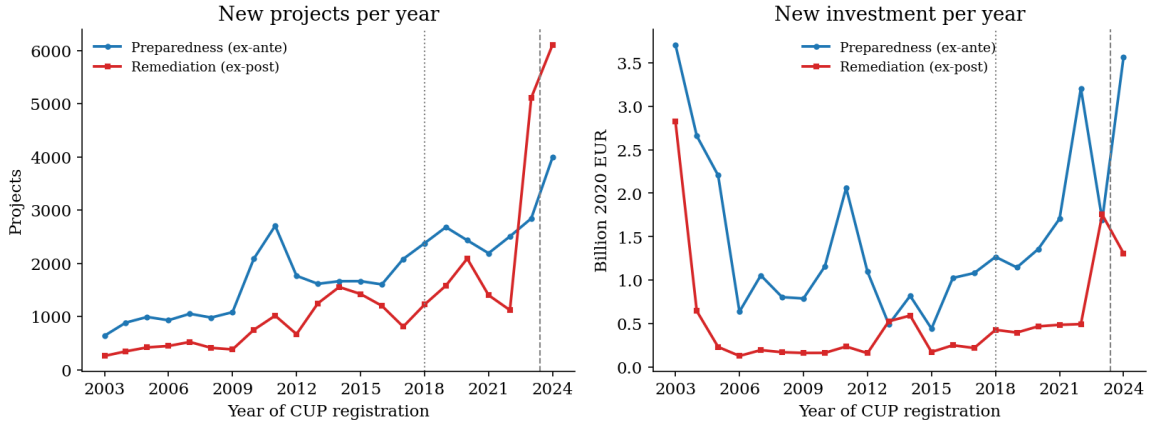
The 2024 data then pose a question the aggregate series cannot settle. One year after the floods, not only did remediation remain at historic highs, but the preparedness share rebounded to 73.2 percent of the budget, representing €3.57 billion across an unprecedented 3,995 projects. One interpretation is that the flood induced belief updating, with the response to May 2023 evolving from immediate repair into a wave of forward-looking protection; another is that the rebound reflects funding that became available on its own schedule, and that its timing relative to the flood is incidental. Either way, the composition margin does not show the adaptation gap closing on its own: absent a shock, the budget is already preventive, and shocks push the mix temporarily towards ex-post repair before it reverts.

1.4.3 The Cost of Under-Implementation

Despite substantial planned investment, a majority of at-risk watercourse segments lacked completed works when the flood arrived. Among the 1,591 segments with positive planned pre-2023 adaptation spending, 965 (60.7 percent) had an incomplete portfolio at flood time (Table A.7), with a mean financial shortfall of €1.84 million on incomplete segments. The finding that state-funded projects are 15 to 20 percentage points less likely to be closed than regionally funded ones, conditional on project size, location, and cohort, points to procurement bottlenecks and administrative burden as the proximate drivers of this gap. State and EU funding pipelines in Italy involve longer approval chains, stricter audit requirements, and more complex tendering procedures than regionally administered funds, all of which extend the time between project approval and financial closure. This institutional explanation is consistent with the stability of the gap across entry cohorts (Figure 2.1): the differential is not a temporary feature of a particular cohort but a structural property of the funding architecture.

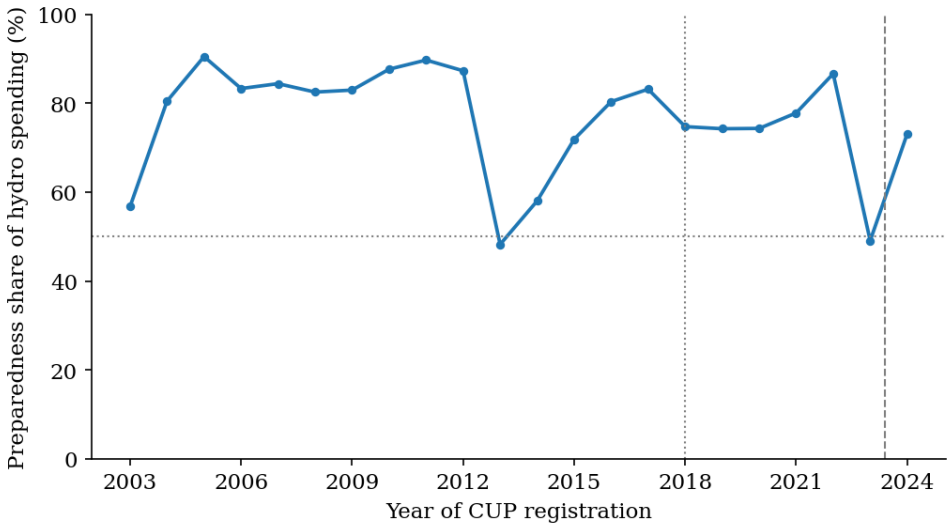
Shin (KPSS) test evaluates the null hypothesis that the series is stationary. Taken together, their results are consistent with a non-stationary series characterised by trend and discrete shifts.

Figure A.5: Public Adaptation Investment Over Time (2003–2024)



Notes: Annual counts (left) and annual investment in 2020 euros (right) of newly registered hydrogeological-adaptation projects in Emilia-Romagna, by category, 2003–2024 (frozen HYDROADAPT run). The dotted vertical line marks 2018, the funding-rule break in preparedness registrations; the dashed vertical line marks the second quarter of 2023, the May 2023 floods.

Figure A.6: Preparedness Share of Total Adaptation Expenditure

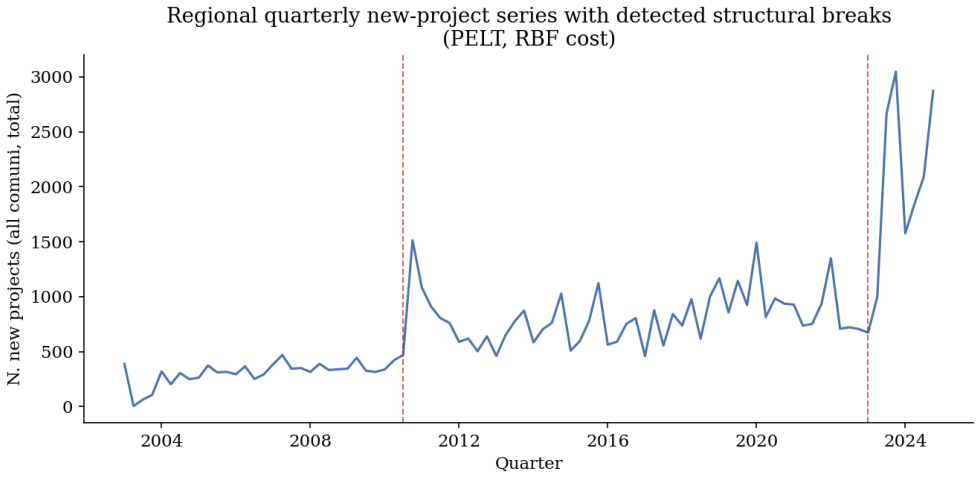


Notes: Preparedness share of total hydrogeological-adaptation spending (2020 euros), by year, 2003–2024. The horizontal dotted line marks an even 50/50 split between preparedness and remediation; the vertical dotted line marks 2018, the funding-rule break; the vertical dashed line marks the second quarter of 2023, the May 2023 floods.

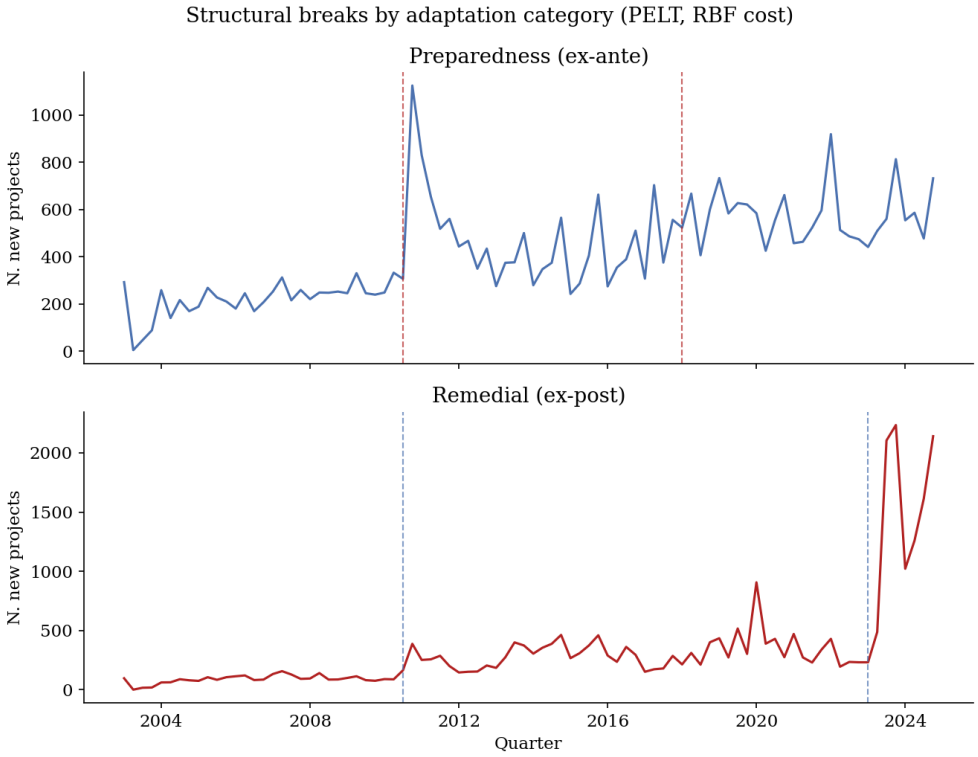
Read against the procurement literature, the pattern is unsurprising in kind but consequential in degree. Bandiera et al. (2009) show that most measurable waste in Italian public spending is passive rather than active, generated by the procedural apparatus itself rather than by theft or bad project choice, and that the burden falls hardest on administrations with the least autonomy over their own procedures. The comparative evidence points the same way: what predicts delivery is not the stringency of the rules on paper but the capacity of the body applying them (Bosio et al., 2022; Decarolis et al., 2020; Best et al., 2023; Rasul and Rogger, 2018), and in Italy contract enforcement conditions vary enough across jurisdictions to move procurement outcomes on their own (Coviello et al., 2018). What the present setting adds is a price. Because the delayed works are flood defences

Figure A.7: Structural breaks in quarterly adaptation investment.

(a) All categories pooled



(b) By category: preparedness and remediation



Notes: Quarterly counts of newly registered hydrogeological-adaptation projects in Emilia-Romagna, 2003Q1–2024Q4. Vertical lines mark breakpoints selected by PELT (RBF cost). Panel (a) pools all categories. Panel (b) reports preparedness and remediation separately, each with its own breakpoints: the 2023Q2 break, coincident with the May 2023 floods, appears only in the remedial series, while the preparedness series breaks in 2018Q1, aligned with no hydrogeological event. The 2010Q3 break, common to all series, reflects the coverage tightening induced by Law 136/2010.

Table A.7: Descriptive statistics of the execution gap (at-risk segments: planned > 0)

	N	Mean	SD	p25	p50	p75
Panel A: Full at-risk sample						
Execution gap (€)	1,591	1,118,803	5,148,079	0	27,508	374,757
Execution gap (count)	1,591	1.96	3.54	0	1	2
Panel B: Completion status						
Completed (gap ≈ 0)	626	39.3% of at-risk segments				
Incomplete (gap > 0)	965	60.7% of at-risk segments				
Panel C: Incomplete segments only						
Execution gap (€)	965	1,844,576	6,509,443	38,016	192,000	1,123,357

Notes: At-risk sample: segments with positive planned pre-2023 spending (N = 1,591). Execution gap = planned − closed, clipped at zero; completed means a gap of at most €100. Constant 2020 euros.

Table A.8: Ex-ante adaptation statistics by breach status

	Non-breached		Breached	
	Mean	Zero (%)	Mean	Zero (%)
Panel A: Planned adaptation				
Expenditure (€M)	1.10	32.1	1.75	20.8
Number of interventions	2.95	32.1	7.63	20.8
Panel B: Closed adaptation				
Expenditure (€M)	0.38	50.8	0.19	37.5
Number of interventions	1.58	50.8	2.04	37.5

Notes: Unit of observation is the river or canal segment; breached segments experienced at least one bank failure during the May 2023 floods. "Zero" is the share of segments with no planned (Panel A) or closed (Panel B) spending, in constant 2020 euros.

and the delay is observed against a realised flood, the cost of passive waste can be measured in the revenues of the firms that stood behind the unfinished levee, rather than in the accounting overrun on the contract.

It is not that public money fails to reach these places. As Section 2 describes, the Italian fiscal response to disasters is substantial and durable, with local expenditure elevated for a decade or more after an event. The resources arrive. They arrive afterwards, and they arrive as repair. What the execution gap measures is the part of the same fiscal effort that was directed at prevention, appropriated in time, and still not standing when the water came.

The policy implication does not concern the *quantity* of planned investment, the regional portfolio is large, growing, and directed towards the most vulnerable watercourses. What the evidence identifies is a cost of *delay*. The revenue losses borne by firms near incomplete works at flood time are large and quantifiable, and they arise not from a failure to plan but from a failure to execute on time. Against a backdrop of increasing flood frequency under climate change (Rentschler et al., 2023), the returns to reducing procurement timelines and administrative bottlenecks in hydraulic adaptation are likely to be substantial. Simplifying tendering and audit procedures, without compromising accountability, and strengthening coordination between regional and national authorities in project sequencing could

Table A.9: ORBIS/BvD Consolidation Codes and Sample Distribution

		Revenue 2016–2024		Geocoded sample	
Code	Description	Firms	Percent	Firms	Percent
C1	Consolidated accounts only	40	0.03	28	0.03
C2	Consolidated with unconsolidated companion	686	0.52	550	0.59
U1	**Unconsolidated accounts only**	**130,615**	**98.98**	**92,207**	**98.95**
U2	Unconsolidated with consolidated companion	624	0.47	401	0.43
Total		131,965	100.00	93,186	100.00

Notes: Definitions follow Bureau van Dijk (ORBIS) conventions. U1 identifies firms with unconsolidated accounts only, i.e. stand-alone firm accounts for which no group-level consolidated companion is available in the database. "Revenue 2016–2024" counts the Emilia–Romagna firms that report revenue in at least one year of the window; "Geocoded sample" counts the firms that survive the geocoding and validation steps of Table A.10.

meaningfully close the execution gap before the next event.

1.5 Firms

1.5.1 Download and Filter

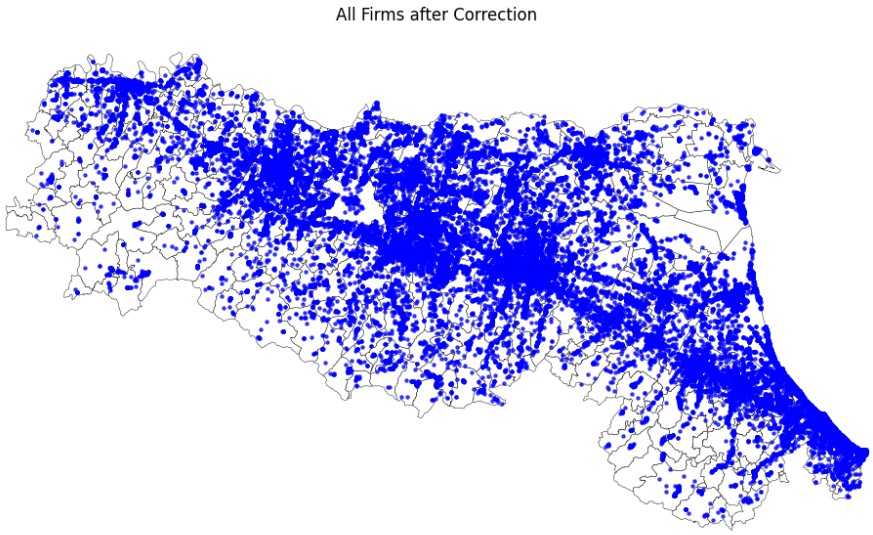
The Orbis extract covers the 684,213 firms registered in Emilia–Romagna, downloaded on 20 March 2026. Most of them cannot be used for balance-sheet analysis, because they report no financial statements or report them without revenue: only 131,965 firms report revenue in at least one year between 2016 and 2024. Of these, I keep the 107,304 that were still active in 2022, the last year before the flood, that is, firms with no recorded cessation or with a cessation dated after 2022. Firms that had ceased before the flood cannot have been affected by it.

The analysis focuses on **U1 companies**, that is, firms with *unconsolidated accounts only*. These are stand-alone firm accounts for which no group-level consolidated balance sheet is available in the database. This choice is motivated by comparability and consistency. First, unconsolidated accounts provide information at the individual-firm level, which is the relevant unit for the empirical analysis. By contrast, consolidated accounts combine the activity of multiple entities within a corporate group and therefore do not map cleanly into firm-level outcomes. Second, restricting the sample to U1 firms avoids mixing stand-alone balance sheets with group accounts, which could introduce substantial heterogeneity in the interpretation of financial variables. U1 firms are 99.0 percent of the firms that report revenue and 98.9 percent of the geocoded sample (Table A.9), so the restriction removes few firms.

1.5.2 Geocoding and cleaning of firm locations

The Orbis extract does not report coordinates, so each firm is located from its registered postal address, street, house number and municipality. Coordinates reported by Orbis would in any case be a weak source: Tarsia (2026), working with Orbis across Western Europe, finds implausible firm coordinates

Figure A.8: Location of Emilia–Romagna firms after geocoding and cleaning



frequent enough to require a dedicated correction procedure. Table A.10 summarises the steps that follow.

The address strings are first standardised. I harmonise case and spacing, expand the abbreviations used in Italian addresses (for example *Loc.* into *Località*), restore accented characters lost in the export, remove duplicated or truncated house numbers, and correct 17 addresses by hand. The 66 firms with no usable address are dropped, which leaves 107,238 firms to locate.

The addresses are then geocoded with Nominatim, the OpenStreetMap geocoder, which locates 91,028 firms (84.9 percent) at the first pass. The 16,205 that fail are retried after a second round of normalisation, which converts Roman numerals in street names, restores missing street types and rewrites the patterns on which Nominatim fails most often, such as hamlets (*località*, *frazione*) and provincial roads. The retry locates 3,229 further firms with Nominatim, 1,057 with Google Maps and 27 by hand from the Chamber of Commerce register. In total 95,341 firms (88.9 percent) are located; the remaining 11,897 cannot be placed reliably and are dropped.

Two validation steps follow. First, I overlay the firms on the regional boundary and drop the 55 that fall outside Emilia–Romagna, eleven of them because the address carries no municipality. Second, I assign each firm to the ISTAT municipality polygon in which its coordinates fall and compare it with the municipality recorded in Orbis, after harmonising the names. The two disagree for 2,100 firms (2.2 percent of those in the region). Rather than correct these cases, I drop them, since a coordinate placed in the wrong municipality may also be wrong within it. The final sample comprises 93,186 firms, every one of which lies in the municipality it declares (Figure A.8).

1.5.3 Construction of firm-level flood exposure

Geometric setup. Firm-level geocoded data are combined with high-resolution information on the geographical extent of the 2023 floods in Emilia–Romagna. Each firm i is represented by a point p_i with

Table A.10: Geocoding and cleaning of firm locations

Step	N firms	Share (%)	Description
Panel A: From the Orbis extract to the firms to locate			
Orbis extract	684,213	—	Firms registered in Emilia–Romagna.
Revenue reported	131,965	—	Revenue in at least one year, 2016–2024.
Active in 2022	107,304	—	No cessation recorded in or before 2022.
Usable address	107,238	100.0	Standardised street, house number and municipality.
Panel B: Geocoding			
Nominatim, first pass	91,028	84.9	OpenStreetMap geocoder on the standardised address.
Nominatim, second pass	3,229	3.0	Retry after further normalisation of failed addresses.
Google Maps	1,057	1.0	Addresses not resolved by Nominatim.
By hand	27	0.0	Located from the Chamber of Commerce register.
Total located	95,341	88.9	
Not located	11,897	11.1	Dropped.
Panel C: Validation			
Outside the region	55	0.1	Coordinates outside Emilia–Romagna; dropped.
Municipality mismatch	2,100	2.0	Coordinates in a municipality other than the one in Orbis; dropped.
Final geocoded sample	93,186	86.9	Every firm lies in its registry municipality.

Notes: Shares are computed relative to the 107,238 firms with a usable address. The second-pass, Google Maps and manual steps apply only to the 16,205 addresses that the first pass failed to resolve. Municipalities are the ISTAT polygons for Emilia–Romagna; names are harmonised before comparison.

coordinates in the projected coordinate reference system ETRS89 / UTM zone 32N (EPSG:25832), so that distances are measured in metres. The flood data consist of a set of polygons $\{S_j\}_{j=1}^{J}$ describing areas that were inundated during the event. For each firm i and flood polygon j the Euclidean distance is defined as

$$d_{ij} = \operatorname{dist}(p_i, S_j),$$

where $d_{ij} = 0$ if the point lies inside the polygon S_j, and $d_{ij} > 0$ otherwise. The minimum distance between firm i and any flooded area is then

$$d_i = \min_j d_{ij}.$$

Exposure categories. Firms are classified into mutually exclusive exposure groups based on d_i, using thresholds at 100, 250, 500 and 1,000 metres. The closest group corresponds to firms whose location falls inside at least one flooded polygon, i.e.

$$\exists j \text{ such that } p_i \in S_j \quad \Leftrightarrow \quad d_i = 0.$$

Exposure category	Number of firms	Share (%)
Inside flooded polygons	2,058	2.21
Within < 100 m	1,308	1.40
Within 100–250 m	2,389	2.56
Within 250–500 m	4,090	4.39
Within 500–1,000 m	6,493	6.97
No flooded polygon within 1 km	76,848	82.47
Total	93,186	100.00

Table A.11: Firm exposure to the 2023 Emilia–Romagna floods

Notes: Shares are percentages of the geocoded analysis sample (N = 93,186). Distance bands are mutually exclusive and defined relative to the nearest flooded polygon boundary.

These firms are labelled as "directly flooded" and are collected in the indicator

$$F_i^{\text{inside}} = \begin{cases} 1 & \text{if } d_i = 0, \\ 0 & \text{otherwise.} \end{cases}$$

For firms with $d_i > 0$, a sequence of non-overlapping distance bands around the flooded areas is constructed. Specifically,

$$F_i^{(0,100]} = \begin{cases} 1 & \text{if } 0 < d_i \leq 100, \\ 0 & \text{otherwise,} \end{cases} \qquad F_i^{(100,250]} = \begin{cases} 1 & \text{if } 100 < d_i \leq 250, \\ 0 & \text{otherwise,} \end{cases}$$

$$F_i^{(250,500]} = \begin{cases} 1 & \text{if } 250 < d_i \leq 500, \\ 0 & \text{otherwise,} \end{cases} \qquad F_i^{(500,1000]} = \begin{cases} 1 & \text{if } 500 < d_i \leq 1000, \\ 0 & \text{otherwise.} \end{cases}$$

Firms for which $d_i > 1000$ are treated as the reference group, with no flooded area within 1 kilometre of their location.

By construction, each firm is assigned to at most one exposure category: either it is directly flooded, or it falls into exactly one of the distance bands, or it lies more than 1 km away from the nearest flooded polygon. Equivalently, every firm belongs to at most one element of the set

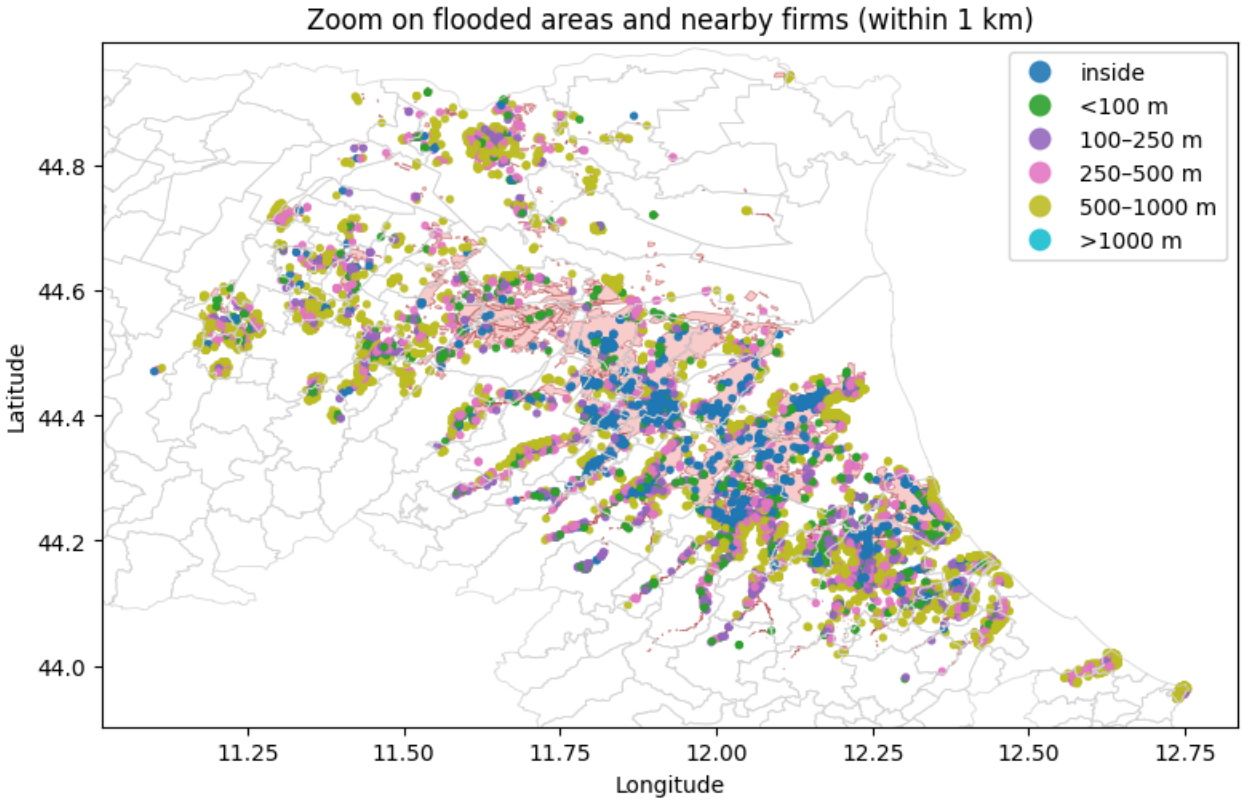
$$\{\text{inside},\ (0,100],\ (100,250],\ (250,500],\ (500,1000],\ >1000\},$$

so the indicators $F_i^{\text{inside}}, F_i^{(0,100]}, F_i^{(100,250]}, F_i^{(250,500]}$ and $F_i^{(500,1000]}$ are mutually exclusive at the firm level.

Multiple flooded polygons within bands. Although each firm is classified using its minimum distance d_i, the information on *all* flooded polygons that lie within the relevant radius is retained. For a firm i assigned to band b, the set

$$\mathcal{J}_{ib} = \{\, j : d_{ij} \leq r_b \,\},$$

Figure A.9: Spatial distribution of firms by flood exposure category.



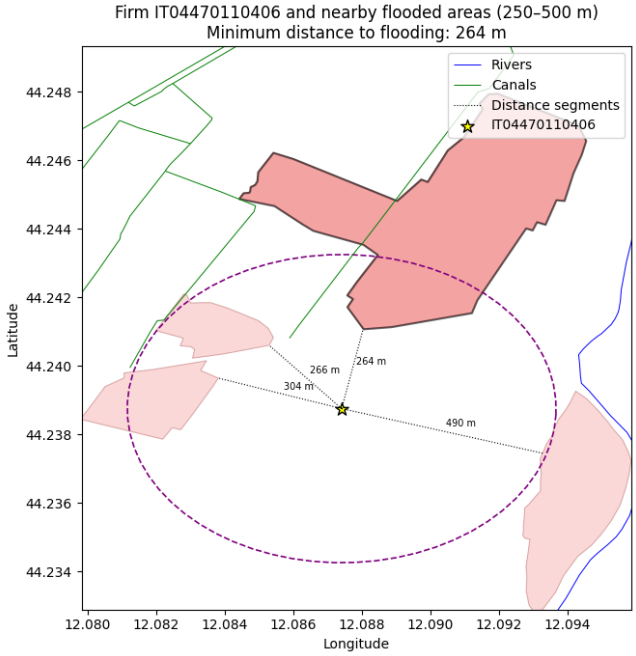
Notes: Colours indicate the firm-level exposure categories reported in Table A.11. Firms inside a flooded polygon appear in the darkest shade. Municipal boundaries and flooded polygons are overlaid for reference.

is recorded, where $r_b \in \{100, 250, 500, 1000\}$ is the upper bound of band b. This allows a firm to be associated with multiple flooded polygons within its assigned band (for example, if several flooded areas lie within 100 metres of the firm), and d_{ij} is computed separately for each of these firm–polygon pairs. In the empirical analysis it is therefore possible to work either with (i) the firm-level categorical exposure dummies, or (ii) the full set of distances d_{ij} and polygon identifiers, depending on the specification.

Figure A.10 illustrates a representative case of a firm that is simultaneously close to several flooded polygons in the 250–500 metre band. The circle shows the relevant distance buffer around the firm, and the straight segments represent the minimum-distance connection between the firm and each neighbouring flooded area. All of these polygons fall within the same band, but the closest one (shown in darker shading) is about 260 metres away. Rivers and canals are also plotted to provide a sense of the local hydrological network that channels the floodwater.

Sample Restrictions for the Adaptation Analysis. Table A.12 traces the restrictions from the firm panel to the sample used in the adaptation specifications. The panel comprises the 74,303 standalone firms with non-missing revenue in 2022. The binding constraint is geographic: only the 15,540 firms within 250 m of at least one planned hydraulic intervention have $W_i > 0$, and these are

Figure A.10: Example of a firm with multiple flooded polygons in the 250–500 m band.



Notes: The point denotes the firm location; circles indicate distance buffers; line segments show the minimum-distance connection to each nearby flooded polygon; rivers and canals are overlaid for reference.

the firms that enter the estimates of Table 5. The remaining 79.1 percent are located further from any planned works.

The restriction follows from the construction of the instrument. For a firm with $W_i = 0$ both Adapt_i and Z_i^{LOO} are identically zero, so the firm carries no variation in the endogenous interaction or in the excluded instrument and contributes only to the rainfall gradient. Keeping these firms does not change the estimate. Re-estimating Column 1 of Table 5 on all 74,303 firms with firm and year fixed effects gives a protective interaction of €109,743 (s.e. €25,028), against €105,348 (s.e. €26,065) on the restricted sample under the same fixed effects. The excluded firms are somewhat larger, with mean 2022 revenue of €2.65 million against €2.23 million (medians €295,000 and €269,000), and they received slightly more rainfall in May 2023, 24.3 against 22.2 mm over three hours. The share located inside the flood perimeter is similar, 2.2 against 2.6 percent.

Panel B reports two consistency checks. First, no firm has $\text{Adapt}_i > 0$ without $W_i > 0$ (row 7), as required by a project lifecycle in which completion presupposes planning. Of the 15,540 firms with planned works nearby, 12,282 also have completed works and 3,258 have none; the latter remain in the sample and supply the comparison at zero completed protection. Second, 595 firms have $W_i > 0$ but $Z_i^{\text{LOO}} = 0$ (row 6): their neighbouring projects all belong to cohorts in which no other project had closed by May 2023, so the leave-one-out completion rate is zero. They remain in the estimation sample with a zero instrument, and their baseline revenue does not differ from that of the other 14,945 firms (difference in mean log 2022 sales −0.014, s.e. 0.090).

Table A.12: Sample Construction

		Restriction	Firms	% of prev.
Panel A: From the firm panel to the estimation sample				
(0)		Standalone firms with non-missing revenue in 2022	74,303	—
(1)	+	At least one planned hydraulic work within 250 m ($W_i > 0$)	15,540	20.9
(2)	+	At least one completed hydraulic work within 250 m	12,282	79.0
(3)	+	Leave-one-out instrument $Z_i^{\text{LOO}} > 0$	14,945	96.2[a]
(4)		Estimation sample of Table 5 (= row 1)	15,540	—
Panel B: Overlap across sample definitions				
(5)		$W_i > 0$ and $\text{Adapt}_i = 0$ (planned works nearby, none completed)	3,258	—
(6)		$W_i > 0$ and $Z_i^{\text{LOO}} = 0$ (planned works nearby, instrument zero)	595	—
(7)		$\text{Adapt}_i > 0$ and $W_i = 0$ (completed works nearby, no planned works)	0	—

Notes: The table traces the reduction in sample size from the panel of standalone firms in Emilia-Romagna with revenue observed in 2022, the last pre-flood year, to the sample used in the adaptation specifications (Section 5.3). All counts are at the firm level in 2022. W_i is the stock of planned pre-2023 hydraulic investment on watercourse segments within 250 m of firm i; Adapt_i is the corresponding stock of projects that reached financial closure before May 2023 (equation 1). The instrument Z_i^{LOO} is defined in equation (8). Firms with $\text{Adapt}_i = 0$ or $Z_i^{\text{LOO}} = 0$ but $W_i > 0$ are retained in the estimation sample. The adaptation regressions are estimated on revenue in levels and so do not require strictly positive sales.
[a] Percentage computed relative to row (1).

1.6 Watercourses

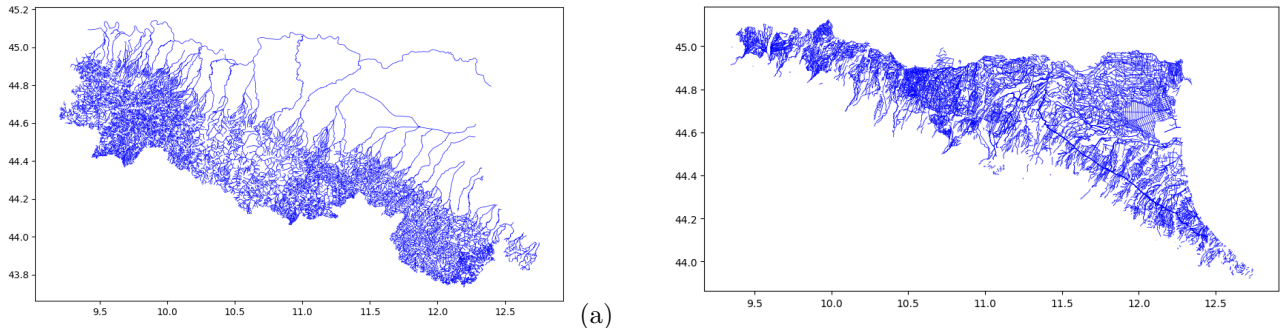
1.6.1 Rivers and Canals

To construct the hydrographic layers used in the analysis, I combined the regional canals shapefile with the full set of river shapefiles for Emilia-Romagna, consisting of **140** map tiles merged into a single river layer. I first verified that both datasets only contained linear features. The river layer includes **7,079** simple line geometries and **1,327** multi-part line geometries, while the canals layer includes **11,619** simple line geometries and **465** multi-part line geometries. I then reprojected both layers into a common geographic coordinate system and plotted them for visual inspection. As Figure A.11 shows, rivers are concentrated mainly in the Apennine area, whereas canals are predominantly distributed across the Po Valley plain. Next, I standardised and cleaned the watercourse names, harmonising abbreviations. I also manually excluded a small number of ambiguous or unidentified entries (97) for which no reliable name attribution could be established. After this cleaning step, I created a single harmonised name for each feature by combining the available type and name information. Since canals do not belong to the natural river hierarchy, I assigned them the lowest order in the classification.

1.6.2 Matching Water Bodies to Project Descriptions

The goal is to identify, for each investment, the specific river or canal segments that are the likely target of the intervention. To do so, I combine the textual descriptions from the investment database with the official hydrographic network for the Emilia–Romagna region.

Figure A.11: Hydrographic network in Emilia-Romagna.



(a) River network

(b) Canal network

Notes: Panel (a) shows the river network, concentrated mainly in the Apennine area. Panel (b) shows the canal network, predominantly distributed across the Po Valley plain. Source: ARPAE Emilia-Romagna.

Data and pre-processing. I start from the full universe of investment records, retaining only one observation for each distinct project description by dropping duplicate CUP codes with identical text. Let T_i denote the free-text description (in Italian) for investment i. I also compile a list of named water bodies from the regional hydrographic layers, extracting the field `APPOSI_NOME` separately for rivers and canals. For each water body w I construct a normalised, lower-case version of its name, and I similarly store a lower-case version t_i of each investment description.

Normalising river and canal names in the text. A distinctive feature of the Italian descriptions is that water bodies are often referred to using a plural generic term followed by a comma-separated list of proper names, without repeating the singular prefix. For example, a typical description might read

fiumi uso, marecchia, ausa, marano.

However, in the hydrographic database each water body is stored with the singular prefix (e.g. *fiume uso*, *fiume marecchia*, . . .). To align the wording of the investment texts with the hydrographic registry, I apply a rule-based text transformation.

Concretely, I implement a set of regular expressions that detect patterns of the form

$$\text{“fiumi” / “canali” / “rii” / “bacini”} + \text{list of names}$$

where the list can include commas and the conjunction “e”. For each match, I expand the plural form into repeated singular forms, attaching the appropriate prefix to every name in the list. In the example above, the text is rewritten as:

fiume uso, fiume marecchia, fiume ausa, fiume marano.

Analogous rules are applied to streams and canals, so that, for instance,

rii enzola, montefalcone e bandirola

Table A.13: Concatenated project text and watercourse-name normalisation

Concatenated OpenCUP fields	Watercourse names after normalisation
Acquisition or provision of services – other – watercourses – levees of the *rivers Uso, Marecchia, Ausa, Panaro e Naviglio*. Preparatory services for interventions aimed at securing riverbanks through the identification of burrows.	River Uso, River Marecchia, River Ausa, River Panaro, Naviglio Canal
Construction of public works – other – watercourses – basins *rivers Montone, Rabbi*. Works for the stabilisation of riverbanks and riverbeds.	River Montone, River Rabbi
Construction of public works – restoration – infrastructure for water resource use – canals *Galasso, Lorno*. Hydraulic safety works and water resource recovery through the construction of a multi-purpose retention basin and complementary interventions.	Galasso Canal, Lorno Canal

Notes: Text translated from Italian for exposition; the routine operates on the original lower-case Italian string. The first segment of each entry concatenates the three standardised administrative fields (`natura_intervento`, `tipologia_intervento`, `categoria_intervento`), followed by the narrative description (`descrizione_sintetica_cup`). Italics mark the plural generic term and the list of proper names that the regular expressions expand into the singular forms under which each watercourse is registered in the hydrographic layers.

is rewritten as

rio enzola, rio montefalcone, rio bandirola,

and phrases such as *bacini dei fiumi torrente taro e ceno* are converted so that each river name appears explicitly with its singular hydrological label.

Table A.13 gives three worked examples. Each row reproduces the concatenated OpenCUP text as it reaches the matching routine, with the plural construction that the rules rewrite in italics, alongside the singular watercourse names recovered from it.

The transformation is applied to the lower-case text t_i and returns a cleaned version $t_i^{(\text{clean})}$ together with an indicator for whether the row has been modified. I use these indicators both to quantify how often the pattern occurs and to focus subsequent manual checks on the small subset of descriptions where plurals remain after the automated pass. As an additional quality-control step, I generate an HTML report that shows, for each edited description, the original and cleaned text side by side, with removed tokens struck through and added tokens highlighted. This makes it possible to visually inspect and validate the behaviour of the regular-expression rules.

I also correct a handful of idiosyncratic typographical errors that would otherwise prevent matching, for example expanding truncated forms such as *t rossenna* into *torrente rossenna*.

Dictionary-based matching of water bodies. Once the descriptions have been normalised, I perform a dictionary-based match between the investment texts and the hydrographic registry. Let $\mathcal{R}$ be the set of river names and $\mathcal{C}$ the set of canal names obtained from the hydrographic layers. For each investment i, I search $t_i^{(\text{clean})}$ for whole-word occurrences of every name in $\mathcal{R}$ and $\mathcal{C}$, using regular expressions with word boundaries to avoid spurious partial matches (for example, avoiding that a river name embedded inside a longer word is mis-classified as a match).

For each record I store:

- the list of river names from $\mathcal{R}$ that appear in the text, and the corresponding count (number of distinct river names mentioned);
- the analogous list and count for canal names from $\mathcal{C}$;
- dummy indicators that equal one if at least one river (respectively canal) name is found in the description.

These variables form the basis for linking each investment to one or more water bodies in the spatial hydrographic network.

Treatment of ambiguous names and quality checks. Some hydronyms are also used as place names or administrative labels, which can generate ambiguous matches. A prominent case in the data is the string *rio bo*, which can refer to a municipality rather than a stream. For such cases I define ad hoc exclusion rules, for example dropping matches to *rio bo* when the surrounding text clearly refers to a municipality rather than to a watercourse.

Finally, I perform a series of diagnostic checks to validate the matching. These include inspecting the frequency distribution of matched water bodies, verifying that key rivers in the study area appear with plausible frequencies, and manually reviewing a sample of descriptions for which matches are particularly dense or potentially ambiguous. The resulting dataset associates to each investment a set of candidate rivers and canals, together with transparent rules that document how the textual descriptions have been normalised and matched to the hydrographic network.

Municipal allocation of project costs (2020 euros). After deduplicating CUP records, I allocate each project's deflated cost across the municipalities it covers. Cost figures are deflated to constant 2020 euros using the national consumer price index from the AMECO database of the European Commission (DG ECFIN), accessed via the Commission's online dashboard[51]. I first standardise `province` and `municipality` (lowercasing, trimming whitespace, removing non-breaking spaces) and map each province to its official number of municipalities (Bologna 55; Ferrara 21; Forlì–Cesena 30; Modena 47; Parma 44; Piacenza 46; Ravenna 18; Reggio nell'Emilia 42; Rimini 27). Projects flagged as region-wide (`province` = `ALL`) are treated as covering all 330 municipalities. Otherwise, within each province I assign the full provincial count whenever any row indicates `municipality` = `all`; if not, I count the distinct municipalities listed for that CUP–province. Let $\text{MunicipalCoverage}_i$ denote the resulting total (summed over provinces) for project i. Given the unique CUP-level cost in constant 2020 euros, ProjectCost2020_i, I compute the per-municipality cost as

$$\text{CostPerMunicipality2020}_i = \frac{\text{ProjectCost2020}_i}{\text{MunicipalCoverage}_i}.$$

[51] See European Commission, AMECO database, national consumer price index (Italy), available via the EC dashboard at https://dashboard.tech.ec.europa.eu/.

Table A.14: Coverage categories (unique CUPs): counts and shares

Category	CUPs	Share of CUPs (%)	Share of cost (%)
Only 1 municipality	6,502	74.91	45.28
Exactly 1 full province	265	3.05	23.61
>1 province (partial)	320	3.69	9.13
All region	108	1.24	1.73
Residual: multi-municipality within one province	1,485	17.11	20.25

Notes: One observation per CUP (unique projects). Categories are defined as follows: *Only 1 municipality* = the project covers exactly one municipality; *Exactly 1 full province* = the project covers all municipalities of a single province and no other jurisdictions; *>1 province (partial)* = the project spans more than one province but does not cover the entire region; *All region* = the project covers all 330 municipalities (or is flagged as region-wide); *Residual: multi-municipality within one province* = the project covers multiple municipalities within a single province, but not the full province. Shares refer to totals over unique CUPs.

Table A.15: Coverage categories (unique CUPs): cost statistics (2020 euros)

Category	Total cost (2020€)	Mean cost (2020€)	Median cost (2020€)
Only 1 municipality	3,766,033,490.38	579,211.55	98,193.23
Exactly 1 full province	1,963,424,898.65	7,409,150.56	224,936.71
>1 province (partial)	759,670,467.12	2,373,970.21	191,381.48
All region	143,555,735.99	1,329,219.78	20,272.98
Residual: multi-municipality within one province	1,684,566,049.94	1,134,387.91	170,068.03

Notes: One observation per CUP (unique projects). *Only 1 municipality*: exactly one municipality. *Exactly 1 full province*: all municipalities of a single province only. *>1 province (partial)*: spans multiple provinces without covering the entire region. *All region*: covers all 330 municipalities (or flagged region-wide). *Residual*: multiple municipalities within one province, but not full coverage. Shares refer to totals over unique CUPs. Monetary values in Table A.15 are in constant 2020 euros (CPI, AMECO).

I then merge back to the working file only these two CUP-level quantities to keep the dataset compact and reproducible.

Table A.14 documents how projects are distributed across territorial coverage, while Table A.15 reports the associated cost statistics (all figures at the CUP level, 2020 euros). Most projects are narrowly targeted: *Only 1 municipality* accounts for 74.9 percent of CUPs but a smaller 45.3 percent of total spending, with a low median cost (about €98k) and a much higher mean (€579k), signalling a right–skewed distribution with a few larger items. By contrast, projects that cover *Exactly 1 full province* are few (3.1 percent) but absorb 23.6 percent of expenditure and exhibit a large mean (€7.41m) relative to the median (€225k), consistent with occasional, capital–intensive programmes. Interventions spanning *>1 province (partial)* are rare (3.7 percent) and contribute 9.1 percent of costs; their mean (€2.37m) again exceeds the median (€191k), indicating skewness. *All region* projects are uncommon (1.2 percent) and fiscally marginal (1.7 percent). The *Residual: multi–municipality within one province* category represents 17.1 percent of CUPs and around one fifth of spending (20.3 percent), suggesting meaningful clustering of works within provinces without full provincial coverage. Overall, the evidence points to a landscape dominated by small, single–municipality projects, punctuated by a limited number of large interventions that drive a disproportionate share of total cost.

1.6.3 River Network Ordering: Flow Direction and Municipality Ranking

Raw data. The river network is sourced from the DBTR (Database Topografico Regionale) of Emilia-Romagna, which provides the hydrographic network as a set of polyline features. Each polyline represents a river segment and is composed of an ordered sequence of vertices. A critical limitation of this dataset is that polylines are stored as *undirected* geometries: segments may be digitised in either direction, so the start and end nodes of a feature do not necessarily correspond to the upstream and downstream endpoints of the physical waterway. Without correcting for this, upstream and downstream would be swapped for a non-trivial share of segments.

Decomposition into micro-segments. To obtain atomic units with unambiguous directionality, each polyline is exploded into consecutive vertex pairs (p_i, p_{i+1}), where p_i and p_{i+1} are adjacent vertices along the original feature. Each such pair constitutes a micro-segment with a well-defined start point, end point, and bearing. This decomposition is necessary because a single polyline can span multiple municipalities and may contain locally inconsistent vertex orderings that are masked at the feature level.

Flow direction correction. In Emilia-Romagna, all rivers descend from the Apennine mountains in the south to the Po plain in the north, so flow direction in the UTM 32N coordinate system runs from high northing (Y) to low northing (Y). For each micro-segment, I compare the northing of the start vertex Y_{start} to the northing of the end vertex Y_{end}. A segment is correctly oriented if $Y_{\text{start}} > Y_{\text{end}}$ (flowing south to north); segments violating this condition are reversed by swapping start and end vertices. Direction assignment is performed manually, cross-validated against altitude data and geographic knowledge of the basin. The resulting micro-segment network has a consistent flow direction throughout.

Municipality ranking. Each corrected micro-segment is spatially joined to its municipality via a point-in-polygon intersection at the segment midpoint. For every river–municipality pair (r, m), I compute the mean endpoint northing

$$\bar{Y}_{r,m} = \frac{1}{|S_{r,m}|} \sum_{s \in S_{r,m}} \frac{Y_s^{\text{start}} + Y_s^{\text{end}}}{2}$$

where $S_{r,m}$ is the set of micro-segments of river r falling within municipality m. Municipalities are then ranked along each river in descending order of $\bar{Y}_{r,m}$: rank 1 is the most upstream municipality (highest mean northing, closest to the Apennines), and the highest rank is the municipality closest to the river mouth.

Upstream and downstream investment. Firm i is located in municipality j. For each river r passing through j, the upstream municipality set is $\mathcal{U}_{i,r} = \{m : \text{rank}(r, m) < \text{rank}(r, j)\}$ and the

Table A.16: Distribution of project costs by watercourse assignment (2020 euros)

Category	Obs.	Mean	Std. dev.	Min	Median	Max
Only rivers	6,529	994,933	9,654,500	38	119,048	323,956,443
Only canals	1,816	839,169	5,081,837	25	99,495	189,873,418
Both	335	887,770	2,924,079	430	127,551	37,291,462
Total (assigned)	8,680					

Notes: One observation per CUP with deflated costs. "Only rivers": projects linked exclusively to river geometries. "Only canals": projects linked exclusively to canals. "Both": projects associated with at least one of each.

downstream set is $\mathcal{D}_{i,r} = \{m : \text{rank}(r,m) > \text{rank}(r,j)\}$. Upstream investment is the sum of completed adaptation expenditure across all municipalities in $\bigcup_r \mathcal{U}_{i,r}$; downstream investment is defined symmetrically. Upstream (downstream) rainfall is the simple average of maximum 3-hour May 2023 precipitation across the corresponding municipality sets, using the ERG5 gridded product matched to each municipality centroid. Firms located in municipalities that do not appear on any ranked river are excluded from the spillover analysis.

1.7 Rainfall

I next examine how rainfall intensity at the grid-cell level relates to the observed flood footprint in May 2023. Using hourly precipitation data aggregated to a regular grid (Fig. A.17), I construct several indicators that capture both cumulative and extreme rainfall: total May precipitation, total precipitation on selected flood days (1, 2, 10, 16, and 17 May), the maximum hourly and daily accumulations, and the maximum precipitation over windows of three consecutive hours and two consecutive days. I then compare these indicators across cells that intersect at least one of the official flood perimeters (EV1 or EV2) and those that remain dry. Table A.18 shows that flooded cells are systematically located in the upper tail of the rainfall distribution: their average total May precipitation is around 287 mm, compared with 233 mm in non-flooded cells, and the differences are even more pronounced for short-run extremes, such as maximum daily and multi-day accumulations. For instance, the mean maximum daily rainfall reaches about 87 mm in flooded cells versus 58 mm in non-flooded cells, and the maximum two-day accumulation is roughly 113 mm versus 81 mm, respectively. Overall, these descriptives confirm that the May 2023 flooding was concentrated in areas that experienced particularly intense and persistent rainfall, consistent with heavy precipitation being a key driver of the event.

2 Identification

2.1 The Funding-Pipeline Instrument

The instrument rests on the claim that projects financed through state and EU channels close more slowly than regionally administered ones for reasons of administrative procedure rather than of what the projects are or where they sit. Two checks bear on that claim, and three further ones, reported at the end of this section, bear on the exclusion restriction.

Table A.17: Grid-cell structure, raw weather variables, and cell-level aggregates (May 2023).

Panel A. Grid-cell information	
Number of grid cells (N_{cells})	1,131
Cell geometry	Polygon, constant within cell over time
Cell area (km^2)	Mean 25.063; Median 25.043 (min 24.788; max 25.455)
Approx. cell side length	$\sqrt{25} \approx 5$ km (about 5×5 km cells)
Spatial reference system	EPSG:4326 (WGS84); areas computed after reprojection to metric CRS
Temporal granularity of raw data	Hourly panel per cell
Panel B. Cell-level aggregates used in analysis (May 2023 floods)	
Aggregate	**Definition**
`tot_prec_may`	Total precipitation in May 2023 (sum of `PREC` over all May hours).
`tot_prec_selected_days`	Total precipitation on selected critical dates (1, 2, 10, 16, 17 May 2023).
`max_prec_3hr`	Maximum precipitation accumulated in any 3 consecutive hours in May 2023 (rolling 3-hour sum).

Notes: Raw data are hourly at the grid-cell level. Precipitation aggregates capture both short-run intensity (hourly and multi-hour bursts) and cumulative saturation (daily and multi-day totals) during the May 2023 flood episode. Non-precipitation summaries characterise contemporaneous local weather conditions.

Table A.18: Rainfall indicators by flood status of grid cell, May 2023

	Non-flooded cells		Flooded cells	
	Mean	Std. dev.	Mean	Std. dev.
Total precipitation in May (mm)	232.6	96.2	287.3	73.5
Total precipitation on selected days (mm)	161.4	91.4	233.5	65.1
Maximum hourly precipitation (mm)	9.7	4.9	12.5	3.8
Maximum daily precipitation (mm)	58.2	33.0	87.2	28.9
Maximum precipitation in 3 consecutive hours (mm)	20.0	7.8	27.1	7.1
Maximum precipitation in 2 consecutive days (mm)	80.6	47.3	113.4	39.3
Number of cells	888		243	

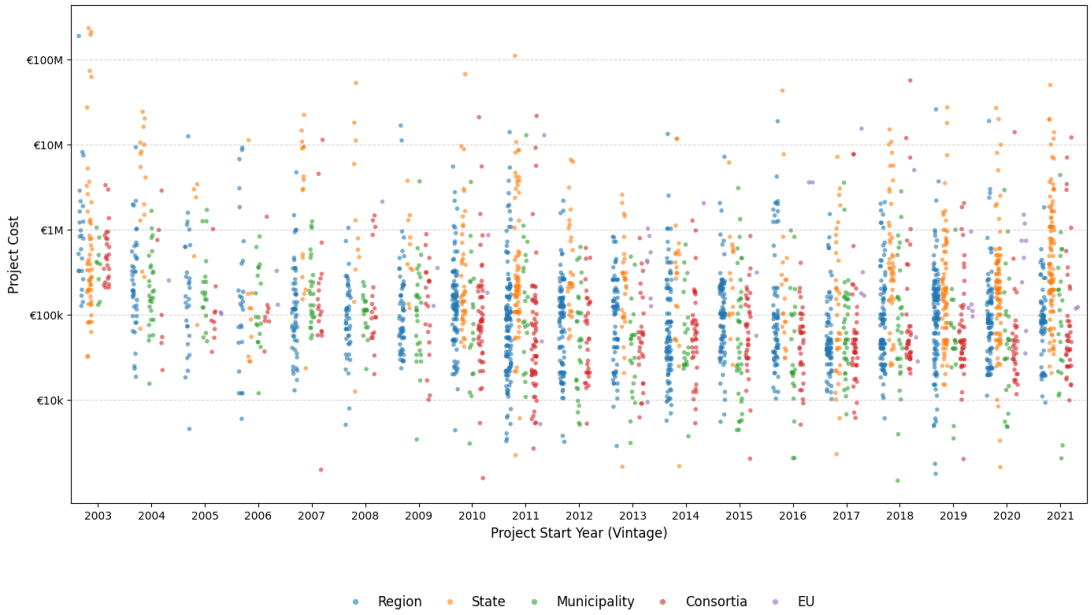
Notes: The table reports means and standard deviations of rainfall indicators by flood status of the grid cell in May 2023. Non-flooded cells are those with no intersection with either EV1 or EV2 flood polygons; flooded cells are those intersecting at least one of the two layers. Rainfall indicators are computed from hourly precipitation at the grid-cell level; the event-window variable is omitted for brevity.

Table A.19: Completion rates by funding source and vintage group

	Completion rate				
Vintage	State		Regional		Gap
2003–2010	41%	(179)	84%	(668)	+43 pp
2011–2015	19%	(229)	65%	(862)	+46 pp
2016–2019	19%	(491)	36%	(679)	+17 pp
Post-2019	*10%*	*(764)*	*17%*	*(373)*	*+7 pp*

Notes: Share of projects reaching financial closure by May 2023, by vintage group and funding source. Observations in parentheses. Post-2019 cohorts are italicised as the gap narrows mechanically for very recent projects.

Figure A.12: Project costs by funding source and start year.

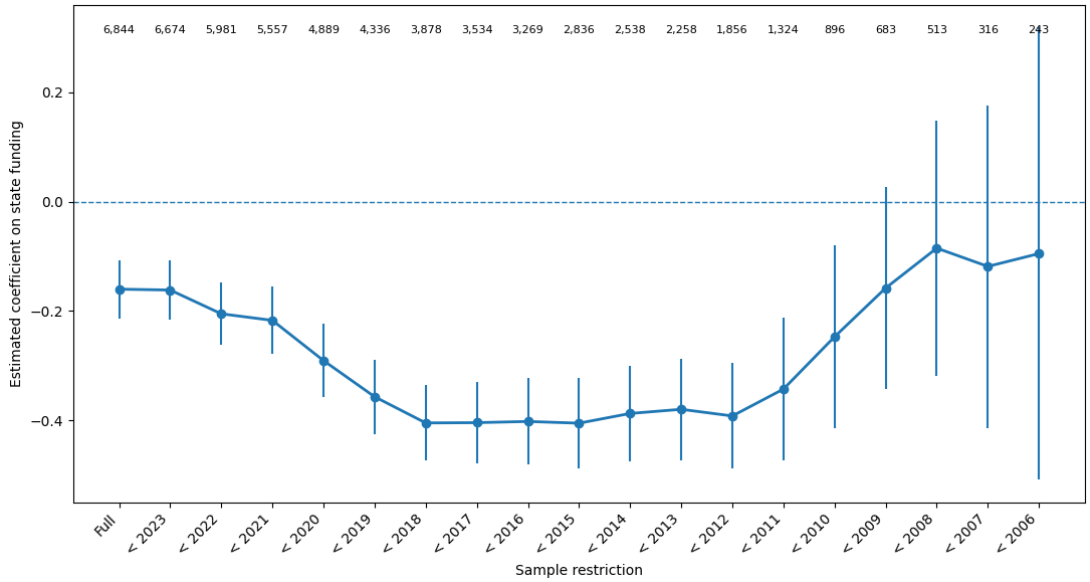


Notes: Each point is one project, plotted by start year (x-axis) and cost in constant 2020 euros (y-axis, log scale), coloured by funding source. At any given cost level and vintage, both regional and state-funded projects are present, confirming substantial overlap in project scale across funding sources.

Common support in project costs. Figure 2.1 plots project costs by funding source and start year for the full sample of hydrological adaptation projects. The cost distributions of regional and state-funded projects overlap substantially across all vintages, confirming that the differential completion speed documented in Table A.19 and Figure 5 cannot be attributed to differences in project scale or physical complexity.

Stability of the funding-source difference. Section 5.2 reports that state-funded projects were 15 to 20 percentage points less likely to have closed by May 2023 than regionally administered ones in the same municipality or on the same watercourse. The obvious alternative reading is mechanical: if state-funded projects were simply started later, they would have had less time to close, and the difference would say nothing about administrative execution speed.

Figure A.13: Estimated coefficient on state funding under progressive exclusion of recent cohorts.



Notes: The figure plots the estimated coefficient on the state-funding indicator from repeated linear probability models with municipality fixed effects, cohort fixed effects, and log project cost. Each point corresponds to a different sample restriction obtained by progressively excluding more recent cohorts. Vertical bars denote 95 percent confidence intervals; numbers at the top indicate sample size.

Figure 2.1 rules this out. It re-estimates the closure regression while progressively excluding the more recent cohorts, and the coefficient on state funding is stable across every restriction. The difference reflects the pipeline a project was assigned to, not its vintage: state and EU funding involve longer approval chains, stricter audit requirements and more complex tendering than regionally administered funds. This is the variation the funding-pipeline instrument exploits.

Risk targeting. Table A.20 regresses each adaptation variable on $\bar{R}_i$, mean May rainfall in the firm's grid cell over 2001–2022, the climatological measure of how often water is a problem where the firm stands. Column 1 shows that planning follows it. Columns 2 and 3 show that completion does not, with and without the planned stock as a control, and Column 4 that the instrument does not either.

Firm characteristics. Table A.21 asks whether the instrument can be predicted from what firms looked like before the flood. Each column enters one pre-flood characteristic on its own, holding the planning margin and the province fixed. None of the four is related to it, and the joint test on the firms for which all four are observed does not reject.

Timing. Table A.22 reports the reduced form estimated year by year, with 2022 as the omitted year. The specification includes the level terms in Z_i and the planned-works terms, both interacted with year; omitting the level terms biases the interactions, since $\text{Post}_t \times Z_i$ enters the pooled reduced form at −125,831. The coefficient on $Z_i \times R_i$ is indistinguishable from zero in each pre-flood year and

Table A.20: Adaptation measures and historical rainfall

	(1) Planned	(2) Completed	(3) Completed	(4) Instrument
$\bar{R}_i$	18.795***	1.981	-0.238	-1.906
	(5.862)	(2.027)	(1.825)	(2.728)
Planned works W_i			0.118	0.308***
			(0.076)	(0.055)
Firms	15,540	15,540	15,540	15,540

Notes: Cross-section of the 15,540 firms in the estimation sample. $\bar{R}_i$ is mean May rainfall in the firm's ERG5 cell over 2001–2022, in 10 mm units; the adaptation variables are in €1 million units. Standard errors clustered at the grid-cell level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.21: The instrument and pre-flood firm characteristics

	Dependent variable: instrument Z_i		
	One at a time (1)	Jointly (2)	Firms
Sales (EUR m)	-0.0020	-0.0034	15,540
	(0.0017)	(0.0021)	
Employees	0.0006	0.0013*	8,788
	(0.0006)	(0.0007)	
Tangible assets (EUR m)	-0.0086*	-0.0132*	15,540
	(0.0049)	(0.0068)	
Total assets (EUR m)	-0.0018*	0.0017	15,540
	(0.0010)	(0.0015)	
Joint F		1.87	
p-value		0.113	
Province fixed effects	✓	✓	
Firms, joint regression		8,788	

Notes: Cross-section of firms with positive planned works within 250 metres, measured in 2022. The dependent variable is the leave-one-out pipeline index Z_i, in €1 million units. Each column is a separate regression in which the characteristic named is the only covariate besides planned works W_i. Standard errors clustered at the grid-cell level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.22: Reduced form by year

	Dependent variable: annual revenue (EUR)					
	2018	2019	2020	2021	2023	2024
$Z_i \times R_i$	-2,546	20,068	17,230	-16,947	73,496***	89,022***
	(40,288)	(35,168)	(37,831)	(12,739)	(21,483)	(30,833)
Firms	15,540					
Observations	94,549					

Notes: Coefficients on $Z_i \times R_i$ from a year-by-year reduced form of equation (5); the dependent variable is annual revenue in euros and 2022 is the omitted year. Year interactions with Z_i, W_i and $R_i \times W_i$ are included throughout. Firm, year and province-by-year fixed effects. Standard errors clustered at the grid-cell level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

becomes large and significant in 2023 and 2024.

2.2 OLS, the Reduced Form and the First Stages

Table A.23 places the estimators side by side on the same sample and with the same controls, so that only the estimator changes. OLS puts the protective interaction at €60,322 against €104,338 under two-stage least squares, a gap of €44,016, and the Wu–Hausman test rejects exogeneity of completed works ($p = 0.039$). The reduced form carries the same message with the instruments in place of the endogenous terms, at €70,496 (€18,757).

Columns (4) and (5) report the two first stages in full rather than the F statistic alone. The loadings on the individual instruments should not be read separately: $\text{Post}_t \times R_i \times Z_i$ and $\text{Post}_t \times Z_i$ have a correlation of 0.97 in the post-flood sample, because rainfall varies little relative to its mean (a coefficient of variation of 0.35). The split of explanatory power between them is therefore unstable, while their joint power is not: the first-stage F is 310 under firm clustering.

3 Robustness Checks

Two questions can be put to the estimates. The first is whether the variation they exploit is the treatment: whether firms were on comparable paths before the flood, and whether the effect appears when and only when the water does. The second is whether the number survives the choices made in estimating it: how adaptation is measured, which fixed effects absorb what, and how uncertainty is computed. This appendix takes them in turn.

3.1 Is the Variation the Treatment?

Does rainfall track the hazard? Of the 801 grid cells in the sample, 234 contain flooded ground. They received 27.1 mm of three-hour rain on average, against 20.1 mm in the rest. Tables A.25 and A.26 put the comparison in a regression: within a province, a cell with 10 mm more three-hour rain is 8.7 percentage points more likely to contain flooded ground. The cumulative measures run

Table A.23: OLS, Reduced Form, Two-Stage Least Squares and the First Stages

	Sales			First stage	
	(1) OLS	(2) Reduced form	(3) 2SLS	(4) $R \times A$	(5) A
Post × Rain	-97,506	-93,597	-129,717*		
	(69,668)	(69,714)	(71,574)		
Post × Rain × Adaptation	60,322***		104,338***		
	(20,608)		(28,366)		
Post × Rain × Instrument		70,496***		-0.090	-0.401***
		(18,757)		(0.114)	(0.038)
Post × Instrument		-125,831***		2.340***	1.857***
		(46,223)		(0.181)	(0.063)
Observations	81,492	81,492	81,492	81,492	81,492
Firms	15,540	15,540	15,540	15,540	15,540

Notes: Estimation window 2018–2023. All columns use the same sample, the same planned-works controls and the same firm, year and province-by-year fixed effects; only the estimator changes. Columns (1)–(3) have annual revenue in euros as the dependent variable. Columns (4) and (5) are the first stages for $\text{Post}_t \times R_i \times \text{Adapt}_i$ and $\text{Post}_t \times \text{Adapt}_i$ respectively; their coefficients are dimensionless. Standard errors clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.24: Within-river logit: marginal effects on breach probability

	(1) Counts only	(2) Counts + expenditure
Planned adaptation (#)	0.0865***	0.0995***
	(0.027)	(0.030)
Closed adaptation (#)	−0.1621***	−0.2169***
	(0.049)	(0.061)
Planned adaptation (€)		−0.0170
		(0.016)
Closed adaptation (€)		0.0227*
		(0.013)
River fixed effects	Yes	Yes
Observations	92	92

Notes: Average marginal effects from logit regressions with river fixed effects; the dependent variable indicates a breach during May 2023. Adaptation is measured as count of interventions (#) and expenditure (€M, constant 2020). Standard errors in parentheses. The sample is the 92 watercourse segments for which breach status could be established: the regional records that document individual breaches are not public, so the exercise cannot be extended beyond them. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.25: Rainfall and the presence of flooded ground in a grid cell

	(1)	(2)	(3)	(4)	(5)	(6)
Max 3-hour	0.2264***	0.0873***				
	(0.0219)	(0.0291)				
Max daily			0.0492***	-0.0162*		
			(0.0047)	(0.0087)		
May total					0.0107***	-0.0083***
					(0.0015)	(0.0025)
Province fixed effects		✓		✓		✓
Grid cells	801	522	801	522	801	522

Notes: One observation per ERG5 grid cell. The dependent variable is an indicator equal to one if any part of the cell lies inside a May 2023 flood perimeter. Each column is a separate regression of the outcome on the rainfall measure named in the row, in 10 mm units. Flooded area is the intersection of the cell with the union of the Event 1 and Event 2 perimeters, computed in EPSG:32632. Heteroskedasticity-robust standard errors. Cells without a province assignment drop out of the fixed-effects columns. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.26: Rainfall and the share of a grid cell flooded

	(1)	(2)	(3)	(4)	(5)	(6)
Max 3-hour	0.0178***	-0.0126				
	(0.0040)	(0.0079)				
Max daily			0.0013**	-0.0158***		
			(0.0006)	(0.0025)		
May total					0.0002	-0.0045***
					(0.0002)	(0.0007)
Province fixed effects		✓		✓		✓
Grid cells	801	522	801	522	801	522

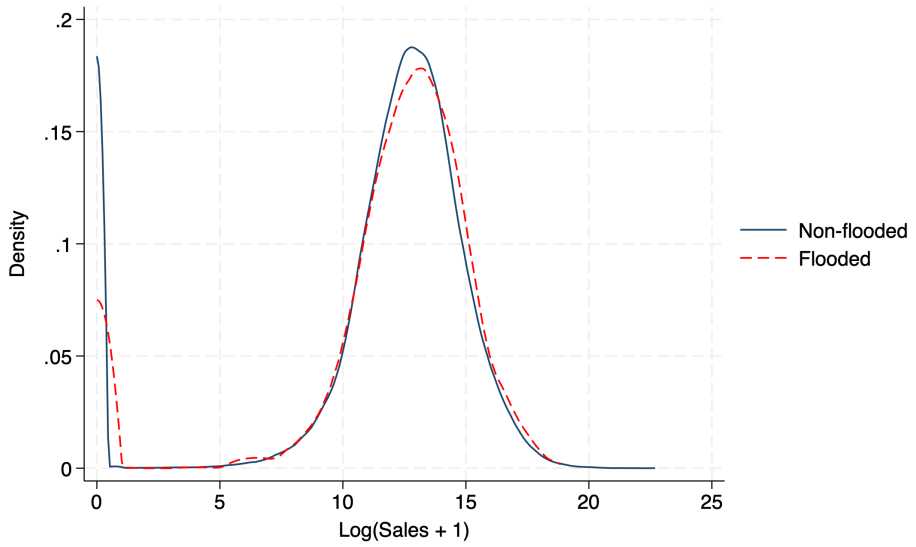
Notes: One observation per ERG5 grid cell. The dependent variable is the share of the cell's area inside a May 2023 flood perimeter. Each column is a separate regression of the outcome on the rainfall measure named in the row, in 10 mm units. Other conventions as in Table A.25. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

the other way, which is what the hydrology implies, since the rain that produced the May floods fell largely on the Apennine slopes while the water accumulated on the plain below.

Comparability of the two exposure groups. Figure 3.1 compares the 2022 revenue distributions of firms inside and outside the flood perimeter. The two are close, so the flooded group is not drawn from a different part of the size distribution than the rest of the region. This bears on interpretation rather than on identification: the firm fixed effects of equation (3) absorb differences in level, so comparable levels are neither necessary nor sufficient for the parallel-trends assumption examined in Section 5. What it does rule out is that the estimated effect describes a comparison between firms of systematically different size.

Pre-trends. The pre-flood event-study coefficients reported in Section 5.1 are $+0.021$, -0.029, $+0.001$ and -0.024, with a mean of -0.008. A joint Wald test of their equality to zero rejects at the ten percent level ($F = 2.07$, $p = 0.082$), so the four coefficients are jointly less reassuring than

Figure A.14: Sales Distribution by Flood Exposure Status, Baseline Year 2022



Notes: The figure displays kernel density estimates of $\log(\text{sales}_i + 1)$ separately for flooded and non-flooded firms in 2022, the last pre-treatment year. The +1 is a display convention that keeps firms reporting zero revenue on the axis; the estimating equations use $\log(\text{sales}_{it})$ on strictly positive sales and treat the extensive margin separately, for the reason given in Section 4.1. Flooded firms are defined as those located within the inundated area (band = 0). The sample is restricted to firms with non-missing sales in 2022.

Table A.27: Placebo Tests on False Event Dates

Assumed event year	2020	2021	2022	2023 (true)
	(1)	(2)	(3)	(4)
Flooded × Post	0.0019	-0.0056	0.0128	-0.0410***
	(0.0203)	(0.0195)	(0.0149)	(0.0144)
Observations	289,214	289,214	289,214	352,707
Grid cells (clusters)	798	798	798	798
Firm fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Each column re-estimates equation (3) assigning the treatment to the year in the column header. Columns 1–3 are estimated on 2018–2022 only, a window in which the flood has not yet occurred, so that any estimated effect must reflect pre-existing divergence rather than the shock. Column 4 is the true event date, estimated on 2018–2023. The outcome is log sales for firm-years with strictly positive sales. Standard errors, in parentheses, are clustered at the ERG5 5 × 5 km grid-cell level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

each of them is on its own.

Placebo event dates. Table A.27 reports the three possible false dates for the exercise described in Section 5.1. None produces an effect: the estimates are 0.002, −0.006, and 0.013, each far smaller than its own standard error, against −0.041 ($p = 0.005$) at the true date in column 4.

3.2 Does the Estimate Survive the Specification?

Alternative measures of adaptation. Table 3.2 re-estimates Column 1 of Table 5, changing only how Adapt_i is measured; the planned-works control and the instrument are redefined to match in each column. Column 1 is the baseline measure, the inverse-distance-weighted stock of completed works on every watercourse within 250 metres of the firm. Column 2 keeps only the works on the single nearest watercourse. Columns 3 and 4 keep the 250-metre buffer but restrict it to works on rivers and on canals respectively. The interaction is positive and significant in all four, at €104,338, €82,547, €103,547 and €90,737 per €1 million of completed works and per 10 mm of rainfall.

The first two columns measure the same thing with different precision. The buffer and the nearest-watercourse measure correlate at 0.92 across firms, and discarding the works on all but the closest watercourse lowers the estimate to €82,547. This is what measurement error produces if the discarded works also protect: the treatment is recorded with more noise and its coefficient is attenuated.

The last two columns are not a second view of the same firms. Of the 15,540 firms in the sample, 4,912 have completed river works within 250 metres and 7,439 have completed canal works, but only 69 have both, and the two measures correlate at -0.045. Rivers and canals therefore deliver two near-independent replications of the estimate, at €103,547 and €90,737. A confounder that made proximity to adapted infrastructure spuriously predictive of resilience to rainfall would have to operate separately in both sets of firms. The agreement also bears on the mechanism: May 2023 was a river flood, so one would expect only river works to have held the water back, and that canal works deliver an attenuation of similar size points to local hydraulic capacity in general, drainage as much as containment, rather than to the specific watercourse that overflowed.

What the table does not deliver is a ratio. The rainfall term is negative in every column and of similar magnitude, between $-83{,}640$ and $-129{,}717$, but it is distinguishable from zero only in Column 1, and there only at the ten percent level; elsewhere the standard errors are of the order of the point estimates. The attenuation is estimated more precisely than the loss it attenuates, so the share of the rainfall-driven loss that protection offsets can be computed only for the baseline measure[52].

Scaling up the fixed effects. Table 3.2 holds the sample and the specification fixed and adds fixed effects one layer at a time. Column (2) is the preferred specification.

The protective effect is stable across the first four columns, between 67,513 and 105,348, and remains significant throughout. It is smallest when province-by-year and sector-by-year effects enter together, which is expected: those absorb part of the variation in completed works that follows administrative and industrial geography.

Column (5) is the demanding test. Grid-cell-by-year effects absorb anything common to firms under the same rainfall, including the rainfall itself, so Post $\times$ Rain drops out and the triple interaction is identified only from differences in completed protection among firms facing an identical meteorological

[52]The 100-metre and 500-metre radii are omitted from Table 3.2. Under grid-cell clustering, the more conservative alternative, their first-stage statistics are 8.3 and 9.9, below the conventional threshold of ten; the four measures reported range from 11.2 to 126.0 on the same basis. At the 100-metre radius the interaction is in any case indistinguishable from zero (€35,674, $p = 0.37$), while at 500 metres it is €102,698 ($p = 0.04$), close to the 250-metre estimate.

Table A.28: Adaptation and Firm Sales: Alternative Measures

	(1) All, 250m	(2) Nearest	(3) Rivers	(4) Canals
Post × Rain	-129,717*	-112,936	-98,530	-83,640
	(71,574)	(70,073)	(69,352)	(70,402)
Post × Rain × Adaptation	104,338***	82,547***	103,547**	90,737*
	(28,366)	(25,892)	(50,175)	(46,510)
First-stage F	310	313	386	2370
grid-cell clustering	11.2	11.6	17.4	126.0
Observations	81,492	81,492	81,492	81,492

Notes: Each column re-estimates Column 1 of Table 5, changing only the definition of Adapt_i and of the corresponding planned-works control and instrument. Column 1 uses the inverse-distance buffer within 250 metres, Column 2 the single nearest watercourse, and Columns 3 and 4 restrict the works to rivers and to canals respectively. The sample is 2018–2023, so Post_t marks the year of the flood. Firm, year and province-by-year fixed effects throughout. The two first-stage F rows report the same Wald statistic on the excluded instruments in the first stage for $\text{Post}_t \times R_i \times \text{Adapt}_i$, computed under firm clustering and under grid-cell clustering respectively. Standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

shock. The estimate is 108,957, close to the preferred one, and the first-stage F rises to 559. Whatever a cell-level confounder could be, it cannot account for the result.

Table A.29: Protective effect as fixed effects are added

	(1) firm + year	(2) + province x year	(3) + sector x year	(4) + province x year + sector x year	(5) + grid cell x year
Post × Rain	-126,627**	-129,717*	-96,428	-125,166*	—
	(62,572)	(71,574)	(63,198)	(70,658)	
Post × Rain × Adaptation	105,348***	104,338***	77,368***	67,513**	108,957**
	(26,065)	(28,366)	(25,449)	(27,295)	(51,110)
First-stage F	268	310	275	318	559
Observations	81,492	81,492	81,492	81,492	81,492
Firms	15,540	15,540	15,540	15,540	15,540

Notes: IV estimates, 2018–2023, standard errors clustered at the firm. Post × Rain is absorbed in column (5) because rainfall does not vary within a grid cell. Planned-works controls are included throughout.

Spatial correlation in the errors. Rainfall is constant within an ERG5 grid cell and firms in neighbouring cells are exposed to correlated weather, so clustering at the firm may understate uncertainty. Table 3.2 re-estimates the preferred specification and changes only the variance estimator: the point estimates are identical across columns by construction.

Moving from firm to grid-cell clustering roughly doubles the standard error on the triple interaction, from 28,366 to 50,670. Conley spatial standard errors at cut-offs of 2, 5 and 10 km give almost exactly the grid-cell figure, between 50,514 and 52,017, which suggests the relevant dependence is local and is already absorbed by the cell. Under every estimator the protective effect remains significant at the 5 percent level. The coefficient on Post × Rain does not: it is significant at 10 percent under firm clustering and not distinguishable from zero under any spatial estimator.

Table A.30: Protective effect under alternative variance estimators

	Clustered		Conley spatial		
	(1) Firm	(2) Grid cell	(3) Conley 2 km	(4) Conley 5 km	(5) Conley 10 km
Post × Rain	-129,717*	-129,717	-129,717	-129,717	-129,717
	(71,574)	(90,968)	(93,572)	(92,764)	(94,416)
Post × Rain × Adaptation	104,338***	104,338**	104,338**	104,338**	104,338**
	(28,366)	(50,670)	(50,514)	(50,811)	(52,017)
Observations	81,492	81,492	81,492	81,492	81,492

Notes: Preferred IV specification, 2018–2023. Point estimates are identical across columns; only the variance estimator changes. Conley cut-offs are great-circle distances.

4 Additional Firm Outcomes

Tables A.31 and A.32 report the estimates for the income-statement and balance-sheet outcomes discussed in Section 5.4. Each outcome is estimated using the specification in Column 1 of Table 5.

Table A.31: Protection and the Income Statement

	(1) Revenues	(2) COGS	(3) Gross profit	(4) EBIT
Post × Rain	-115,200	-93,681	-21,593	389
	(78,897)	(63,466)	(28,014)	(9,084)
Post × Rain × Adaptation	100,713***	67,574***	33,153**	12,237**
	(28,894)	(22,648)	(13,304)	(4,905)
Observations	72,777	72,694	72,734	72,777

Notes: Each column reports the specification of Column 1 of Table 5 with a different dependent variable. Financial-statement items are taken from Orbis and merged to the estimation sample by firm and year, which accounts for the difference in the number of observations relative to Table 5. Completed works within 250 metres are instrumented with the leave-one-out predicted completion index; planned-works controls are included throughout. The sample is 2018–2023. Firm, year and province-by-year fixed effects. Standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Revenues are operating revenue and turnover; gross profit is revenue net of cost of goods sold.

Table A.32: Protection and the Balance Sheet

	(1) Tangible assets	(2) Total assets	(3) Current assets	(4) Solvency	(5) Current ratio
Post × Rain	10,157	107,502	-7,803	-0.617*	-0.097
	(14,707)	(76,577)	(42,950)	(0.352)	(0.121)
Post × Rain × Adaptation	-1,292	-53,294	-4,655	0.203	-0.045
	(12,402)	(75,651)	(59,594)	(0.156)	(0.052)
Observations	72,777	72,777	72,737	71,751	71,376

Notes: Each column reports the specification of Column 1 of Table 5 with a different dependent variable. Financial-statement items are taken from Orbis and merged to the estimation sample by firm and year, which accounts for the difference in the number of observations relative to Table 5. Completed works within 250 metres are instrumented with the leave-one-out predicted completion index; planned-works controls are included throughout. The sample is 2018–2023. Firm, year and province-by-year fixed effects. Standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Columns 1–3 are in euros; Columns 4 and 5 are ratios: solvency is shareholders' funds over total assets and the current ratio is current assets over current liabilities.

Table A.33 reports the effect on value added, which Section 5.3 uses to measure the return to protection. Value added is reported for about half of the firms, mostly the larger ones, so its coefficient is not comparable with the revenue coefficient of Table 5. Each pair of columns therefore estimates revenue and value added on the same observations.

Table A.33: Value Added and Revenue on the Same Firms

	Value added reported		Value added in every year	
	(1)	(2)	(3)	(4)
	Revenue	Value added	Revenue	Value added
Post × Rain	-194,308	-8,806	-275,216	4,148
	(127,449)	(26,866)	(194,722)	(39,448)
Post × Rain × Adaptation	83,614***	30,586***	129,552***	49,783***
	(28,658)	(10,456)	(42,082)	(15,752)
Value added / revenue effect		0.37		0.38
First-stage F	149	149	85	85
Observations	44,189	44,189	27,564	27,564
Firms	9,951	9,951	4,594	4,594

Notes: Each column reports the specification of Column 1 of Table 5 with the dependent variable named in the header. Columns 1 and 2 use every firm-year with value added reported; Columns 3 and 4 restrict the sample to firms that report value added in every year of 2018–2023. Within each pair, revenue and value added are estimated on identical observations. The row "Value added / revenue effect" is the ratio of the two interaction coefficients. Firm, year and province-by-year fixed effects; standard errors, in parentheses, are clustered at the firm level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.