Flooded Cities*

Adriana Kocornik-Mina, Thomas K.J. McDermott,
Guy Michaels, Ferdinand Rauch†

May 21, 2018

Abstract

Does economic activity relocate away from areas that are at high risk of recurring shocks? We examine this question in the context of floods, which are among the costliest and most common natural disasters. Over the past thirty years, floods worldwide killed more than 500,000 people and displaced over 650,000,000 people. This paper analyzes the effect of large scale floods, which displaced at least 100,000 people each, in over 1,800 cities in 40 countries, from 2003-2008. We conduct our analysis using spatially detailed inundation maps and night lights data spanning the globe's urban areas, which we use to measure local economic activity. We find that low elevation areas are about 3-4 times more likely to be hit by large floods than other areas, and yet they concentrate more economic activity per square kilometer. When cities are hit by large floods, these low elevation areas also sustain damage, but like the rest of the flooded cities they recover rapidly, and economic activity does not move to safer areas. Only in more recently populated urban areas, flooded areas show a larger and more persistent decline in economic activity. Our findings have important policy implications for aid, development and urban planning in a world with rapid urbanization and rising sea levels.

Keywords: Urbanization, Flooding, Climate change, Urban recovery *JEL Classification: R11, Q54*

*Corresponding author: Michaels (g.michaels@lse.ac.uk). For helpful comments and discussions, we thank David von Below, Vernon Henderson, Matthew Kahn, Somik Lall, Gianmarco León, Steve Pischke, Tony Venables as well as seminar participants at the CEPR workshop on Political Economy of Development and Conflict at UPF, Bonn University, the OxCarre Annual Conference in Oxford, the Irish Economic Association 2015, EAERE 2015, RSAI-BIS 2015, the Grantham Research Institute, UCC, NUI Galway, the World Bank Research Workshop on Urbanization in Developing Countries, and the World Bank and George Washington University-IIEP Conference on Urbanization and Poverty Reduction. We are grateful to the Centre for Economic Performance, the Global Green Growth Institute, the College of Business and Law, UCC and the Strategic Research Fund, UCC for their generous financial support. We also gratefully acknowledge the generous support of a Global Research Program on Spatial Development of Cities, funded by the Multi Donor Trust Fund on Sustainable Urbanization of the World Bank and supported by the UK Department for International Development. In Adriana Kocornik-Mina's case much of the work was carried out while at LSE's Grantham Research Institute. In Tom McDermott's case, much of the work was carried out while at LSE's Grantham Research Institute, and at UCC's School of Economics. Any opinions expressed and any mistakes are our own.

†Wageningen Environmental Research, NUIG and LSE, LSE, and Oxford respectively.

1 Introduction

Does economic activity within cities readjust in response to major shocks, which are potentially recurrent, and which disproportionately threaten specific neighborhoods? We examine this question in the context of floods, which are among the costliest and most recurring natural disasters.

According to media reports collated by the Dartmouth Flood Observatory, from 1985-2014 floods worldwide killed more than 500,000 people, displaced over 650,000,000 people and caused damage in excess of US$800 billion, adjusted using the CPI for 2000 USDs (Brakenridge 2015). Other datasets tell of even farther reaching impacts: according to the International Disaster Database (EM-DAT – see Guha-Sapir, Below and Hoyois, 2015), in 2010 alone 178 million people were affected by floods and total losses exceeded US$40 billion. To these direct costs we should add longer term costs due to disruptions of schooling, increased health risks, and disincentives to invest.

If there were perfect housing markets one might argue that these risks must be balanced by gains to be had from living in flood-prone areas. But as Kydland and Prescott (1977) show in their Nobel-prize winning contribution, flood plains are likely to be overpopulated, because the cost of building flood defenses tends to be borne in part by people who reside in safer areas. This problem is exacerbated by the fact that reconstruction costs in the aftermath of floods are usually also borne in part by non-residents.[1] This situation creates potential for misallocation of resources, and forces society to answer difficult distributional questions. Our paper examines how prevalent it is for economic activity to concentrate in flood-prone areas, and whether cities adapt to major floods by relocating economic activity to safer areas.

To frame our analysis, we outline a simple model, which considers how a large flood may affect an individual's decision to locate in safe or risky locations. The model predicts that flooding may cause people to relocate away from risky areas because of either Bayesian updating on the probability of a flood, or because floods destroy sunk investments, thereby reducing the cost of moving relative to staying. At the same time, the model suggests that even large floods may induce little new relocation of economic activity if past adaptation means that recent flooding conveys relatively little new information.

In our empirical analysis we study the local impact of large-scale urban floods. We use new data from spatially disaggregated inundation maps of 53 large floods, which took place from 2003-2008. The floods that we study affected 1,868 cities in 40 countries around the globe, but mostly in developing countries. These floods were all ruinous, displacing over 100,000 people each. We study the economic impact of the floods using satellite images of night lights at an annual frequency.

Our data show that the global exposure of urban areas to large scale flooding is substantial, with low elevation urban areas flooded much more frequently. Globally, the average annual risk of a large flood hitting a city is about 1.3 percent for urban areas more than 10 meters above sea level, and 4.9 percent for urban areas less than 10 meters above sea level. These estimates

[1] In some poor countries insurance markets are dysfunctional, so it is understandable that the government takes on some role in providing flood insurance. But even in those circumstances people who live on flood plains place a burden on taxpayers who live elsewhere.

likely represent a lower bound on urban flood risk since we do not have detailed flood maps for all the large flood events in the period we study. Of course, this average risk masks considerable variation across locations. Local flood risk results from a complex combination of local climate, permeation, and topography, among other factors. Some urban areas – even if located at low elevation – will flood rarely, if ever, while others are exposed to recurrent flooding. For example, even in our relatively short sample period (January 2003 to May 2008), a substantial number of cities were flooded repeatedly in our data. Out of 34,545 cities in the world a little over 5 percent (1,868 cities) get flooded at least once in our data. Conditional on being flooded, about 16 percent were flooded in more than one year. This is consistent with systematically higher risk of flooding in these locations.

In spite of their greater exposure to large flood events, we find that across the globe, urban economic activity, as proxied by night light intensity, is concentrated disproportionately in low elevation areas.[2] This disproportionate concentration of economic activity in flood-prone areas is found even for areas that are prone to extreme precipitation.

When we analyze the local economic impact of large floods, we find that on average they reduce a city's economic activity, as measured by night time lights, by between 2 and 8 percent in the year of the flood. The larger estimates come from using measures of extreme precipitation, rather than flood maps. The fact that we find similar patterns with extreme precipitation is reassuring that our results on floods are not driven by selection of particularly devastating floods, or by the omission of floods that we cannot observe for lack of inundation maps.[3]

Our results also show that recovery is relatively quick – lights typically recover fully within a year of a major flood, even in low elevation areas. This suggests that there is no significant adaptation, at least in the sense of a relocation of economic activity away from the most vulnerable locations.[4] With economic activity fully restored in vulnerable locations, the scene is then set for the next round of flooding.[5]

Some readers may contend that this pattern of rapid recovery shows that the expected damage reflected in our estimates is somehow tolerable. But the actual damage caused by floods is in all likelihood substantially larger than the temporary loss of economic activity, when one takes account of for example the cost of restoring damaged infrastructure, the opportunity cost of adults pulled away from productive work, and the loss of human capital as children are withdrawn from schooling and exposed to an unhealthy flooded environment. The longer-term costs of disruptions to human capital formation in particular could represent a multiple of immediate damages and death tolls from disasters (Antilla-Hughes and Hsiang, 2013).

A possible motivation for restoring vulnerable locations is to take advantage of the trading

[2]Below and in the Online Appendix we discuss in considerable detail the relationship between night lights and economic activity at the local level.

[3]We know that the floods (at least temporarily) displaced population, so it is highly unlikely that the effects we find are purely due to power outages.

[4]One possibility is that as some people permanently move to the safer parts of cities following the floods, newcomers (arriving for example from surrounding rural areas), take their place. However, we may expect such a pattern to increase population density - and therefore night light intensity - in the safer parts of flooded cities, which is not what we observe.

[5]We cannot rule out adaptation in the form of new or improved flood defenses. But most of the world's flooded urban areas are too poor to finance substantial flood defenses. Even where such defenses are built, they often represent a publicly funded solution, rather than private adaptation.

opportunities – and amenity value – offered by water-side locations. But we find that economic activity, as measured by night time lights, is fully restored even in low elevation locations that do not enjoy the offsetting advantages of being near a river or coast. Our results are also robust to excluding cities that are entirely less than 10m above sea level, where movement to higher ground within the existing urban area in response to a flood is not an option.

While we find that in most cities economic activity does not shift permanently in response to large floods, there is one important exception. We find a stronger and more lasting shift in response to floods in newly populated parts of cities. These areas, which we define as unlit during the first year that we observe night lights (1992), account for just 13 percent of the urban areas that we study. We find that in these recently populated urban areas, flooded areas show a larger and more persistent decline in night light intensity, indicating a stronger and more persistent relocation of economic activity in response to flooding. These results might be due to information updating about the riskiness of these locations; or it may be because fewer sunk investments were made in these locations.

In an attempt to distinguish learning from sunk costs (path dependence) we also compare areas that were unlit in 1992 to areas that were poorly lit in 1992. We think that it is plausible to assume that in both types of areas (the poorly lit and the unlit) there was very little sunk investment. But the two types of areas likely differ more in the information that people had about the risk of flooding. We think that unlit areas are more likely to have been uninhabited. Newly arrived residents to these areas would have less information available to them on the local flood history of the area. As a consequence there were more opportunities to learn from realized flooding in those areas. In contrast, poorly lit areas are more likely to have been inhabited for longer (prior to 1992). Residents of these areas would have a longer history of experience with local flooding to draw from, and therefore less scope to learn from a new flood realization. The results of this comparison suggest that the effect of flooding in unlit areas is significantly stronger and more persistent than in poorly lit areas. But there also seems to be an effect in poorly lit areas, which is a little larger than for the average area (which is better lit). Taken together, these results suggest that there may be a role for both learning and sunk costs, as our model suggests.

Our results are important for a number of reasons.

First, our findings contribute to a broader literature on urban responses to large scale shocks (e.g. Davis and Weinstein 2002). Much of this literature has focused on one-off shocks (e.g. due to war), whereas we examine the effects of large recurrent shocks. Moreover, our data allow us to combine global coverage with localized (within city) analysis. In the context of floods, Gallagher (2014) describes movements in the insurance market after rare flood events. While most of the literature has interpreted the recovery from shocks in a positive way, our finding of rapid recovery is not necessarily such a positive message, given the recurrent risk of flooding.

Second, the trend towards increased global urbanization is ongoing; presently, just over half of the world's population lives in urban areas, and this is expected to rise (United Nations 2008). As urbanization progresses, it is important to know whether cities have ways to adapt and avoid dangerous areas. Our results suggest that flooding poses an important challenge for urban planning because adaptation away from flood-prone locations cannot be taken for granted even in the aftermath of large and devastating floods.

Third, floods disproportionately affect poor countries. Given the scale of human devastation, and its potential to affect the formation of human capital (for example disruptions to study or health damages) this is an important issue for growth and development. Specifically, in developing countries planning and zoning laws and their enforcement are weak. Globally, it is estimated that more than 860 million people live in slums (Marx et al. 2013), which tend to develop on cheap land with poor infrastructure, including on flood-prone land (Handmer, Honda et al 2012). Our finding that low elevation areas concentrate much of the economic activity even in poor urban areas that are prone to extreme rainfall highlights the tragedy of the recurring crisis imposed by flooding.

Fourth, global warming and especially rising sea levels are expected to further exacerbate the problem of flooding. The threat of rising sea levels is not confined to developing countries and small island nations. Based on the extent of sea level rise that we now expect given cumulative emissions through 2015, Strauss, Kulp and Levermann (2015) identify 414 US municipalities that would see over half of their population-weighted area below future high tide levels. For continued, business-as-usual emissions scenarios, by 2100 this estimate rises to some 1,540 municipalities, which currently are home to more than 26 million people. Hallegatte et al. (2013) find that global average annual flood losses of US$6 billion in 2005 could reach US$52 billion by 2050. Under a scenario characterized by climate change and subsidence but no adaptation this amount could increase to US$1 trillion or more per year.

Work by Weitzman (2009) and Barro (2015) also argues that the costs of climate change could be high due to an increase in the odds of rare disasters. Our findings inform these studies in two ways. One positive message is that the economic impact of large-scale floods during our period of study was largely contained and localized. But the more negative implication is that a lack of adaptation could set the scene for more catastrophic events in the future, if sea levels rise and flooding becomes more widespread and devastating.

Understanding the extent to which people relocate away from stricken areas is important not only for assessing costs of flooding, but also for estimating the likely future global concentration of economic activity (Desmet and Rossi Hansberg 2013, Kahn and Walsh 2014). In a related paper Desmet et al (2015) analyze optimal migration responses to coastal flooding, with a view to the long run of gradual shifts in fundamentals. If sea levels rise considerably, and technological solutions (e.g. coastal barriers) are impractical, then people may eventually abandon some urban areas. If people interpret future flooding as indicating increased risk they may take measures to adapt. But our results indicate that even large and devastating floods do not permanently displace population, so it may prove challenging to convince people to leave as sea levels rise. In other words, in the medium run, adaptation may be painfully slow.[6]

Fifth, recovery assistance after flooding is an important part of international aid. While there are good reasons to restore some urban areas when they flood, we find that economic activity resumes even in poor locations with high risk of future flooding and without clear offsetting advantages (such as proximity to rivers and coasts). This suggests that those engaged in aid and

[6]While the period of our study is too short to separate the effect of recurrent shocks from permanent changes in locational fundamentals, one historical example may illustrate a painfully slow adaptation process. Dunwich, on England's eastern coast, was one of the country's largest towns in 1086. But it was subjected to repeated flooding and land losses, and its population, which peaked at around 5,000, declined by more than 95 per cent. However, this decline took centuries to unfold, and even today it remains a populated village (Sear et al. 2011, Patel 2009).

reconstruction should consider moving economic activity away from some of the most dangerous locations, in order to mitigate the risk of recurrent humanitarian disasters, and reduce the costs of bailing out future flood victims.

Finally, our results are relevant for discussions of the costly effects of path dependence (Bleakly and Lin 2012, Michaels and Rauch 2018). Our findings suggests that cities and parts of cities, which are built in flood-prone areas, may be locking in exposure to flood risk for a long time, even when circumstances and the global climate change.

The remainder of the paper is structured as follows. We outline a simple framework that allows us to consider how individuals may respond to a large flood in Section 2, discuss related literature in Section 3, describe the data in Section 4, present our main results in Section 5 and conclude in Section 6.

2 Theory

To frame our empirical investigation, we begin with a simple framework that allows us to consider how individuals may respond to a large flood. The framework, which is outlined in Online Appendix A, considers a discrete-time model, where people choose between two locations, which differ in their flood risk.

In the model, there are several factors that affect people's locational choice, including preferences over locations; the risk of floods multiplied by the net loss from floods (which also reflects the transfers that they receive in the aftermath of floods); and non-residential consumption. Rental costs may differ by location, and individuals moving - and those hit by floods - have to pay an additional cost. The model also includes an element of Bayesian updating: in any period a flood occurs (or does not), people increase (decrease) their assessment of the risk of future floods.

If people do update and some want to move from the risky area to the safe one, then the population or the prices in the safe area (or both) will increase relative to the risky area. If housing supply is inelastic, then the price of housing in the safe areas will increase relative to the risky areas, but there will be no change in the population ratio between the two locations and we might not detect any change in night time light intensity. But if housing supply is not completely fixed, then we expect both the price of housing and the population of the safe area to increase relative to the risky area so that updating will result in changes in economic activity as reflected by night time lights.

This simple model guides our empirical investigation in the following ways. First, we investigate the link between risk and low elevation locations. Anticipating and quantifying flood risk in the real world is a complicated endeavor, but we ask specifically how much more susceptible to large scale flooding are low elevation locations, compared to high elevation ones. Second, we ask whether people generally reside in riskier low elevation urban areas. In the model, the benefits to living in risky areas, or moving costs, might make it prohibitively expensive to relocate. Third, floods may cause people to leave the riskier areas because of either Bayesian updating, or because floods reduce the cost of moving to safer areas (relative to staying in the riskier ones). Our paper examines the extent to which large floods move economic activity away from risky areas towards safer ones. Fourth, the model suggests that there will be more updating in newly populated urban areas. In the empirical analysis we examine whether there is more relocation from riskier to safer areas in the aftermath of a flood in urban areas that concentrated no (measurable) economic activity until recently. Finally, we examine whether the presence of higher risk of flooding due to climatic factors shifts people towards safer areas, as our model suggests.

In addition to the issues raised by our model, an important question is whether rising sea levels and a changing climate will affect the aggregate global concentration of economic activity in flood prone areas. Of course, rising sea levels may increase the risk of flooding both in low elevation areas and in other areas that are currently safer. But it seems plausible to assume that at least in the near future, it is in the low elevation areas that rising sea levels will have a greater effect. In our analysis we will shed some light on the aggregate concentration of urban economic activity in low elevation areas over a longer period of time.

Another normative issue is how much ex-post transfers should victims receive, and in what form. In the model, higher compensation to flood victims makes movement away from risky areas less likely. From the perspective of a donor, if a property is frequently flooded, the costs of repeatedly paying compensation might be high. In developing countries where institutions are weak, finding private flood insurance may be a difficult challenge, especially for the poor. Ex-post disaster relief, including from large scale floods, is therefore a task that governments and non-government organizations around the world engage in from time to time. The main policy issue that we raise is whether it should be possible, in certain circumstances, to concentrate public reconstruction efforts towards safer areas, in order to avoid the high risk of recurrent disasters.

3 Related Literature

This paper contributes to a number of active strands of literature in urban economics, economic development, and the economics of disasters and climate change.

First, our paper speaks to the literature on the economic impact of floods and other natural disasters. Closely related to our study is Boustan, Kahn, and Rhode (2012), who look at the migration response to natural disasters in the US during the early twentieth century. They find movement away from areas hit by tornadoes but towards areas prone to flooding, possibly due to early efforts to build flood mitigation infrastructure. In a more modern setting, differences in migration responses by disaster type were also observed in research by Mueller, Gray and Kosec (2014) on determinants of out-of village migration in Pakistan. They find that heat stress, and not high precipitation or flooding, is associated with long-term migration. Elliott et al. (2015) study the effect of typhoons on night lights, and find a significant but short term effect. Also closely related is Hornbeck and Naidu (2014), who examine the Mississippi Flood of 1927, which led to out-migration of African Americans and a switch to more capital intensive farming. Our paper differs from most of these studies in its scope (we examine areas around the world, especially in developing countries), its timing (we examine much more recent floods), and its focus on urban areas and recurrent shocks. Our findings are also different, indicating that persistence of economic activity in risky areas is a concern.

A related strand of literature examines the updating of beliefs and changes in risk perceptions in the aftermath of natural disasters. Turner (2012) presents a model of Bayesian learning, where individuals update their risk assessments based on recent experience of disasters. Using data on US county level population, Turner finds evidence that population declines are more pronounced following a larger than previously experienced hurricane. Related papers include Cameron and Shah (2010) who find evidence of increased risk aversion among individuals in rural Indonesia who had over the past three years experienced first-hand a flood or an earthquake. Similarly, Eckel et al. (2006) note, based on interviews with a sample of Hurricane Katrina evacuees, that psychological factors such as levels of stress in the aftermath of an event influence individual risk aversion. Other related case studies include papers on the effect of Hurricane Katrina on the development of New Orleans and its residents (Glaeser 2005, Basker and Miranda 2014, Deryugina, Kawano and Levitt 2014), on the consequences of the Tsunami of 2004 (de Mel, McKenzie and Woodruff 2012), and Typhoons in China (Elliott et al. 2015). Also related are studies of the effect of flooding on house prices in the Netherlands (Bosker, Garretsen et al. 2015). Global studies include Hsiang and Jina (2014), who study the effect of cyclones on long run economic growth worldwide, and Cavallo et al (2013), who study the effect of natural disasters on GDP. Floods are generally more difficult to locate with great precision than, say, earthquakes or tropical storms.[7] Our innovation is to combine detailed inundation maps with information on elevation, which is well measured globally and at high resolution. This approach allows us to conduct within-city analysis at the global level - the first such analysis for floods that we are aware of.

Second, our study is related to the broader analysis of urban responses to large scale shocks. Two other recent papers that analyze the adaptation that takes place within cities in response to large scale shocks are Hornbeck and Keniston (2014), who analyze the recovery of Boston from the fire of 1872, and Ahlfeldt et al. (2015), who analyze the reorganization of Berlin in response to its division and reunification. Both are important case studies of large once-off shocks, whereas the shocks we study are more recurrent. Several other papers investigate urban destruction and recovery in the aftermath of wars, epidemics and other calamities (Davis and Weinstein 2002, Brakman et al 2004, Miguel and Roland 2011, Paskoff 2008, Beeson and Troesken 2006). Our study adds both a global perspective, since we analyze shocks around the world, but also a more localized perspective, since we examine what happens within cities. Whereas most of this literature has interpreted the recovery from shocks in a positive way, our finding that there is no shift in economic activity towards higher ground is not necessarily such a positive message.

Third, our study relates to a growing literature on urbanization in developing countries (Barrios et al. 2006, Marx et al 2013, Henderson et al 2015, and Jedwab et al 2015). Our paper also relates to a literature on the use of night lights data for empirical analysis of economic growth (Henderson et al 2012, Michalopoulos and Papaioannou 2014). The correlation between lights and economic activity in the cross section has long been noted (Donaldson and Storeygard, 2016), while Henderson et al. (2012) were perhaps the first to formally test the relationship between changes in lights and economic growth, using GDP data at the national level. Since then numerous studies have used the lights data as a proxy for economic activity or prosperity at a local (sub-national) level (Ghosh et al. 2013 and Donaldson and Storeygard 2016, summarize

[7] Of course, flooding is sometimes the result of tropical storms – as is the case for 10 of the 53 large flood events included in our sample. These storms include hurricanes, cyclones, and typhoons, which are different names given to the same type of tropical storm that occurs in different parts of the world. While wind field models, combined with detailed storm track data, can allow precise estimation of the location and intensity of winds associated with tropical cyclones (see e.g. Strobl 2011), this method may not identify the precise extent of associated flooding.

some of these applications). A number of studies have noted truncation of the night lights data, both at the lower and upper ends of the spectrum, resulting in weak predictive power in settings with little economic activity (e.g. rural areas, particularly in developing countries - see Chen and Nordhaus 2011) and where lights are saturated (e.g. in urban areas in rich country settings - see Mellander et al. 2015). However, recent verification exercises on the lights-GDP relationship using regional, city and prefecture level data (e.g. Hodler and Raschky 2014, Storeygard 2016) show a close correspondence with the central lights-GDP estimate from Henderson et al. (2012). Further discussion of this literature is included in Online Appendix B.

Local GDP is often preferred as a measure of local economic activity, even though it is measured for coarser geographies than night lights and often suffers from measurement error. To further examine the relationship between local GDP, night lights, and floods using within-country variation, we focus on the case of India, and we report our results in Online Appendix C. In summary, the relationship between night lights and GDP (both in logarithms) in India is strong both in a cross-section and in a panel of districts. This relationship is of a similar order of magnitude to what the above-mentioned literature finds. At the same time, our results show that variation within only one country gives an imprecise relationship between floods and both GDP and night lights. Nevertheless, the confidence intervals that we find using the Indian data are quite wide, possibly because Indian districts are large and populous, while the flooding measures we discuss below are more geographically precise. In the rest of the analysis we therefore focus on night lights as an outcome, since they can be measured at a fine spatial scale and with greater consistency across the world's cities, even when GDP data is unavailable or poorly measured.

We acknowledge that a limitation with the use of night lights is that the effect of disasters on power plants may be hard to distinguish from the destruction of buildings and other infrastructure. For this reason, we are cautious in interpreting the magnitudes of our regression coefficients, despite the close relationship between night lights and GDP at the local level, as discussed above.

Finally, our paper also relates to the literature estimating the costs of climate change and sea level rise (Hanson et al 2011, Hallegatte et al 2013, Desmet et al. 2015, Tessler et al. 2015). Coastal cities feature prominently in this large literature, given their current and future exposure to flooding in particular. One important factor in assessing the long term impact of flooding is adaptation, or the degree to which people move away from environmentally dangerous locations. Our study suggests that adaptation responses may be inadequate, and consequently the costs of increases in future flooding may be higher than anticipated.

4 Data

The dataset that we compile for our empirical analysis comprises data on flood locations, physical characteristics of locations (including elevation and distance to rivers and coasts), precipitation, urban extents, population density and night light intensity, all mapped onto an equal area one kilometer-squared grid covering the entire world (using the Lambert cylindrical equal area projection). The data are drawn from a number of sources as detailed below.

Floods

The primary data for our analysis are the flood maps that we use to identify flooded locations. These come from the Dartmouth Flood Observatory (DFO) archive (Brakenridge 2015). The DFO database includes information on the location, timing, duration, damage, and other outcomes for thousands of flood events worldwide from 1985-2015. These data were compiled from media estimates and government reports. While we use this database to derive general statistics about floods, our paper is focused mostly on a subset of floods for which DFO provides detailed inundation maps (which we discuss in more detail below). These maps were produced predominantly for the period 2003-2008, and even for that period they do not cover all large floods (see below).[8]

In this paper we focus on the most devastating flood events, which (according to DFO) displaced at least 100,000 people each, to which we sometimes refer in short as "large floods".[9] Our focus on large floods with available inundation maps, left us with a sample of 53 large flood events that affected 1,868 cities in 40 countries worldwide from 2003-2008. This sample represents a majority (55 percent) of displacement-weighted events, which took place during this period, according to the DFO database (see Table 1). Table 1 also provides a count of events displacing more than 100,000 people per year, based on the complete DFO database. The table suggests that the period of our main sample (2003 - 2008) was one with a particularly high number of large flood events. The higher frequency of large floods during our period of analysis compared to other periods could reflect an actual change in flood devastation over time and/or more intensive documentation by DFO, as suggested by the availability of detailed inundation maps for this period.

The locations of the large flood events in our sample are illustrated on the world map in Figure 1. The map shows all urban areas in the world (in light grey). City sizes are inflated – even more so for flooded cities – in order to make them more clearly visible on a map of the entire world. The map shows locations that were affected by large floods, with darker shades representing higher frequencies of flooding. The number of floods in the legend refers to the number of years during our main sample period (2003-2008) in which each city was affected by a flood that displaced a total of 100,000 people or more. As the map illustrates, large urban floods are especially common in South and East Asia, but they also afflict parts of Africa and the Americas.[10]

The patterns that the map reveals are not coincidental. Large-scale flooding usually involves heavy precipitation, so it mostly occurs in tropical or humid sub-tropical areas. Of course other areas are not immune from large floods due to tropical storms (e.g. hurricane Sandy in the New York area in 2012) and tsunamis (the 2011 Tsunami in Japan), which fall outside our period of analysis. Large-scale urban flooding also typically occurs more often in densely populated areas, such as the basins of the Ganges, Yangtze, and Yellow rivers. Finally, large-scale flooding more

[8]Some maps for earlier and more recent events exist on the DFO website, which were less detailed and/or not fully processed and were therefore not directly comparable.

[9]For comparability we used displaced as indicator of intensity instead of the traditional 1 in 10 year flood, 1 in 100 year flood, etc. Our choice is motivated by our interest in floods that are devastating to human lives in an absolute sense, and not just relative to local precipitation patterns. We also note that DFO 'displaced' figures are an attempt to estimate the number of people who were evacuated from their homes due to floods. These estimates are not exact, and may cover both temporary displacement and events where people's homes were permanently destroyed. Even though the overall human cost from this sample of floods is enormous we refer to them as 'large floods' to emphasize that the effect on each affected flooded city need not be such.

[10]Europe and Australia are also not immune from large floods, but during the period that we examine they were not affected by large floods covered by DFO inundation maps.

commonly occurs in developing countries, where flood defences are weaker (or non-existent). But again the examples mentioned above, and the large flooding events in Louisiana and Florida (shown on our map) show that rich nations are by no means immune.

The DFO flood maps are constructed from satellite images. Flood outlines based on satellite imagery are translated by DFO into Rapid Response Inundation maps showing the extent of area that is flooded – often for different days during a given flood event. It is very likely that the DFO maps understate the true extent of flooding in each event, in part due to cloud cover obstructing the view from the satellites, or in part because the extent of flooding is not documented for every point in time. Furthermore, as explained above, some large flood events do not appear on any inundation maps. For this reason, cities that never appear in the database might nonetheless be flooded in a given year, and we restrict most of our analysis to cities that appear as flooded in at least one inundation map. Since we are concerned that the documented high water marks of floods within cities might understate the actual one, we do not use information on the extent of flooding within cities. Instead, we define a city as flooded in a given year if at least one gridpoint within it is flooded (by a large flood) in that year. An example of one of our flood maps, in this case the flooding associated with Hurricane Katrina in the city of New Orleans and its environs in 2005, is given in Panel A of Figure 2.[11]

Several types of extreme events caused the 53 large floods that we analyze: heavy precipitation (42 events, of which 12 are due to monsoonal rain), tropical storms (10 events), and a tidal surge (the 2004 Tsunami). Since tropical storms can cause damage from wind as well as flooding, we discuss regression results showing that precipitation, rather than wind damage, is likely the main driver of our results. Taken together, DFO estimates suggest that the 53 flood events displaced almost 90 million people, of which 40 million were displaced in the 2004 floods in India and Bangladesh.

Night-time light data

To identify the economic effects of floods at a fine spatial scale, we use data on night lights as a proxy for economic activity. These data are collected by satellites under the US Air Force Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS). The satellites circle the earth 14 times each day, recording the intensity of Earth-based lights. Each satellite observes every location on the planet every night at some instant between 8:30 and 10:00 pm local time (Henderson et al. 2012). NOAA's (National Oceanic and Atmospheric Administration) National Geophysical Data Center (NGDC) processes the data and computes average annual light intensity for every location in the world. An average 39.2 (s.d. 22.0) nights are used for each satellite-year dataset. Light intensity can be mapped on approximately one-kilometer squares and are thus available at much higher spatial resolution than standard output measures. The data are available annually from 1992 - 2013. For some years more than one dataset is

[11] This and other DFO inundation maps are available as image files from `http://floodobservatory.colorado.edu/Archives/MapIndex.htm`. Different color codes are used in these images to indicate flood extents at different points in time (and also across flood events). Our approach to digitizing these images captures the mainly red and pink hues used by DFO to show the flooded areas. Specifically, we use the following code to capture flood extents: $((\text{“}MAP_{id}.jpg - Band_1\text{”} > 240)\&(\text{“}MAP_{id}.jpg - Band_2\text{”} < 180)\&(\text{“}MAP_{id}.jpg - Band_3\text{”} < 180))|((\text{“}MAP_{id}.jpg - Band_1\text{”} > 80)\&(\text{“}MAP_{id}.jpg - Band_2\text{”} < 10)\&(\text{“}MAP_{id}.jpg - Band_3\text{”} < 10))|((\text{“}MAP_{id}.jpg - Band_1\text{”} > 245)\&(\text{“}MAP_{id}.jpg - Band_2\text{”} < 215)\&(\text{“}MAP_{id}.jpg - Band_3\text{”} < 215))$. We georeferenced each map in ArcGIS to identify its location, enabling the creation of a digital shape file identifying locations affected by each of the events included in our sample.

available. Where this is the case, we chose datasets so as to minimize the number of different satellites used to collect the data.[12]

While these data are well suited for studying local economic developments on a global scale, they are not without limitations. One concern is that the use of different satellites for different years may result in measurement error. We address this concern by including year fixed effects in all our specifications. Another limitation is that the lights data range from 0-63, where 63 is a top-coded value. While imperfect, we note that most of the floods that we analyze affect developing countries where much of the light activity is below the top-coded level, and this point emerges clearly from the descriptive statistics.[13] For our main sample of cities affected by at least one of the large floods the proportion of top-coded cells varies from just 1.4 percent to 5 percent over the period 2003-2008. Lastly, we note that the lights datasets also include light related to gas flares. Our data processing included the removal of gas flaring grid points from the data (as in Elvidge, Ziskin, Baugh et al. 2009).[14]

An example of a light intensity map for the city of New Orleans and its environs is provided for the years 2004, 2005 and 2006 in Panels B, C and D of Figure 2. The three panels illustrate how light intensity in the city looked in the year prior to the flood caused by Hurricane Katrina (2004 - Panel B), in the year of the flood (2005 - Panel C) and in the year following the flood (2006 - Panel D). One can see a distinct dimming of the lights in the year of the flood (2005 - Panel C), relative to the previous year (2004 - Panel B). This pattern is particularly pronounced in the North-Eastern parts of the city, corresponding to the worst affected areas, according to the DFO flood map in Panel A of Figure 2. The light intensity map in Panel D of Figure 2 (2006) appears to show a restoration of light intensity in the city to levels that are fairly close to those observed prior to the flood. The example of Hurricane Katrina also demonstrates that despite the top-coding and any measurement error, even in a rich country such as the US, the effects of floods are visible from light activity. Nevertheless, we should emphasize that New Orleans is atypical of our data; the vast majority of the large flood events that we analyze take place in poorer countries.

Urban extents

We focus our analysis on urban areas, as defined by the Global Rural-Urban Mapping Project (GRUMP) urban extent grids from the Center for International Earth Science Information Network (CIESIN) at Columbia University, for the year 1995 (GRUMPv1, 2015). To keep the analysis tractable, we treat these boundaries as fixed. Urban extents are defined either on the basis of contiguous lighted cells using night-time light data or using buffers for settlement points with population counts in 1995 greater than 5,000 persons (CIESIN 2011). For our analysis we split urban areas that span multiple countries into distinct units, so that we can assign each urban area to the country in which it lies. This gives us a total of 34,545 urban areas. We can assess to what degree lighting correlates with population growth. We estimate regressions where the dependent variable is population growth from 1995-2000 (once in levels and once in

[12] We use data from Satellite F10 for 1992-1993; from Satellite F12 for 1994-1999; from Satellite F15 for 2000-2007; from Satellite F16 for 2008-2009; and from Satellite F18 for 2010- 2012.

[13] Aside from top-coding, the specification of the light intensity measure involves low levels of light set to zero. This might be a further source of measurement error, although is less likely a concern for our analysis, given the focus on urban areas. In our data there are only about 5.5 percent of observations coded zero, which is not surprising given how urban extents are identified in the GRUMP data (see subsection).

[14] Only 0.0057 of gridpoints fall in this category.

logarithms), and the regressor of interest is the level of lights in 2010 for areas that were unlit in 1992, since by construction the level of light in the base period here is zero. In both regressions we control for country fixed effects and cluster the standard errors by country. Both regression estimates are positive and statistically significant. However, for our main specifications, we restrict our analysis to urban areas that were hit at least once by a large flood in our data - a sample of 1,868 cities. We also take population density data at one-kilometer square resolution from the same source.[15]

Other data

In our analysis we also use data on elevation (in meters above sea level), which are taken from the US Geological Survey (USGS), and data on distance to (nearest) coasts and rivers (in kilometers) from the same source. The elevation data come from the GTOPO30, a global digital elevation model.[16] The data on elevation are spaced at 30-arc seconds and cover the entire globe. As with all our data these are projected from geographical coordinates to an equal area projection (Lambert cylindrical equal area) and fitted onto our 1 square kilometer grid.

We also obtain monthly precipitation data on a 0.5 x 0.5 degree cells resolution from the Climatic Research Unit (CRU) at the University of East Anglia (Jones and Harris, 2013).[17] We use these data to construct extreme precipitation indicators for locations that experience monthly precipitation in excess of 500mm (or 1000mm) at least once in a given year.[18] Although extreme precipitation is by no means a perfect predictor of flooding for a particular location, it has the advantage of being an exogenous source of variation in flood location and timing. We use these extreme precipitation indicators as alternative explanatory variables to mitigate against endogeneity concerns with respect to our flood indicator.

5 Results

We begin with a cross-sectional analysis of flood exposure and the concentration of economic activity, by location, using the full sample of all urban areas in the world.

We first examine the nature of global urban flood risk, using information from our inundation maps. In Table 2 we test how exposure to large urban flooding (events that displace at least 100,000 people) depends upon location characteristics. We regress a measure of the frequency of flooding on an indicator for low elevation (being no more than 10 meters above sea level) and controls, for the full sample of all urban areas. The regressions reported in Table 2 are of the

[15] We use population density data adjusted to match UN total estimates ("ag") not national censuses ("g").

[16] GTOPO30 is the product of collaboration among various national and international organizations under the leadership of the U.S. Geological Survey's EROS Data Center. See `https://lta.cr.usgs.gov/GTOPO30`.

[17] A 0.5 x 0.5 degree cell measures approximately 60km x 60km at the equator. At higher latitudes the East-West dimension of these cells becomes smaller. For example, the highest latitude city in our main sample is located at about 39 degrees North. At this latitude, a 0.5 x 0.5 degree cell measures roughly 42km (East to West).

[18] These are relatively rare events. About 15 percent of urban gridpoints in the world have experienced monthly precipitation exceeding 500mm at least once during the period 1992-2012, while only 1.2 percent of urban gridpoints in the world experienced monthly precipitation exceeding 1000mm at least once during the period 1992-2012.

following form:

$$FloodFreq_{ik} = \beta_{11} + \beta_{12}(Elev < 10m)_i + \beta_{13}River_i + \beta_{14}Coast_i + Country_k + \epsilon_{ik}. \quad (1)$$

The left hand side represents the frequency of flooding for a given location, measured as the number of years during our main sample in which each location is hit by at least one large flood event, divided by the length of the sample.[19] The sample here is all urban gridpoints in the world, based on the 1995 GRUMP definitions, discussed above. The right hand side includes dummy variables for locations that are less than 10m above sea level ($Elev < 10m_i$), less than 10km from the nearest river ($River_i$) or coast ($Coast_i$). Columns (5) to (8) include country fixed effects. To account for spatial correlation, we cluster the standard errors by country, which is a more conservative approach than that taken in most of the literature.

We find that globally, urban flood risk by this measure is around 1.3 percent per year for areas at least 10m above sea level (based on the intercept of Column 1). Low elevation areas are substantially more likely to be in a city affected by flooding. For urban areas less than 10m above sea level, the annual risk of being hit by a large flood rises to about 4.9 percent[20], i.e. an annual probability of almost one in 20 of being hit by a flood that displaces at least 100,000 people. That is likely an underestimate of global flood risk, since there may on (rare) occasions be more than one event per city per year, and also because we only have inundation maps for fewer than half the events in our sample period (January 2003 - May 2008). At the same time, it is possible that the period we study may have been especially bad. From the information in Table 1 it does appear that 2003-2008 was a period with a relatively high number of large flood events.

Looking beyond the means, cities close to coastlines or rivers do not appear to face significantly higher flood risk than other urban areas, according to our data, although the estimates for rivers are non-trivial in magnitude and marginally significant; see Columns (2)–(4), and (6)–(8) of Table 2. We also note that the estimated effect of elevation is a bit less precise when we control for country fixed effects, although the magnitude is fairly similar to the estimates without fixed effects.

We next investigate whether economic activity concentrates disproportionately in flood-prone urban areas – specifically locations that are low elevation, and those that are exposed to extreme precipitation, or both. To investigate this question, we regress light intensity at each gridpoint (in 2012) on an indicator for low elevation (being less than 10m above sea level), an indicator for being exposed to high levels of precipitation in a single month, an interaction of the two, and controls, for the full sample of all urban areas. The precise specifications reported in Table 3 are of the following form:

$$ln(Y_{ilk}) = \beta_{21} + \beta_{22}(Elev < 10m)_i + \beta_{23}Precip_l + \beta_{24}Precip_l \times (Elev < 10m)_i + Country_k + \epsilon_{ilk}, \quad (2)$$

[19]In practice, we only have data on floods up to May 2008, so that our sample spans five years and five months. To capture the likelihood of flooding per year for a given location, the dependent variable here is generated by dividing the number of years (2003-2008) in which a location is hit by a large flood, by the length of the sample, i.e. five years and five months (or 65/12).

[20]Summing the intercept and the coefficient on the low elevation indicator, i.e. $0.013 + 0.036$.

where the left hand side is the natural log of mean light intensity (in 2012) at each gridpoint i (located in grid cell l, in country k).[21] The right hand side includes the low elevation indicator, an indicator for areas that have experienced extreme precipitation in a single month at least once in the period 1992-2012, and the interaction of these two indicators. Each specification also includes country fixed effects. Columns (4), (7) and (10) add city fixed effects. Columns (3), (4), (6), (7), (9) and (10) add river and coast dummies, defined above. We include three different versions of the extreme precipitation indicator: These indicate locations that experience more than 1000mm (500mm) of precipitation in a single month at least once in the period 1992–2012, or monthly precipitation of 500mm or more at least twice during that period.

The results reported in Table 3 show that low elevation areas are more lit relative to country averages – as indicated by the coefficients on the elevation dummy (in the first row), which are all positive (and significant in Columns 1–3, 5–6 and 8–9). These results suggest a greater concentration of economic activity in low elevation areas. These areas are also, as we might expect, more vulnerable to large floods – i.e. they get hit more frequently – as demonstrated in the previous analysis (described above and reported in Table 2). Even in the specifications that include city fixed effects (Columns 4, 7 and 10), the coefficients on the elevation dummy are positive, although only in one of them (Column 4) is the estimate precise at conventional levels.

Looking at the interactions between the low elevation indicator and the extreme precipitation indicators (in Columns 2–10), we find again that even in areas that experience monthly precipitation exceeding 500mm at least once, low elevation areas are still more lit relative to country averages (Columns 5 and 6), and no less lit than city averages (Column 7). For areas that experience monthly precipitation exceeding 500mm at least twice, low elevation areas are again found to be more lit relative to country averages (Columns 8 and 9), and no less lit than city averages (Column 10). For areas exposed to monthly precipitation exceeding 1000mm at least once, low elevation areas are no less lit than country or city averages (Columns 2–4). All of these findings are also robust to including controls for proximity to the nearest river or coast (Columns 3, 4, 6, 7, 9 and 10).

Taken in the aggregate, the results in Table 3 indicate that globally, urban economic activity as measured by night light intensity is concentrated disproportionately in low elevation areas, which are more prone to flooding, and this is even true in regions that are prone to extreme rainfall.

We also experiment with variations of the specifications presented in Table 3 to investigate if certain types of countries are better at avoiding concentrating economic activity in low elevation, flood-prone locations. Specifically, we examine the effects of national income and democracy on the location of economic activity, as proxied by the intensity of night lights. The results are reported in Table A1, and they suggest that democracies (classified as having a Polity IV score in 2008 greater than or equal to five) are better at avoiding concentrating economic activity in flood-prone locations. On the other hand, we find that richer countries are not significantly different from poorer ones in avoiding flood-prone areas.

We next move to the panel analysis of the local economic impacts of large urban floods. Here the dataset consists of a panel of gridpoints (i) located in city (j) in country (k) with time dimension

[21] Precipitation is measured at the grid cell level, where cells measure 0.5 x 0.5 degrees.

(t). In order to focus the analysis on changes over time within areas that are prone to large-scale flooding, we restrict the sample to gridpoints in cities that are affected by a large flood at least once in the sample, excluding other cities, many of which may be qualitatively different and may never flood.

In our analysis, we use variation over time in the occurrence of flooding (and later, also extreme precipitation), by estimating equations of the form:

$$ln(Y_{ijkt}) = \beta_{31} + \beta_{32}Flood_{jt+s} + Gridpoint_i + Year_t + Country_k \times Trend_t + \epsilon_{it}, \qquad (3)$$

where Y_{ijkt} is mean light intensity in gridpoint i (located in city j, in country k) in year t and $Flood_{jt+s}$ is a flood dummy, indicating whether or not city j was hit by a large flood in year $t+s$. We include gridpoint and year fixed effects and country-specific trends.[22] As a robustness check, we re-estimate our regressions with dynamic panel specifications, which include a lagged dependent variable, instrumented by a second lag (Arellano and Bond 1991).[23]

Estimation results of equation 3 are reported in Table 4. Column (1) of Table 4 shows that a flooded city darkens in the year in which it is flooded. This effect is also present when controlling for the instrumented lagged dependent variable, in Column (4). The magnitude is similar in both specifications, at -0.021 and -0.023, respectively, and statistically significant at the five percent level in both cases. This can be interpreted as a 2.1 (or 2.3) percent reduction in average light intensity of urban gridpoints in the year of the flood. Although we note that this represents the average effect for all gridpoints in a flooded city, including areas that are unaffected by the flood. Flooded gridpoints are likely to experience greater changes in light intensity, but the quality of our flood maps only allows us to use variation in flooding at the city level – and later interact it with measures of flood-proneness due to low elevation.

How should we interpret the magnitude of these estimates? Henderson, Storeygard and Weil (2012) relate the change in lights to changes in economic activity. Their main estimate of the GDP to lights elasticity is between 0.3 and 1. Based on this range of estimates, the percentage reductions in light intensity associated with floods that we estimate could be interpreted as percentage reductions in economic activity, although the relationship between lights and economic activity estimated by Henderson et al. (2012) could be different at the local level. A number of papers have estimated the lights-GDP relationship using sub-national data, finding a similar elasticity of GDP to lights as Henderson et al. (2012). The evidence in Online Appendix C is also broadly consistent with Henderson et al. (2012).

Our estimates of the effects of flooding focus on the reduction in economic activity captured by the night lights. These do not include the costs of rebuilding houses and other infrastructure. In fact, if reconstruction efforts temporarily increase night time lights – and these efforts occur in the same year as the flood – this could mask the true economic impact of the flood, which could be larger than we estimate.

[22] These country trends account for the differences across the world in the changes in lit areas.

[23] Our main specification is a panel with fixed effects, which controls for country-specific trends. Nevertheless, since our dataset is a short panel of the type "small T, large N", this may bring up concerns of a potential Nickell bias (Nickell 1981). To mitigate this potential concern we use the Generalized Method of Moments (GMM) estimator developed by Arellano and Bond (1991). As we explain in the footnotes to the relevant tables, our implementation uses the second lag of the dependent variable as an instrument for the first lag.

Some readers may be concerned about possible endogeneity of our flood indicators, with respect to economic activity, given that we identified large floods as those that displaced at least 100,000 people. To mitigate such concerns, Table 4 also includes results using extreme precipitation indicators, in place of the flood dummy. The extreme precipitation indicators, $Precip_{ljt}$, indicate whether or not grid cell l in city j experienced monthly precipitation exceeding 500mm (or 1000mm) in year t. These are not common occurrences; about 15 percent of urban gridpoints in the world have experienced monthly precipitation exceeding 500mm at least once during the period 1992-2012, while only 1.2 percent of urban gridpoints in the world experienced monthly precipitation exceeding 1000mm at least once during the period 1992-2012.

The results of these specifications are reported in Columns 2–3 and 5–6 of Table 4. The effect of an episode of monthly precipitation exceeding 500mm is similar in magnitude to that of a large flood, at between −0.025 and −0.027 (Columns 2 and 5). The rarer event of monthly precipitation exceeding 1000mm has a substantially larger effect on light intensity in affected cities, leading to average dimming of between −0.080 and −0.083 (Columns 3 and 6). The coefficients on the extreme precipitation indicators are statistically significant at the one percent level in each of these specifications.[24]

We also repeat our main analysis at the city-wide level, aggregating the data to city level, with observations weighted by city population. At the city level the specification becomes:

$$ln(Y_{jkt}) = \beta_{41} + \beta_{42} Flood_{jt+s} + City_j + Year_t + Country_k \times Trend_t + \epsilon_{jt}. \qquad (4)$$

which is essentially the same as 3 but with city fixed effects now replacing gridpoint fixed effects.

The estimation results for this specification, now across rather than within cities, are reported in Table A2. We find a similar pattern of results as before. The effect of a large flood on city-wide light intensity is still statistically significant at the five percent level, albeit slightly smaller in magnitude at between −0.017 and −0.019 (Columns 1 and 4). Episodes of extreme precipitation also reduce light intensity at the city-wide level, with the effects significant at the one percent level in each case (see results in Columns 2–3 and 5–6 of Table A2).

We next investigate patterns of recovery of urban economic activity in the aftermath of floods. In Table 5 we report the results of estimating versions of Equation 3 including lagged versions of the flood indicator (up to $t-4$) – i.e. testing the effects of large floods on light intensity at up to four years after the flood. Columns (1) and (6) of Table 5 repeat Columns (1) and (4) of Table 4 for ease of comparison. The remaining Columns of Table 5 show that the statistically significant impact of the flood on light intensity in the year of the event disappears at $t-1$ and does not reappear at further lags.[25] These results indicate that urban economic activity is fully

[24] We do not use extreme rainfall to instrument for flooding, because extreme rainfall can adversely affect a city's economic fortunes even if far fewer than 100,000 people are displaced. In technical terms, this amounts to a violation of the exclusion restriction. This problem, coupled with the relative rarity of large floods (a small first stage), implies that 2SLS estimates of the effects of large floods using extreme rainfall as an instrument are much larger than the OLS estimates that we report in the paper. We therefore prefer to focus on the OLS estimates, which we find more credible.

[25] In the Online Appendix we include an event study (Figure A1) showing the coefficients reported in Table 5, and two additional years before the flood, not shown in the table. The figure illustrates that the shock of the flood is concentrated in the year in which it took place.

restored just one year after a large flood strikes a city. This pattern of rapid recovery is also found for cities affected by episodes of extreme precipitation (see results in Tables A3 and A4). We also test for recovery at the city-wide level, running versions of Equation 4 with lags of the flood indicator (results presented in Table A5). Again, we find a similar pattern, with the effect of the flood on city-wide light intensity disappearing after just one year.

We next consider heterogeneity of floods' effects within cities. In particular, we are interested in the differential effect of large floods by elevation. We test this by interacting the flood indicator with an elevation band indicator. Returning to the panel of gridpoint-years, the regression specification now becomes:

$$ln(Y_{ijkht}) = \beta_{51} + \Sigma_h \beta_{52h} Flood_{jt+s} \times Elevation_h + Gridpoint_i + Year_t + Country_k \times Trend_t + \epsilon_{it}, \tag{5}$$

where $Elevation_h$ is a dummy for elevation band h. In practice we interact the flood indicator with an indicator for urban locations that are less than 10m above sea level (and an indicator for areas that are 10m or more above sea level). The results of these specifications are reported in Table 6. As before, these regressions include year and gridpoint fixed effects, as well as country-specific trends (in Columns 1–6). In Columns (7) to (8) we replace the country-specific trends with city-specific trends, to account for different city-specific changes in light intensity over time. We also re-estimate our main regressions using the dynamic panel specification described previously (results reported in Columns 4–6).

The results in Columns (1) and (4) show that low elevation areas within cities are hit harder than other areas when a city is struck by a large flood. The effect on light intensity for areas less than 10m above sea level is estimated at between -0.027 and -0.028. This effect is even slightly stronger (-0.030) when accounting for city-specific trends in Column (7). These effects are statistically significant at the one percent level. The estimated effects for areas more than 10m above sea level are smaller in magnitude, and not statistically significant in Columns (1) and (4). When we include city-by-year fixed effects, the finding that low elevation areas are hit harder does not survive this more demanding specification, which only uses variation in elevation within cities exposed to floods in a given year. Similar specifications, where instead of elevation we interacted floods with indicators for distance to nearest river or coast, found no such significant pattern of heterogeneity.[26]

The effects at low elevation are even stronger when using extreme precipitation to identify affected locations. Specifically, the results in Table A6 show that light intensity for locations less than 10m above sea level is reduced by up to -0.122 in years with episodes of monthly precipitation in excess of 1000mm.[27]

The interaction of the flood indicator with an indicator for low elevation areas captures the impact of floods on the riskiest parts of cities. As we might expect the effects identified for low-elevation areas are both stronger in magnitude and more precisely estimated than the average effects reported in Table 4. However, as pointed out previously, the effects we report are still

[26] These alternative specifications are not reported in this version of the paper. Results available on request.

[27] A similar analysis for episodes of monthly precipitation in excess of 500mm did not find any significant heterogeneity by elevation. Results not reported here, but available on request.

average effects across all gridpoints (in this case, all gridpoints less than 10m above sea level) in affected cities. The effects for gridpoints experiencing the worst actual flooding could well be stronger again than those reported here.

We know that the floods in our sample (at least temporarily) displaced population, so it is highly unlikely that the effects we find are purely due to power outages. The heterogeneous impacts by elevation that we identify here would seem to support this conclusion. Damage to energy infrastructure would likely reduce light across the city, in both its higher and lower elevation neighborhoods. The heterogeneous impact is suggestive that lights within cities indeed correlate with local economic activity.

Table 6 also shows the pattern of recovery following a flood event for urban locations at different elevations. Again the effects of the flood disappear just one year after the event, even for the harder hit low elevation areas (those less than 10m above sea level) – as demonstrated by the results in Columns (2), (5) and (8) of Table 6. The positive and significant coefficients on the interaction of the $flood_{t-2}$ indicator with the $elev < 10m$ indicator in Columns (3), (6) and (9) of Table 6 indicate some over-shooting in the recovery of low elevation areas. Two years after a flood event, the light intensity in the hardest hit areas of flooded cities is above its (country-specific or city-specific) trend. A similar specification, where instead of the flood indicator we interacted elevation with an indicator of extreme precipitation, found a temporary increase in light intensity one year after experiencing monthly precipitation of 1000mm or more. However, this increase disappears in the following year, which might have to do with aid and reconstruction efforts.[28]

The pattern of results presented in Table 6 – both the heterogeneous impacts by elevation and the rapid recovery of even the harder hit low elevation areas – is robust both to the exclusion of locations within 10km of rivers and coasts (see Table A7) and to the exclusion of cities that are entirely less than 10m above sea level (see Table A8). The rapid recovery of low-elevation areas, even when excluding locations within 10km of rivers and coasts, suggests that this recovery process is not simply being driven by the attractiveness or greater productivity of water-side locations. Similarly, the rapid recovery of low-elevation areas is observed even when we exclude (the small number of) cities that are entirely less than 10m above sea level – where relocating economic activity to higher ground (within the city) is unfeasible. In other words, people in flooded areas typically have the option to move to higher ground even within their own city, but either cannot afford the move or choose not to do so. This result supports our conclusion that this process is being driven by an important economic problem, and not simply a technical or geographic constraint on adaptation.

We also test for differential effects of floods within cities by newly populated versus existing locations. These specifications are similar to Equation 5, but instead of an elevation dummy, here we interact the flood indicator with a dummy for newly populated areas (locations that were dark, with $lights = 0$ in 1992) and a dummy for existing areas (those that were not dark in 1992). The results of these regressions are reported in Table 7.

The results show larger and more persistent impacts of flooding on newly populated areas, with the negative effect of the flood on lights persisting, and intensifying, for about three to four years after the event. As before, we have also tested versions of the specification reported in Table

[28]This alternative specification is not reported in this version of the paper. Results are available on request.

7 including city-by-year fixed effects. In this case, the main results reported in Columns 1-3 of Table 7 do survive the more demanding specification, either at the 5 percent or 10 percent level.[29] These findings are in line with the predictions of our model; because updating decreases in t (where t is the length of the sample of information that an individual has on past flooding), we expect that there will be more updating in newly populated urban areas.

The persistent negative effect of floods in areas that were not lit in 1992 stands in contrast with the return to pre-existing conditions elsewhere in cities. The negative effect of floods in new areas is (from the year after the flood) roughly an order of magnitude larger than in the areas that were settled in 1992. And the negative effects are significant at the 5-10 percent levels for four years. Even five years after the flood, and despite the limitations of our short panel and conservative inference, the point estimate of the flood is still larger than in the year of impact. These results suggest that there is some adaptation to floods, but only in areas that are newly populated, where the risk of flooding may not have been fully realized, and substantial sunk investments may not yet have been made.

We also compare newly lit areas with areas that were poorly lit in the first year we observe night lights (1992). For comparison the mean night lights intensity in urban areas in our sample in 1992 was close to 21. We think that it is plausible to assume that in both types of areas (the poorly lit and the unlit) there was very little sunk investment. But the two types of areas likely differ more in the information that people had about the risk of flooding. We think that unlit areas are more likely to have been uninhabited. Newly arrived residents to these areas would have less information available to them on the local flood history of the area. As a consequence there were more opportunities to learn from realized flooding in those areas. In contrast, poorly lit areas are more likely to have been inhabited for longer (prior to 1992). Residents of these areas would have a longer history of experience with local flooding to draw from, and therefore less scope to learn from a new flood realization. The results of this comparison suggest that the effect of flooding in unlit areas is significantly stronger and more persistent than in poorly lit areas. But there also seems to be an effect in poorly lit areas, which is a little larger than for the average area (which is better lit). Taken together, these results suggest that there may be a role for both learning and sunk costs, as our model suggests. These additional specifications are reported in Appendix Figure A2.

Finally, we examine whether rising sea levels are gradually shifting aggregate urban economic activity away from flood prone areas irrespective of the incidence of large-scale floods. According to the IPCC mean global sea levels rose about 20cm over the 20th century (Myhre et al 2013). In fact, we find that the share of economic activity concentrated in areas less than 10 meters above sea level is fairly constant and close to 10 percent throughout the 20 years from 1992 to 2012, suggesting that gradual adjustment to climate change is not what we observe in our data.[30]

[29] These results are not reported in this version of the paper, but are available on request.

[30] In a separate exercise, not reported in the paper, we add up the light intensity for low lying urban areas (those less than 10 meters above sea level) and compare this to higher elevation urban areas. We find that the share of lights in the low lying areas barely changed from 1992 (when it was 0.1038) to 2012 (when it was 0.1040).

6 Discussion and conclusions

In this paper we study the effect of large floods on local economic activity in cities worldwide. In particular, we examine (i) whether economic activity concentrates disproportionately in flood-prone urban areas; (ii) whether higher risks of extreme precipitation affect the concentration of economic activity in areas with higher risk of flood; and (iii) whether large floods cause economic activity to shift to safer urban areas or safer parts within the same urban area.

Our analysis indicates that urban areas globally face substantial flooding risk. In particular, in our data, low elevation urban areas – those less than 10m above sea level – on average face a one in 20 risk of a large scale flood, displacing at least 100,000 people, hitting their city. This is likely an underestimate of the true risk, given that we have incomplete coverage of the events that meet this criterion during our main sample period (January 2003 to May 2008).[31] We also find that large scale flooding represents a recurrent risk for certain urban locations. Of the 1,868 cities affected by a large flood in our data, about 16 percent get hit in more than one year during our brief sample period. In spite of the greater vulnerability of low elevation areas to flood risk, we find that global urban economic activity as measured by night light intensity is disproportionately concentrated in these areas.[32] This concentration of urban economic activity in flood-prone areas is found to hold even in regions that are prone to extreme precipitation.

Urban flood risk is likely to increase with trends such as population growth and urbanization, which are more intensive in areas currently most at risk – e.g. South Asia and sub-Saharan Africa – along with the potentially exacerbating effects of climate change and rising sea levels.

When we analyze the local economic impact of large floods, we find that on average they reduce the intensity of night lights by between 2 and 8 percent in the year of the flood, depending on the method used to identify affected locations. Our results also show that recovery is relatively quick, with lights fully recovered in the year after the flood. These results – which are consistent across our various specifications and robust to excluding areas within 10km of the nearest river or coast, and to excluding cities that are entirely less than 10m above sea level – indicate a lack of adaptation, in the sense of a movement of economic activity away from the most vulnerable locations within cities. One exception to this appears to be in newly populated areas, where the observed decline in light intensity is both stronger and more persistent.

Projections of future losses associated with climate change rest heavily on assumptions about the degree of adaptation we can expect in response to changing risk profiles. While the potential for human and economic systems to adapt may be high, our results indicate that even large and devastating floods do not permanently displace population, so it may prove challenging to convince people to leave as sea levels rise. In other words, in the medium run, adaptation may be painfully slow.

Defensive investments, involving the building of more robust infrastructure and flood protection schemes, may mitigate some of the risks associated with extreme precipitation and coastal

[31] Although this may have been an especially destructive period, as suggested by the data in Table 1.

[32] This concentration in vulnerable locations also appears to have intensified over time. Looking at changes in light intensity from 2000-2012, we find that low elevation areas have grown more rapidly relative to average country trends (but less rapidly relative to city trends). Looking at city averages, low elevation cities have also grown faster relative to country trends. Results of this analysis available on request. See also Ceola, Laio and Montanari (2014).

flooding. However, these are not costless.[33] Moreover, it is often the case that money and effort are more readily expended in disaster recovery than prevention.[34] Motivating the latter faces political challenges. Aside from the issue of political myopia, it has also been shown that voters are more likely to reward highly visible recovery efforts than preventive actions (Healy and Malhotra, 2009).

Our findings highlight the costs associated with the path dependence of urban locations, and stress the existence of barriers to change in the spatial distribution of economic activity across cities. From a policy perspective, this suggests that incorporating flood risk (and adaptation) into development and urban planning is an important challenge. Making progress on this front is most urgent in developing countries where rapid population growth and urbanization, combined with weak planning and zoning laws, contribute to the high levels of flood risk.

[33]Nor are fiscal costs of natural disasters low. When non-disaster government transfers are added to disaster-specific aid fiscal costs of exogenous shocks in US counties increase almost three-fold (Deryugina 2013).

[34]It has been estimated that $7 of international aid flows are spent on disaster recovery for every $1 spent on prevention (Kellett and Caravani, 2014). Following the 2014 floods in the UK, Prime Minister David Cameron stated that "money is no object in this relief effort" ("Flood simple: the UK flooding crisis explained", Guardian, 13 February 2014, `http://www.theguardian.com/uk-news/2014/feb/13/uk-floods-essential-guide` accessed on 29 April 2015.

References

[1] Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (forthcoming), "The economics of density: Evidence from the Berlin Wall," Econometrica.

[2] Anttila-Hughes, J.K., and S.M. Hsiang (2013), "Destruction, disinvestment, and death: Economic and human losses following environmental disaster," available at SSRN 2220501.

[3] Arellano, M. and S. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," The Review of Economic Studies 58.2: 277-297.

[4] Barrios, S., L. Bertinelli and E. Strobl (2006), "Climatic change and rural-urban migration: The case of sub-Saharan Africa," Journal of Urban Economics 60.3: 357-371.

[5] Barro RJ. Environmental Protection, Rare Disasters and Discount Rates. Economica. 2015 Jan 1;82(325):1-23.

[6] Basker, E. and J. Miranda (2014), "Taken by storm: Business financing, survival, and contagion in the aftermath of Hurricane Katrina," University of Missouri Working Paper 14-06.

[7] Beeson, P.E. and W. Troesken (2006), "When bioterrorism was no big deal," National Bureau of Economic Research Working Paper 12636.

[8] Bleakley, H. and J. Lin (2012), "Portage and path dependence," The Quarterly Journal of Economics, 127(2), 587.

[9] Bosker, M., H. Garretsen, G. Marlet and C. van Woerkens (2014) "Nether Lands: Evidence on the price and perception of rare environmental disasters," Centre for Economic Policy Research DP10307, December.

[10] Boustan, L.P., M.E. Kahn and P.W. Rhode (2012), "Coping with economic and environmental shocks: Institutions and outcomes – Moving to higher ground: Migration response to natural disasters in the early Twentieth Century," American Economic Review: Papers & Proceedings 102.3: 238-244.

[11] Brakenridge, R. and E. Anderson (2006), "MODIS-based Flood Detection, Mapping and Measurement: The Potential for Operational Hydrological Applications" in J. Marsalek et al. (eds), Transboundary Floods: Reducing Risks Through Flood Management, Nato Science Series: IV: Earth and Environmental Sciences 72: 1-12.

[12] Brakenridge, R., "Global Active Archive of Large Flood Events", Darthmouth Flood Observatory, University of Colorado, available from `http://floodobservatory.colorado.edu/Archives/index.html` Accessed in 2015.

[13] Cameron, L. and M. Shah (2015),"Risk-taking behavior in the wake of natural disasters," The Journal of Human Resources 50.2: 484-515.

[14] Caselli, F. and Michaels, G. (2013). "Do Oil Windfalls Improve Living Standards? Evidence from Brazil," American Economic Journal: Applied Economics 5(1): 208-238

[15] Cavallo, E., S. Galiani, I. Noy and J. Pantano (2013), "Catastrophic natural disasters and economic growth," Review of Economics and Statistics 95.5: 1549-1561.

[16] Ceola, S., F. Laio and A. Montanari (2014), "Satellite nighttime lights reveal increasing human exposure to floods worldwide," Geophysical Research Letters 41, 7184-7190, doi: 10.1002/2014GL061859.

[17] Chen, Xi, and William D. Nordhaus. 2011. "Using Luminosity Data as a Proxy for Economic Statistics." PNAS 108(21): 8589-94.?

[18] Davis, D.R. and D.E. Weinstein (2002), "Bones, Bombs, And Break Points: The Geography of Economic Activity," American Economic Review 92.5: 1269.

[19] De Mel, S., D. McKenzie and C. Woodruff (2012), "Enterprise recovery following natural disasters," The Economic Journal 122.559: 64-91.

[20] Desmet, K. and E. Rossi-Hansberg (2012), "On the spatial economic impact of global warming," National Bureau of Economic Research W18546.

[21] Desmet, K., D.K. Nagy and E. Rossi-Hansberg (2015), "The geography of development: Evaluating migration restrictions and coastal flooding", National Bureau of Economic Research Working Paper 21087, April.

[22] Deryugina, T., L. Kawano and S. Levitt (2014), "The economic impact of hurricane Katrina on its victims: evidence from individual tax returns," National Bureau of Economic Research Working Paper 20713.

[23] Deryugina, T. and B. Kirwan (2014), "Does the Samaritan's Dilemma matter? Evidence from crop insurance," July 14, Mimeo.

[24] Deryugina, T. (2013), "The role of transfer payments in mitigating shocks: Evidence from the impact of hurricanes," Available at SSRN 2314663.

[25] Donaldson, D. and A. Storeygard (2016), "The View from Above: Applications of Satellite Data in Economics", Journal of Economic Perspectives 30 (4): 171-198.

[26] Eckel, C.C., M.A. El-Gamal and R.K. Wilson (2009), "Risk loving after the storm: A Bayesian-Network study of Hurricane Katrina evacuees," Journal of Economic Behavior & Organization 69: 110-124.

[27] Elliott, R., E. Strobl and P. Sun (2015), "The local Impact of typhoons on economic activity in China: A view from outer space", Journal of Urban Economics 88: 50-66.

[28] Elvidge, C.D, D. Ziskin, K. E. Baugh, B.T. Tuttle, T. Ghosh, D.W. Pack, E.H. Erwin and M. Zhizhin (2009), "A fifteen year record of global natural gas flaring derived from satellite data," Energies 2: 592-622.

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

[30] Ghosh, T., S. J. Anderson, C. D. Elvidge and P. C. Sutton (2013), "Using nighttime satellite imagery as a proxy measure of human well-being", sustainability, 5: 4688-4719.

[31] Glaeser, E.L. (2005), "Should the government rebuild New Orleans, or just give residents checks?" The Economists' Voice 2.4.

[32] Global Rural-Urban Mapping Project (GRUMPv1), Urban Extents Grid v1 (1995). Center for International Earth Science Information Network (CIESIN), Earth Institute, Columbia University. Available at `http://sedac.ciesin.columbia.edu/data/collection/grump-v1/sets/browse`. Last accessed on 23 October 2015.

[33] Guha-Sapir, D., R. Below and Ph. Hoyois "EM-DAT: The CRED/OFDA International Disaster Database," available from `www.emdat.be`. Universite Catholique de Louvain, Brussels, Belgium. Last accessed on 23 October 2015.

[34] Hallegatte, S., C. Green, R.J. Nicholls and J. Corfee-Morlot (2013), "Future flood losses in major coastal cities," Nature Climate Change 3.9: 802-806.

[35] Handmer, J., Y. Honda, Z.W. Kundzewicz, N. Arnell, G. Benito, J. Hatfield, I.F. Mohamed, P. Peduzzi, S. Wu, B. Shertstyukov, K. Takahashi, and Z. Yan (2012), "Changes in impacts of climate extremes: Human systems and ecosystems" in Managing the risks of extreme events and disasters to advance climate change adaptation [Field, C.B., V. Barrios, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 231-290.

[36] Hanson, S., R. Nicholls, N. Ranger, S. Hallegatte, J. Corfee-Morlot, C. Herweijer and J Chateu (2011), "A global ranking of port cities with high exposure to climate extremes," Climatic Change 104: 89-111.

[37] Harari, M. (2015), "Cities in Bad Shape: Urban Geometry in India", MIT working paper.

[38] Healy, A. and N. Malhotra (2009), "Myopic voters and natural disaster policy," American Political Science Review 103: 387-406.

[39] Henderson, J.V., A. Storeygard and U. Deichmann (2015), "Has climate change driven urbanization in Africa?" Working paper, London School of Economics.

[40] Henderson, J.V., A. Storeygard and D.N. Weil (2012), "Measuring economic growth from outer space." American Economic Review 102.2: 994-1028.

[41] Hodler, R. and P. A. Raschky (2014), "Regional Favoritism", The Quarterly Journal of Economics : 995-1033.

[42] Hornbeck, R. and S. Naidu (2014), "When the Levee Breaks: Black Migration and Economic Development in the American South," American Economic Review, 104(3): 963-90.

[43] Hornbeck, R. and D. Keniston (2014), "Creative destruction: Barriers to urban growth and the Great Boston Fire of 1872," NBER Working Paper No. 20467.

[44] Hsiang, S.M. and A.S. Jina (2014), "The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones," National Bureau of Economic Research Working Paper 20352.

[45] Jedwab, R. L. Christiaensen and M. Gindelsky (2015), "Demography, urbanization and development: Rural push, urban pull and... urban push?" Policy Research Working Paper WPS7333.

[46] Jones, P.D. and I. Harris. CRU TS3.21: University of East Anglia Climatic Research Unit (CRU) Time-Series (TS) Version 3.21 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901- Dec. 2012). NCAS British Atmospheric Data Centre, 24 September 2013. doi:10.5285/D0E1585D-3417-485F-87AE-4FCECF10A992, available from `http://dx.doi.org/10.5285/D0E1585D-3417-485F-87AE-4FCECF10A992`

[47] Kahn, M.E. and R. Walsh (2014), "Cities and the Environment," NBER Working Paper Series. Working Paper 20503, available from `http://www.nber.org/papers/w20503`.

[48] Kellett, J. and A. Caravani (2014), "Financing Disaster Risk Reduction: A 20 year story of international aid," Overseas Development Institute, London.

[49] Kydland, F.E., and E.C. Prescott (1977), "Rules rather than discretion: The inconsistency of optimal plans," The Journal of Political Economy, pp. 473-491.

[50] Marx, B., T. Stoker and T. Suri (2013), "The economics of slums in the developing world," The Journal of Economic Perspectives, pp. 187-210.

[51] Mellander, C., J. Lobo, K. Stolarick, Z. Matheson (2015), "Night-time light data: A good proxy measure for economic activity?", PLoS ONE 10 (10): e0139779

[52] Michaels, G. and F. Rauch (2018), "Resetting the urban network: 117-2012." Economic Journal, 128(608): 378-412.

[53] Michalopoulos, E. and E. Papaioannou (2014), "National institutions and subnational development in Africa," The Quarterly Journal of Economics 129(1): 151-213.

[54] Mueller, V., C. Gray and K. Kosec (2014), "Heat stress increases long-term human migration in rural Pakistan," Nature Climate Change 4: 182-185.

[55] Myhre, G., et al. "Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change." K., Tignor, M., Allen, SK, Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, PM, Cambridge University Press Cambridge, United Kingdom and New York, NY, USA (2013).

[56] Nickell, S. (1981). "Biases in Dynamic Models with Fixed Effects". Econometrica, 49(6): 1417-1426.

[57] Paskoff, P. F. (2008), "Measures of war: A quantitative examination of the civil war's destructiveness in the confederacy," Civil War History 54(1): 35-62.

[58] Patel, S. S. 2009., "Swept Away," Archaeology 62(1).

[59] Redding, S.J., D.M. Sturm and N. Wolf (2011), "History and industry location: Evidence from German airports," Review of Economics and Statistics 93.3: 814-831.

[60] Sear, D.A., Bacon,S.R., Murdock, A.P., Donaghue, G., Baggaley, P., Sera, C., and Lebas, T.P., (2011), "Cartographic, geophysical and diver surveys of the medieval town site at Dunwich, Suffolk, England," International Journal of Nautical Archaeology 40(1) 113-132.

[61] Storeygard, A. (2016), "Farther on down the Road: Transport costs, trade and urban growth in Sub-Saharan Africa", Review of Economic Studies 83; 1263-1295.

[62] Strauss, B. H., S. Kulp and A. Levermann (2015), "Carbon choices determine US cities committed to futures below sea level," PNAS, www.pnas.org/cgi/doi/10.1073/pnas.1511186112.

[63] Strobl, E. (2011), "The economic growth impact of hurricanes: Evidence from U.S. coastal counties," The Review of Economics and Statistics, 93(2): 575-589.

[64] Tessler, Z.D., C.J. Vorosmarty, M.Grossberg, I. Gladkova, H. Aizenman, J.P.M. Syvitski and E. Foufoula-Georgiou (2015), "Profiling risk and sustainability in coastal deltas of the world," Science 349(6248): 638-643.

[65] Turner, G. (2012), "Natural disasters, risk expectations, and population growth. A Study of US counties since 1850," working paper, available from http://www.webmeets.com/files/papers/EAERE/2012/941/NatRiskLocation_12.01.31.pdf.

[66] United Nations (2008), "World Urbanization Prospects - The 2007 Revision," available from www.un.org/esa/population/publications/wup2007/2007wup.htm.

[67] University of East Anglia Climatic Research Unit (CRU) [Phil Jones, Ian Harris] (2008), "CRU Time Series (TS) high resolution gridded datasets," NCAS British Atmospheric Data Centre. Accessed 20 May 2014, available from http://badc.nerc.ac.uk/.

[68] Weitzman ML. "On modeling and interpreting the economics of catastrophic climate change". The Review of Economics and Statistics. 2009 Feb 1;91(1):1-9.

Figure 1: Map of all urban areas in the world showing the locations of cities affected by our sample of large flood events. City sizes are inflated in order to make them visible on a map of the entire world. Smaller dots correspond to cities not affected by any of the floods in our sample. The number of floods in the legend refers to the number of years from 2003-2008 during which each city was affected by a flood that displaced a total of 100,000 people or more. This map uses the Mercator (WGS 1984) projection. The rest of the analysis in the paper uses equal area projections.

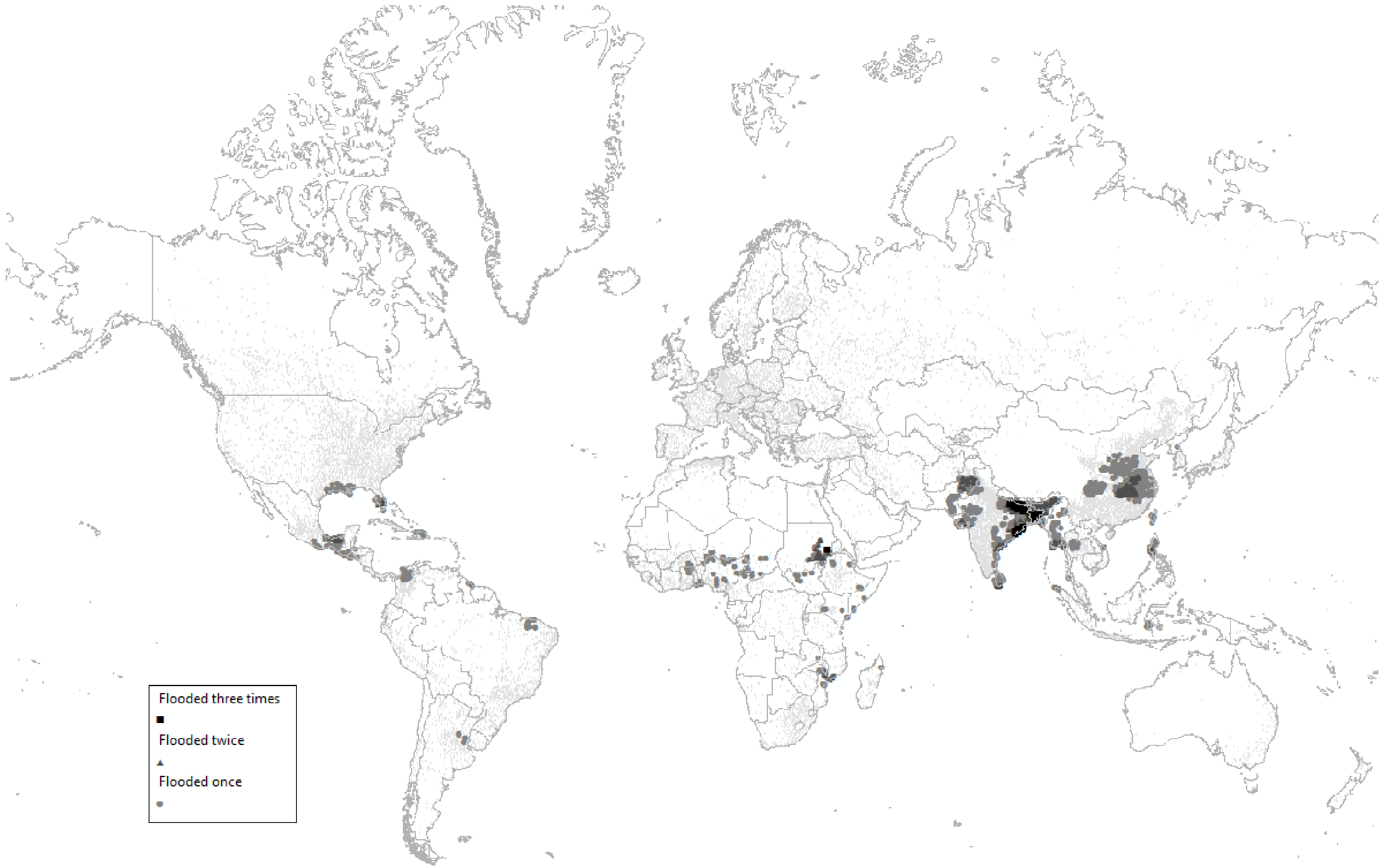

Figure 2: Inundation and light intensity maps for Hurricane Katrina, New Orleans. Panel A shows a detail from one of the inundation maps associated with Hurricane Katrina, concentrated on the area around the city of New Orleans. The map displays in red and pink the areas that were inundated during the flooding. Panels B, C and D show the average annual light intensity in 2004, 2005, 2006 respectively, for the city of New Orleans. There is a notable dimming of lights city-wide in 2005. This is particularly pronounced in the eastern parts of the city, which were worst affected by the flood. In Panel D a recovery of light intensity is apparent. We are unable to observe any decline in light intensity in the range above the top coded light intensity level of 63. However, New Orleans is rich compared with the rest of our sample, where top coding tends to be less frequent.

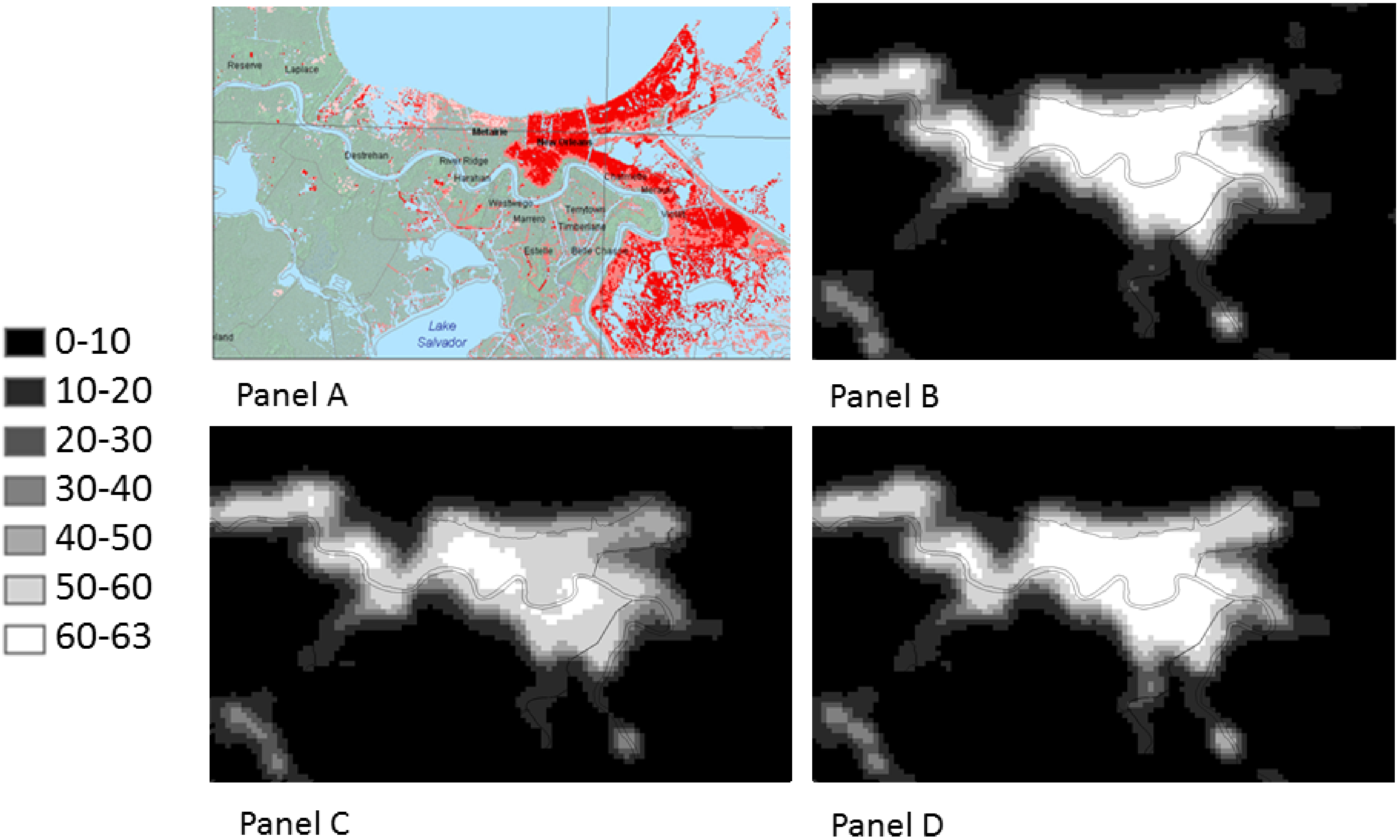

Table 1: Flood events displacing at least 100,000 people, by year (1988-2014)

Year	Number of events	(in our sample)	Millions of people displaced	(in our sample)
1988	22		19.1	
1989	20		8.0	
1990	18		14.2	
1991	21		16.9	
1992	12		12.6	
1993	16		34.2	
1994	15		7.8	
1995	24		47.4	
1996	18		12.1	
1997	21		5.6	
1998	23		41.7	
1999	22		56.4	
2000	20		49.2	
2001	13		9.4	
2002	16		19.0	
2003	16	(13)	20.6	(19.9)
2004	19	(15)	50.0	(49.1)
2005	30	(8)	21.8	(5.8)
2006	25	(7)	16.7	(5.2)
2007	30	(9)	33.2	(8.2)
2008	24	(1)	20.7	(1.5)
2009	17		7.8	
2010	17		19.8	
2011	14		6.9	
2012	12		5.1	
2013	14		6.2	
2014	9		3.1	
Total	508	(53)	565.6	(89.6)

Notes: Data from the Dartmouth Flood Observatory (DFO) database. Our sample refers to the 53 flood events from the DFO database for which we have detailed inundation maps, as discussed in the text.

Table 2: Flood odds by location characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$	$FloodFreq_{ik}$
$Elev < 10m_i$	0.036			0.041	0.033			0.034
	(0.017)			(0.018)	(0.018)			(0.019)
$River_i$		0.010		0.009		0.009		0.009
		(0.006)		(0.006)		(0.006)		(0.006)
$Coast_i$			-0.001	-0.010			0.007	-0.001
			(0.007)	(0.003)			(0.006)	(0.001)
$Constant$	0.013	0.014	0.016	0.012				
	(0.005)	(0.004)	(0.005)	(0.004)				
Observations	3,807,799	3,807,799	3,807,799	3,807,799	3,807,799	3,807,799	3,807,799	3,807,799
Country fixed effects	No	No	No	No	Yes	Yes	Yes	Yes

Notes: The regressions reported in this Table correspond to Equation 1, and include the full global sample of all urban areas.
The dependent variable $FloodFreq_{ik}$ measures the odds of flooding per year for a given location, defined as the number of years during our main sample in which each location is hit by at least one large flood event, divided by the length of the sample (five years and five months).
$Elev < 10m_i$ is a dummy variable for locations that are less than 10m above sea level.
$River_i$ and $Coast_i$ indicate locations that are less than 10km from the nearest river or coast, respectively.
Robust standard errors, clustered by country, in parentheses.

Table 3: Light intensity by elevation and exposure to extreme precipitation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$	$ln(Y_{ilk})$
$Elev < 10m_i$	0.182	0.184	0.137	0.059	0.157	0.110	0.038	0.172	0.125	0.047
	(0.037)	(0.037)	(0.032)	(0.027)	(0.033)	(0.028)	(0.033)	(0.034)	(0.029)	(0.028)
$Elev < 10m_i \times Precip > 1000mm_l$		-0.034	-0.032	-0.030						
		(0.075)	(0.075)	(0.048)						
$Precip > 1000mm_l$		-0.063	-0.081	-0.021						
		(0.079)	(0.074)	(0.054)						
$Elev < 10m_i \times Precip > 500mm_l$					0.113	0.112	0.078			
					(0.044)	(0.048)	(0.064)			
$Precip > 500mm_l$					-0.043	-0.054	-0.078			
					(0.064)	(0.062)	(0.054)			
$Elev < 10m_i$								0.067	0.065	0.049
$\times Precip > 500mm(twice)_l$								(0.026)	(0.028)	(0.038)
$Precip > 500mm(twice)_l$								-0.050	-0.066	-0.114
								(0.075)	(0.073)	(0.056)
Observations	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083	3,642,083
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	No	No	No	Yes	No	No	Yes	No	No	Yes
River & Coast FE	No	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
$Prec > 1000mm$	Mean	0.012								
$Prec > 500mm$	Mean	0.153								
$Prec > 500mm(twice)$	Mean	0.111								

Notes: The regressions reported in this Table correspond to Equation 2 and include the full global sample of all urban areas.
The dependent variable in all regressions $ln(Y_{ilk})$ is the natural log of mean light intensity (measured in 2012) at each gridpoint i (located in grid cell l, in country k).
$Elev < 10m_i$ is a dummy variable for locations that are less than 10m above sea level.
$Precip > 1000mm_l$ ($> 500mm_l$) indicates locations that have experienced monthly precipitation of 1000mm (500mm) or more at least once, and in the case of $Precip > 500mm(twice)_l$ monthly precipitation of 500mm or more at least twice, in the period 1992-2012.
Regressions with river and coast controls include dummies for locations within 10km of the nearest river or coast.
Robust standard errors, clustered by country, in parentheses.

Table 4: Main effects of flood on light, gridpoint year panel

	(1)	(2)	(3)	(4)	(5)	(6)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Flood_{jt}$	-0.021			-0.023		
	(0.010)			(0.010)		
$Precip > 500mm_{lt}$		-0.025			-0.027	
		(0.008)			(0.008)	
$Precip > 1000mm_{lt}$			-0.080			-0.083
			(0.018)			(0.018)
$ln(light_{t-1})$				Yes	Yes	Yes
Observations	1,422,018	1,422,018	1,422,018	1,392,501	1,392,501	1,392,501
No. of gridpoints	243,303	243,303	243,303	235,460	235,460	235,460

Notes: The results presented in this Table correspond to Equation 3 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt}$ is a dummy indicating whether or not city j was hit by a large flood in year t.
$Precip > 1000mm_{lt}$ ($> 500mm_{lt}$) indicates locations that experienced monthly precipitation of 1000mm (500mm) or more in year t.
All regressions include year fixed effects, gridpoint fixed effects and country-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table 5: Recovery, gridpoint year panel

	(1) $ln(Y_{ijkt})$	(2) $ln(Y_{ijkt})$	(3) $ln(Y_{ijkt})$	(4) $ln(Y_{ijkt})$	(5) $ln(Y_{ijkt})$	(6) $ln(Y_{ijkt})$	(7) $ln(Y_{ijkt})$	(8) $ln(Y_{ijkt})$	(9) $ln(Y_{ijkt})$	(10) $ln(Y_{ijkt})$
$Flood_{jt}$	-0.021					-0.023				
	(0.009)					(0.011)				
$Flood_{jt-1}$		-0.003					-0.021			
		(0.012)					(0.014)			
$Flood_{jt-2}$			0.017					0.017		
			(0.015)					(0.015)		
$Flood_{jt-3}$				0.005					-0.002	
				(0.004)					(0.007)	
$Flood_{jt-4}$					0.001					0.001
					(0.004)					(0.005)
$ln(light_{t-1})$						Yes	Yes	Yes	Yes	Yes
Observations	1,422,018	1,414,781	1,417,877	1,421,167	1,420,548	1,392,501	1,386,261	1,380,492	1,375,245	1,374,842
No. of gridpoints	243,303	243,292	244,256	245,077	245,018	235,460	235,421	235,302	234,838	234,962

Notes: The results presented in this Table correspond to Equation 3 and use the sample of cities affected by at least one of the large flood events in our data.

The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.

$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.

All regressions include year fixed effects, gridpoint fixed effects and country-specific trends.

Columns (6) to (10) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.

Robust standard errors, clustered by country, in parentheses.

Table 6: Interactions with elevation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Flood_{jt} \times elev_{<10i}$	-0.027			-0.028			-0.030		
	(0.006)			(0.006)			(0.005)		
$Flood_{jt} \times elev_{10+i}$	-0.019			-0.021			-0.017		
	(0.012)			(0.013)			(0.007)		
$Flood_{jt-1} \times elev_{<10i}$		0.009			-0.014			0.003	
		(0.009)			(0.009)			(0.011)	
$Flood_{jt-1} \times elev_{10+i}$		-0.007			-0.023			-0.006	
		(0.014)			(0.018)			(0.012)	
$Flood_{jt-2} \times elev_{<10i}$			0.042			0.043			0.036
			(0.016)			(0.017)			(0.020)
$Flood_{jt-2} \times elev_{10+i}$			0.008			0.007			0.011
			(0.011)			(0.011)			(0.010)
$ln(light_{t-1})$				Yes	Yes	Yes			
Country-specific trends	Yes	Yes	Yes	Yes	Yes	Yes			
City-specific trends							Yes	Yes	Yes
Observations	1,422,018	1,414,781	1,417,877	1,392,501	1,386,261	1,380,492	1,422,018	1,414,781	1,417,877
No. of gridpoints	243,303	243,292	244,256	235,460	235,421	235,302	243,303	243,292	244,256

Notes: The results presented in this Table correspond to Equation 5 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
$Elevation_h$ is a dummy for elevation band h, where h is either less than 10m above sea level, or 10m or more above sea level.
All regressions include year fixed effects and gridpoint fixed effects. Columns (1) to (6) include country-specific trends and Columns (7) to (9) include city-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table 7: Flood impacts by newly populated vs existing locations

	(1) $ln(Y_{ijkt})$	(2) $ln(Y_{ijkt})$	(3) $ln(Y_{ijkt})$	(4) $ln(Y_{ijkt})$	(5) $ln(Y_{ijkt})$	(6) $ln(Y_{ijkt})$	(7) $ln(Y_{ijkt})$	(8) $ln(Y_{ijkt})$	(9) $ln(Y_{ijkt})$	(10) $ln(Y_{ijkt})$	(11) $ln(Y_{ijkt})$	(12) $ln(Y_{ijkt})$
$Flood_{jt} \times New_i$	-0.023						-0.026					
	(0.007)						(0.012)					
$Flood_{jt} \times Old_i$	-0.021						-0.022					
	(0.010)						(0.011)					
$Flood_{jt-1} \times New_i$		-0.065						-0.078				
		(0.025)						(0.019)				
$Flood_{jt-1} \times Old_i$		0.003						-0.016				
		(0.011)						(0.014)				
$Flood_{jt-2} \times New_i$			-0.073						-0.067			
			(0.032)						(0.021)			
$Flood_{jt-2} \times Old_i$			0.025						0.023			
			(0.014)						(0.015)			
$Flood_{jt-3} \times New_i$				-0.094						-0.074		
				(0.043)						(0.032)		
$Flood_{jt-3} \times Old_i$				0.014						0.003		
				(0.005)						(0.009)		
$Flood_{jt-4} \times New_i$					-0.072						-0.066	
					(0.040)						(0.034)	
$Flood_{jt-4} \times Old_i$					0.008						0.005	
					(0.007)						(0.006)	
$Flood_{jt-5} \times New_i$						-0.039						-0.033
						(0.025)						(0.020)
$Flood_{jt-5} \times Old_i$						0.004						0.004
						(0.007)						(0.007)
$ln(light_{t-1})$							Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,422,018	1,414,781	1,417,877	1,421,167	1,420,548	1,185,258	1,392,501	1,386,261	1,380,492	1,375,245	1,374,842	1,141,486
No. of gridpoints	243,303	243,292	244,256	245,077	245,018	244,711	235,460	235,421	235,302	234,838	234,962	233,110

Notes: The results presented in this Table are variations on Equation 5 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
New_i is a dummy for locations that were unlit ($lights = 0$) in 1992. Old_i is a dummy for locations that were lit ($lights > 0$) in 1992.
All regressions include year fixed effects, gridpoint fixed effects and country-specific trends.
Columns (7) to (12) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

7 Online Appendix A: Theory

To frame our empirical investigation, we outline a simple framework that allows us to consider how individuals may respond to a large flood. We consider a discrete-time model, where a person has to choose between two locations, one to which we refer as "Risky" (indexed by R) and another which we will for simplicity consider "Safe" (indexed by S).

The person in question resides initially in the risky location, and considers whether to relocate to the safe location. The period utility of the person from the risky location is

$$U_R = V_R - P_F(D_F - T_F) + u(c), \tag{6}$$

where V_R is the utility from residing in the risky location; P_F is the assessed probability of a flood, which we discuss below; D_F and T_F are the damage from a flood and the transfers received in the aftermath of a flood, expressed in utility terms; and $u(c)$ is the utility from non-residential consumption. We drop time subscripts to increase simplicity. The period utility from the safe location is

$$U_S = V_S + u(c), \tag{7}$$

where V_S is the utility from residing in the safe location. But in order to move the person has to pay relocation costs M, which capture the cost of moving. We also assume that once a flood has hit the person has to pay the cost M regardless of whether they move or stay, since the flood implies paying costs of renovating over and above those captured by D_F. The point of this simplifying assumption is that when a flood hits, the cost of moving (compared to staying) is lower than in the absence of the flood. We assume that non-residential consumption is a numeraire good, whose price is normalized to one. The budget constraint is therefore:

$$I = p(L) + c + M \times 1_{move},$$

where I is income (which we assume to be constant);[35] $p(L)$ is the rental price of residing in location $L \in \{R, S\}$, which is paid to absentee landlords; and 1_{move} is an indicator for moving. The choice over relocation represents an infinite horizon problem, with discount rate θ. Given the simple structure of the model, and holding prices fixed, our individual relocates from the risky to the safe location if $V_S - V_R + P_F(D_F - T_F)$ is sufficiently large. An important factor in this model is how the person assesses the probability of a flood. Following Turner (2012) we model flooding through a Beta-Bournoulli Bayesian learning model.[36] We assume that the risk of a flood (by which we mean a large flood) in a given year is x. Our resident's prior is that x is

[35]We could assume that income depends on location, but this would not substantively change the model.

[36]Gallagher (2014) provides evidence of Bayesian learning in the context of floods. As we explain below this is a simplification, since this probability can rise with climate change, or decline with public investment to address climate change.

distributed according to a Beta distribution: $x \sim \beta(\alpha, \beta)$. The probability distribution function is:

$$f(x|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, x \in [0, 1], \alpha > 0, \beta > 0, \tag{8}$$

where the normalization constant is the Beta function:

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx. \tag{9}$$

The prior probability of a flood is therefore

$$P_F = E[x] = \frac{\alpha}{\alpha + \beta}. \tag{10}$$

After observing t years, during which a flood has occurred S_t times, the updated posterior is:

$$E[x|t, S_t] = \frac{\alpha + S_t}{\alpha + \beta + t}. \tag{11}$$

In other words, for an individual who has information on flood events in the past t years, the expected probability of a flood next year increases by $1/(t+\alpha+\beta)$ if a flood took place in year t compared to the case where it did not. As t approaches infinity there is no updating. The model captures the intuition of Bayesian learning: as t approaches infinity there is no more updating, since the degree of risk is known.

If as a result of the flood people update the risk of flooding, some of them may want to relocate from the risky area to the safe one. Similarly, if there is no flood and people update, there may be movement in towards the risky areas. But without loss of generality we will focus on the case where a flood does occur.

If people do update and some want to move from the risky area to the safe one, then the population or the prices in the safe area (or both) will increase relative to the risky area. If housing supply is inelastic, then the price of housing in the safe areas will increase relative to the risky areas, but there will be no change in the population ratio between the two locations and we might not detect any change in night time light activity. But if housing supply is not completely fixed, then we expect both the price of housing and the population of the safe area to increase relative to the risky area so that updating will result in changes in economic activity as reflected by night time lights.

This simple model guides our empirical investigation in the following ways. First, we investigate the link between risk and low elevation locations. Anticipating and quantifying flood risk in the real world is a complicated endeavor, but we ask specifically how much more susceptible to large scale flooding are low elevation locations, compared to high elevation ones. This informs us about the approximate magnitude of P_F.

Second, we ask whether people generally reside in riskier low elevation urban areas. In the model, the benefits to living in risky areas (if $V_R > V_S$), or moving costs, M, might make it prohibitively expensive to relocate. One set of advantages for risky areas could be that living near coasts or rivers makes seaborne activities, such as trade and fishing, less costly. At the same time, living in flood prone areas may be the legacy of historical lock-in (Bleakly and Lin 2012; Michaels and Rauch 2013).

Third, floods may cause people to leave the riskier areas because of either Bayesian updating, or because floods reduce the cost of moving to safer areas (relative to staying in the riskier ones). Our paper examines the extent to which large floods move economic activity away from risky areas towards safer ones.

Fourth, because updating decreases in t, we expect that there will be more updating in newly populated urban areas. In the empirical analysis we examine whether there is more relocation from riskier to safer areas in the aftermath of a flood in urban areas that concentrated no (measurable) economic activity until recently.

Fifth, we examine whether the presence of higher risk of flooding due to climatic factors shifts people towards safer areas. In our model, an increase in P_F holding all else constant, shifts people away from risky low elevation areas.

In addition to the issues raised by our model, an important question is whether rising sea levels and a changing climate will affect the aggregate global concentration of economic activity in flood prone areas. Of course, rising sea levels may increase the risk of flooding both in low elevation areas and in other areas that are currently safer. But it seems plausible to assume that at least in the near future, it is in the low elevation areas that rising sea levels will have a greater effect. In our analysis we will shed some light on the aggregate concentration of urban economic activity in low elevation areas over a longer period of time.

Our analysis also touches upon a number of normative considerations. As Kydland and Prescott (1977) note, flood protection may exacerbate the moral hazard problem of living on the flood plains. By spending public money to reduce the risk borne by those living in flood prone areas, such flood protection involves a cost. At the same time, as our paper shows, people may be reluctant to relocate away from risky areas. As sea levels rise and the world becomes richer, the tradeoffs between flood protection and the relocation of economic activity to safer areas are likely to become an important issue for public debate (see Strauss, Kulp and Levermann, 2015).

Another normative issue is how much ex-post transfers should victims receive, and in what form. In the model, a larger value of T_F makes movement away from risky areas less likely. From the perspective of a donor, if a property is frequently flooded, the costs of repeatedly paying compensation might be high. In developing countries where institutions are weak, finding private flood insurance may be a difficult challenge, especially for the poor. Ex-post disaster relief, including from large scale floods, is therefore a task that governments and non-government organizations around the world engage in from time to time. The main policy issue that we raise is whether it should be possible, in certain circumstances, to concentrate public reconstruction efforts towards safer areas, in order to avoid the high risk of recurrent disasters.

8 Online Appendix B: What do variations in night lights capture?

Our aim in this paper is to examine how prevalent it is for economic activity to concentrate in flood-prone areas, and whether cities adapt to major floods by relocating economic activity to safer areas. In particular, we ask does economic activity within cities readjust in response to major shocks, which are potentially recurrent, and which disproportionately threaten specific neighborhoods? This research question requires data on economic activity (or a reasonable proxy thereof), with temporal variation, high spatial resolution, and global coverage. In our analysis we observe variations in night light intensity on a 1km grid, in more than 1800 cities in 40 countries around the world in response to large-scale flooding (or extreme precipitation) events. We undertake the usual data cleaning exercises as described in the data section.

As Donaldson and Storeygard (2016) point out, the correlation between lights and economic activity in the cross section has long been noted (see for example Croft 1973; Doll, Muller, and Morley 2006), while Henderson et al. (2012) were perhaps the first to formally test the relationship between changes in lights and economic growth, using GDP data at the national level. Since then numerous studies have used the lights data as a proxy for economic activity or prosperity at a local (sub-national) level. An example closely related to our paper, in terms of the challenge of finding income data at the city level or finer spatial resolution is Storeygard (2016) who uses lights to proxy for economic activity for a sample of cities in Africa. Lights data offer the opportunity to study variation in economic activity where traditional economic data (especially on income and especially in a panel) are generally not available - as is the case at the city level in Africa.

Ghosh et al. (2013) provide a relevant summary of the various ways in which night time light data have been used to measure human wellbeing at the subnational level, while Donaldson and Storeygard (2016) highlight a number of novel uses of night lights data as a proxy for economic activity within small geographic units. They note that "[lights data] can plausibly be used as a proxy for economic activity under the assumption that lighting is a normal good" (p.183).

In spite of the increasing prevalence of studies exploiting the night lights data as a proxy for local economic activity, an important question is how lights respond to changes in economic activity over various spatial and temporal scales.

Using annual data for a panel of countries from 1992 to 2008, Henderson, Storeygard and Weil (2012) find evidence of a linear relationship between lights and GDP, with an estimated lights-GDP elasticity of around 0.3. The estimated relationship is of similar magnitude for a restricted sample of low and middle-income countries. At national level, some light emission growth variables correlate stronger with growth of GDP, non-agricultural GDP and manufacturing value-added (Addison and Stewart 2015).

More relevant to us are studies that examine the relationship between lights and economic activity at subnational scales. An early example is Sutton, Elvidge and Ghosh (2007) who find that "night lights track economic output" at the state or province level, for four countries (China, India, Turkey and the USA).

Taking a global perspective, Chen and Nordhaus (2011) compare night lights to economic output measured on a 1-degree grid (approx. 100km by 100km). They find a strong positive relationship between luminosity and output at the grid cell level globally. They note that the relationship is weak at the low end of the output/luminosity spectrum, specifically for log output density of less than -9, which is equivalent to about $100,000 per sq km. They conclude that estimates of income are only substantially improved for countries or regions within countries with relatively weak or non-existent traditional economic statistics.

Hodler and Raschky (2014) also estimate the relationship between lights and GDP at the regional level using a broader panel dataset of regional GDP (assembled by Gennaioli et al. 2013), which includes data for 1,503 regions in 82 countries. Similarly to Henderson et al. (2012) they find a lights-GDP elasticity of around 0.3, with a slightly larger estimate (0.386) for the short-run relationship than for the long-run relationship (0.227). Their evidence also points to a roughly linear relationship between lights and GDP at the regional level. Hodler and Raschky (2014) conclude "the relationship between night-time light and GDP is linear and thereby similar across regions with different nighttime light intensity and income levels" (p. 1030).

As mentioned above, a more closely related study, in terms of the data challenges that we face, is Storeygard (2016) who uses lights data to proxy for city-level income for a sample of cities in Africa, where income data are unavailable. As a verification exercise, Storeygard tests the relationship between lights and city (or prefecture) level GDP using Chinese data, finding that the elasticity of GDP with respect to light is significant and positive for a long difference specification (from 1990/92 - 2005). The point estimate (of around 0.25) is very similar (using either the city or prefecture data) to that found for the global sample at the country level.

As part of our own robustness tests, we estimated a lights-GDP elasticity of 0.2 for an annual panel of Indian districts over our study period of 2003-2008, which is again quite similar to the findings in Henderson et al. (2012), Hodler and Raschky (2014) and Storeygard (2016). The consistency of this finding is encouraging.

Mellander et al. (2015) examine the strength of the relationship between night time lights and economic activity using fine-grained official socio-economic data on individuals and establishments in Sweden (on a 250m grid for urban areas and 1000m for rural areas). They find that night time light has a relatively weak relationship with economic activity as measured by people's wages (consumption) or wages by establishment (production), but a relatively strong relation with population density, at this fine spatial scale. Their findings clearly indicate the limitations of the lights data for analyses in high income countries, where top-coding in the lights data becomes an important constraint. As most of our sample is in low to middle income countries, the share of top coded cells in our dataset is small (see discussion in the data section).

The relationship between lights and income or wealth has also been tested at the micro level for developing countries, for example using data from the Demographic and Health Survey (DHS), e.g. Weidmann and Schutte 2017 (see also Michalopoulos and Papaioannou 2013). Weidmann and Schutte (2017) use the DHS data to compare lights to a wealth index constructed at the household level (based on assets) for a sample of 34,047 clusters (typically a village or neighborhood within a city) from 56 surveys in 39 countries for the years 2003 and 2012. They find a correlation between lights and the wealth index that averages 0.73 across all 56 surveys included. While some of this correlation is accounted for by the differences across rural and urban locations, they also estimate the relationship for rural and urban locations separately, finding that the average correlation between lights and wealth is slightly higher for urban clusters (0.62) compared to rural clusters (0.42). They also test the relationship between predictive accuracy and the size of the buffer around the cluster location point used to measure the light intensity associated with that point; the minimum is 2km for urban clusters and 5km for rural clusters, reflecting the random artificial error introduced to the location information in the DHS data (for the sake of preserving anonymity of survey respondents). Weidmann and Schutte also experiment with buffer radii of 5km, 10km and 20km. They find that buffer size matters a lot - as they increase the minimum radius there is an increase in prediction error, which they interpret as "a clear indication that the local levels of night lights - and not the emissions across a region - seem to matter for prediction [of wealth]" (p.131).

9 Online Appendix C: The relationship between night lights, GDP, and floods in India

The evidence that we present in the paper focuses on the effect of large floods on night time lights. As we discuss in the literature survey (in Part B of this Online Appendix), a recent literature confirms that there is a strong relationship between night lights and economic activity. This makes night time lights a useful measure of GDP, since it is available at the local-annual level around the globe, even in locations where local GDP is missing or mis-measured. This strongly suggests that our results reflect the effect of floods on economic activity around the world's flooded cities.

In this section we further explore the relationship between night time lights, GDP, and floods, using data from one particular country - India. We focus on India because its cities were affected by large floods covered in our dataset in five of the six years from 2003-2008. No other country in our dataset of global cities was affected for more than three years. In addition, over the period of our study, large floods affected districts in roughly three quarters of India's states. This means that India, in addition to its size and population, exhibited a fair degree of temporal and spatial variation in the occurrence of large floods, making it an interesting case study.

Data

We obtain GDP data on Indian districts from the Indian government's planning commission (`http://planningcommission.nic.in/plans/stateplan/index.php?state=ssphdbody.htm`). We complement these with administrative boundary shapefiles from the GADM database of Global Administrative Areas (`http://www.gadm.org`).

Starting with the GDP data, we note that there were changes in districts in some states over time. To harmonize district definitions we made the following changes, which involved merging districts (listed below by state): Assam: Bongaigaon includes Chirang; Barpet includes Baksa (an imperfect match); Darrang includes Udalguri. Haryana: Gurgaon includes Mewat (an imperfect match). Jharkhand: Dumka includes Jamtara; Gumla includes Simdega; Palamu includes Latehar; Singhbhum includes Saraykela Kharsawa. West Bengal: Midnapore covers Midnapore East and Midnapore West.

Next, we merge the GDP data to the administrative boundary shapefile. This involved the following steps. We start with district-level outcomes on 24 states. One state's name was changed from Orissa to Odisha. Another state, Telangana, was later created out of Andhra Pradesh. We are missing district GDP data for some states, most of which have relatively small population (by India's standards). The states and Union Territories (UTs) for which we miss district data are, in alphabetical order: Chhattisgarh, Dadra and Nagar Haveli, Daman and Diu, Goa, Gujarat, Jammu and Kashmir, Lakshadweep, Nagaland, NCT of Delhi, Puducherry, and Tripura. Nevertheless, the states that we do have information on cover 87.5% of India's population, 71% of Indian GDP and 89.7% of its area in 2000.

After dropping the 11 states and UTs for which we have no GDP data, the remaining 25 states and UTs contain a total of 581 districts, according to the GADM administrative boundaries. The GDP data for the same 25 states and UTs contain a total of 522 districts.

An initial computer match of the GADM and GDP data, based on matching district and state names (using the reclink command in Stata), returned 403 exact matches, and a further 59 matches (using a minscore=0.95). A further iteration of this process (dropping the minscore to the default 0.6) returned a further 22 matches. These computer matches were checked manually line-by-line to ensure they were accurate and no false positives were kept.

The remaining districts were matched manually - many were identifiable from variations in the district name, while others (some 44) involved districts that had been newly created in the years after the end of our short time series of GDP data (which covers the period 2000-2008). Only two districts in the GADM data remained unmatched: Balod and Surajpur (combined population of approx. 1.5m), both in the state of Chhattisgarh.

Having matched the district-year GDP data to the GIS map, we now proceed to match in the night lights and flooding data. We define a flood indicator to take a value of 1 if any point in a district and year was flooded. We take the mean light intensity value of all the pixels in a district as the light variable. The accounting year for population and GDP does not overlap with calendar years. For this reason, and following the accounting year dates, we compute a weighted average that gives three quarter weight to the current year population, and one quarter to last year's. GDP and population data are provided at district level. We also have a district map that we use to get flood and light data via GIS. We can match the districts on the map with the districts in the census data with a few corrections of abbreviated or slightly differently spelled names.

We combine all these data to construct a panel of district-year observations. We have a balanced panel for 519 districts in the years 2000-2004 inclusive. Because of missing data this number becomes 452 in 2005, 196 in 2006 and 105 in 2007. For 2008 we only have 13 observations in the dataset.

Results

We begin to use our dataset by examining the relationship between night lights and GDP in Indian districts as reported in Appendix Table A9. In Column (1) we regress the logarithm of night lights on the logarithm of GDP at the district level.[37] The regression estimates for the first year of our study, 2003, give a precisely estimated coefficient of 0.44, with a standard error of about 0.04. This coefficient is quite stable when we add state fixed effects (Column (2)), and is broadly similar for other years. When we reverse the order of the left hand side and right hand side variables of interest, regressing the logarithm of GDP on the logarithm of light, we get a coefficient of about 1.11 (s.e. 0.12) without state fixed effects and 0.93 (s.e. 0.1) with state fixed effects. All these regressions suggest a strong correlation between GDP and night light activity at the district level in India.

We next examine the same relationship using panel data, and controlling for district and year fixed effects. The panel of lights and GDP data covers the period 2000-2008 (Columns 5-6). The panel of lights, GDP and floods covers the period 2003-2007 (Columns 7-10). The regression coefficient from this regression is around 0.19, with a standard error of around 0.05. When we now reverse the order of the variables of interest, however, we get an imprecise estimate of 0.17 with a standard error of 0.12 (Columns (3) and (4)). Taken together these results still suggest that local GDP and night lights, which are both measured with error, do tend to co-move in

[37]While in the main part of the paper we cluster the standard errors by country, here we have only one country, so we cluster the standard errors by state.

panel data, but the relationship between them is not as precise. It seems plausible that the elasticity of GDP with respect to night lights in India is around 1, but there is quite a bit of uncertainty about this magnitude.

When we then use the Indian district-level panel to regress the logarithm of night lights on the flood indicator, again controlling for district and year fixed effects and clustering by state, the coefficient estimate is around -0.02 with a standard error of around 0.014. In other words, the point estimate is quite similar to what we find in our main regressions, but it is less precise. We think that this lower precision is reasonable, since we are only using a fraction of the data that we use in the main analysis. When we repeat the regression with lagged floods we get an imprecisely estimated coefficient of -0.02. When we repeat the analysis using the logarithm of GDP instead of the logarithm of lights as the dependent variable, the coefficients are -0.005 and 0.015 for the flood indicator and the lagged flood indicator, respectively.

Taken together, these results suggest that in India night lights and GDP tend to co-move, as the literature finds. But using variation across districts in India alone, we do not have enough power to detect effects of floods on night lights or local GDP. This is perhaps unsurprising, since our regression estimates using all the world's cities fall well within the standard confidence intervals estimated using the Indian data.

References

[1] Addison, D. and B. Stewart (2015) "Nighttime Lights Revisited: The Use of Nighttime Lights Data as a Proxy for Economic Variables", Policy Research Working Paper 7496, World Bank Group, November 2015.

[2] Bleakley, H. and J. Lin (2012), "Portage and path dependence," The Quarterly Journal of Economics, 127(2), 587.

[3] Chen, Xi, and William D. Nordhaus. 2011. "Using Luminosity Data as a Proxy for Economic Statistics." PNAS 108(21): 8589-94.?

[4] Croft, Thomas A. 1973. "Burning Waste Gas in Oil Fields." Nature 245: 375-76

[5] Doll, Christopher N. H., J.-P. Muller, and Jeremy G. Morley. 2006. "Mapping Regional Economic Activity from Night-time Light Satellite Imagery." Ecological Economics 57(1):75-92.

[6] Donaldson, D. and A. Storeygard (2016), "The View from Above: Applications of Satellite Data in Economics", Journal of Economic Perspectives 30 (4): 171-198.

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

[8] Gennaioli, N. R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2013), "Growth of Regions", NBER Working Paper 18937.

[9] Ghosh, T., S. J. Anderson, C. D. Elvidge and P. C. Sutton (2013), "Using nighttime satellite imagery as a proxy measure of human well-being", sustainability, 5: 4688-4719.

[10] Henderson, J.V., A. Storeygard and D.N. Weil (2012), "Measuring economic growth from outer space." American Economic Review 102.2: 994-1028.

[11] Hodler, R. and P. A. Raschky (2014), "Regional Favoritism", The Quarterly Journal of Economics : 995-1033.

[12] Kydland, F.E., and E.C. Prescott (1977), "Rules rather than discretion: The inconsistency of optimal plans," The Journal of Political Economy, pp. 473-491.

[13] Mellander, C., J. Lobo, K. Stolarick, Z. Matheson (2015), "Night-time light data: A good proxy measure for economic activity?", PLoS ONE 10 (10): e0139779

[14] Michaels, G. and F. Rauch (2013), "Resetting the urban network: 117-2012." Bureau for Research and Economic Analysis of Development (BREAD).

[15] Michalopoulos, S. and E. Papaioannou (2013), "Pre-colonial ethnic institutions and contemporary African development", Econometrica 81 (1): 113-152.

[16] Storeygard, A. (2016), "Farther on down the Road: Transport costs, trade and urban growth in Sub-Saharan Africa", Review of Economic Studies 83; 1263-1295.

[17] Strauss, B. H., S. Kulp and A. Levermann (2015), "Carbon choices determine US cities committed to futures below sea level," PNAS, www.pnas.org/cgi/doi/10.1073/pnas.1511186112.

[18] Sutton, Elvidge and Ghosh (2007), "Estimation of gross domestic product at sub-national scales using nighttime satellite imagery," International Journal of Ecological Economics and Statistics 8 (S07): 5-21.

[19] Turner, G. (2012), "Natural disasters, risk expectations, and population growth. A Study of US counties since 1850," working paper, available from `http://www.webmeets.com/files/papers/EAERE/2012/941/NatRiskLocation_12.01.31.pdf`.

[20] Weidmann, N. B. and S. Schutte (2017), "Using night light emissions for the prediction of local wealth", Journal of Peace Research 54 (2): 125-140.

10 Appendix Tables

(not necessarily for publication)

Table A1: Light intensity by elevation, democracy and income levels

	(1) $ln(Y_{ik})$	(2) $ln(Y_{ik})$	(3) $ln(Y_{ik})$	(4) $ln(Y_{ik})$	(5) $ln(Y_{ik})$	(6) $ln(Y_{ik})$	(7) $ln(Y_{ik})$	(8) $ln(Y_{ik})$	(9) $ln(Y_{ik})$
$Elev < 10m_i$	0.182	0.309	0.272	0.053	0.390	0.364	-0.028	0.004	0.024
	(0.037)	(0.060)	(0.056)	(0.012)	(0.389)	(0.410)	(0.267)	(0.270)	(0.209)
$Elev < 10m_i \times DemocracyIndicator_k$		-0.175	-0.191	0.007				0.009	-0.225
		(0.067)	(0.076)	(0.037)				(0.029)	(0.078)
$Elev < 10m_i \times ln(GDPpercapita)_k$					-0.021	-0.023	0.009	0.005	0.028
					(0.037)	(0.040)	(0.029)	(0.029)	(0.023)
Observations	3,642,083	3,610,249	3,610,249	3,610,249	3,562,613	3,562,613	3,562,613	3,543,409	3,543,409
Country FE	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
City FE	No	No	No	Yes	No	No	Yes	Yes	No
River and Coast FE	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes

Notes: The regressions reported in this Table are variations on Equation 2 and include the full global sample of all urban areas.
The dependent variable in all regressions $ln(Y_{ik})$ is the natural log of mean light intensity (measured in 2012) at each gridpoint i (located in in country k).
$Elev < 10m_i$ is a dummy variable for locations that are less than 10m above sea level.
$ln(GDPpercapita)_k$ is the natural log of GDP per capita (in 2011) in country k (data are from the Penn World Tables v8).
$DemocracyIndicator_k$ is a dummy for countries with a Polity IV score (in 2008) greater than or equal to 5.
Regressions with river and coast controls include dummies for locations within 10km of the nearest river or coast.
Robust standard errors, clustered by country, in parentheses.
Robust standard errors are clustered by country.

Table A2: Main effects of flood on light, city-year panel

	(1)	(2)	(3)	(4)	(5)	(6)
	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$
$Flood_{jt}$	-0.017			-0.019		
	(0.007)			(0.006)		
$Precip > 500mm_{lt}$		-0.039			-0.040	
		(0.011)			(0.015)	
$Precip > 1000mm_{lt}$			-0.057			-0.058
			(0.014)			(0.014)
$ln(light_{t-1})$				Yes	Yes	Yes
Observations	10,363	10,363	10,363	9,878	9,878	9,878
No. of urban areas	1,817	1,817	1,817	1,702	1,702	1,702

Notes: The results presented in this Table correspond to Equation 4 and use the sample of cities affected by at least one of the large flood events in our data. The dependent variable in all regressions $ln(Y_{jkt})$ is the natural log of mean light intensity for each city j (located in country k) in year t.
$Flood_{jt}$ is a dummy indicating whether or not city j was hit by a large flood in year t.
$Precip > 1000mm_{lt}$ ($> 500mm_{lt}$) indicates locations that experienced monthly precipitation of 1000mm (500mm) or more in year t.
All regressions include year fixed effects, city fixed effects and country-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A3: Recovery, gridpoint year panel, extreme precipitation (500mm) indicator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Precip > 500mm_{jt}$	-0.025					-0.027				
	(0.008)					(0.008)				
$Precip > 500mm_{jt-1}$		0.004					0.003			
		(0.012)					(0.011)			
$Precip > 500mm_{jt-2}$			0.002					0.003		
			(0.006)					(0.007)		
$Precip > 500mm_{jt-3}$				-0.013					-0.012	
				(0.012)					(0.012)	
$Precip > 500mm_{jt-4}$					-0.009					-0.009
					(0.010)					(0.010)
$ln(light_{t-1})$						Yes	Yes	Yes	Yes	Yes
Observations	1,422,018	1,422,018	1,422,018	1,422,018	1,422,018	1,392,501	1,392,501	1,392,501	1,392,501	1,392,501
No. of gridpoints	243,303	243,303	243,303	243,303	243,303	235,460	235,460	235,460	235,460	235,460

Notes: The results in this Table are variations on Equation 3 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
$Precip > 500mm_{jt+s}$ indicates locations that experienced monthly precipitation of 500mm or more in year $t+s$.
All regressions include year fixed effects, gridpoint fixed effects and country-specific trends.
Columns (6) to (10) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A4: Recovery, gridpoint year panel, extreme precipitation (1000mm) indicator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Precip > 1000mm_{jt}$	-0.080					-0.083				
	(0.018)					(0.018)				
$Precip > 1000mm_{jt-1}$		0.054					0.053			
		(0.033)					(0.034)			
$Precip > 1000mm_{jt-2}$			0.004					0.002		
			(0.020)					(0.019)		
$Precip > 1000mm_{jt-3}$				0.002					0.004	
				(0.013)					(0.012)	
$Precip > 1000mm_{jt-4}$					0.001					0.003
					(0.029)					(0.028)
$ln(light_{t-1})$						Yes	Yes	Yes	Yes	Yes
Observations	1,422,018	1,422,018	1,422,018	1,422,018	1,422,018	1,392,501	1,392,501	1,392,501	1,392,501	1,392,501
No. of gridpoints	243,303	243,303	243,303	243,303	243,303	235,460	235,460	235,460	235,460	235,460

Notes: The results in this Table are variations on Equation 3 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
$Precip > 1000mm_{jt+s}$ indicates locations that experienced monthly precipitation of 1000mm or more in year $t+s$.
All regressions include year fixed effects, gridpoint fixed effects and country-specific trends.
Columns (6) to (10) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A5: Recovery, city-year panel

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$	$ln(Y_{jkt})$
$Flood_{jt}$	-0.017					-0.019				
	(0.007)					(0.006)				
$Flood_{jt-1}$		-0.003					-0.008			
		(0.014)					(0.019)			
$Flood_{jt-2}$			0.017					0.017		
			(0.016)					(0.015)		
$Flood_{jt-3}$				0.004					-0.008	
				(0.009)					(0.021)	
$Flood_{jt-4}$					0.014					0.014
					(0.009)					(0.011)
$ln(light_{t-1})$						Yes	Yes	Yes	Yes	Yes
Observations	10,363	10,281	10,315	10,352	10,338	9,878	9,869	9,833	9,785	9,796
No. of urban areas	1,817	1,814	1,820	1,818	1,819	1,702	1,707	1,707	1,703	1,712

Notes: The results presented in this Table correspond to Equation 4 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{jkt})$ is the natural log of mean light intensity in each city j (located in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
All regressions include year fixed effects, city fixed effects and country-specific trends.
Columns (6) to (10) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A6: Interactions with elevation, extreme precipitation (1000mm) indicator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Precip > 1000mm_{jt} \times elev_{<10i}$	-0.120			-0.122			-0.100		
	(0.019)			(0.020)			(0.016)		
$Precip > 1000mm_{jt} \times elev_{10+i}$	-0.052			-0.056			-0.031		
	(0.022)			(0.021)			(0.011)		
$Precip > 1000mm_{jt-1} \times elev_{<10i}$		0.111			0.111			0.117	
		(0.015)			(0.016)			(0.017)	
$Precip > 1000mm_{jt-1} \times elev_{10+i}$		0.020			0.018			0.033	
		(0.035)			(0.035)			(0.030)	
$Precip > 1000mm_{jt-2} \times elev_{<10i}$			0.006			0.007			-0.022
			(0.009)			(0.014)			(0.011)
$Precip > 1000mm_{jt-2} \times elev_{10+i}$			0.003			0.001			-0.005
			(0.028)			(0.026)			(0.023)
$ln(light_{t-1})$				Yes	Yes	Yes			
Observations	1,422,018	1,422,018	1,422,018	1,392,501	1,392,501	1,392,501	1,422,018	1,422,018	1,422,018
No. of gridpoints	243,303	243,303	243,303	235,460	235,460	235,460	243,303	243,303	243,303

Notes: The results presented in this Table correspond to Equation 5 and use the sample of cities affected by at least one of the large flood events in our data.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Precip > 1000mm_{lt+s}$ indicates locations that experienced monthly precipitation of 1000mm or more in year $t+s$.
$Elevation_h$ is a dummy for elevation band h, where h is either less than 10m above sea level, or 10m or more above sea level.
All regressions include year fixed effects and gridpoint fixed effects.
Columns (1) to (6) include country-specific trends. Columns (7) to (9) include city-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A7: Interactions with elevation, excluding locations within 10km of rivers and coasts

	(1) $ln(Y_{ijkt})$	(2) $ln(Y_{ijkt})$	(3) $ln(Y_{ijkt})$	(4) $ln(Y_{ijkt})$	(5) $ln(Y_{ijkt})$	(6) $ln(Y_{ijkt})$
$Flood_{jt} \times elev_{<10i}$	-0.030			-0.032		
	(0.007)			(0.006)		
$Flood_{jt} \times elev_{10+i}$	-0.020			-0.021		
	(0.013)			(0.014)		
$Flood_{jt-1} \times elev_{<10i}$		0.015			-0.009	
		(0.011)			(0.011)	
$Flood_{jt-1} \times elev_{10+i}$		-0.001			-0.016	
		(0.012)			(0.017)	
$Flood_{jt-2} \times elev_{<10i}$			0.037			0.038
			(0.017)			(0.017)
$Flood_{jt-2} \times elev_{10+i}$			0.017			0.017
			(0.014)			(0.014)
$ln(light_{t-1})$				Yes	Yes	Yes
Observations	814,294	810,524	812,476	795,536	792,801	790,097
No. of gridpoints	139,712	139,683	140,298	134,640	134,600	134,474

Notes: The results presented in this Table correspond to Equation 5 and use the sample of cities affected by at least one of the large flood events in our data, restricted to exclude gridpoints within 10km of the nearest river or coast.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t+s$.
$Elevation_h$ is a dummy for elevation band h, where h is either less than 10m above sea level, or 10m or more above sea level.
All regressions include year fixed effects and gridpoint fixed effects and country-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A8: Interactions with elevation, excluding cities entirely less than 10m above sea level

	(1)	(2)	(3)	(4)	(5)	(6)
	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$	$ln(Y_{ijkt})$
$Flood_{jt} \times elev_{<10i}$	-0.021			-0.022		
	(0.006)			(0.007)		
$Flood_{jt} \times elev_{10+i}$	-0.019			-0.020		
	(0.012)			(0.012)		
$Flood_{jt-1} \times elev_{<10i}$		0.012			-0.006	
		(0.008)			(0.010)	
$Flood_{jt-1} \times elev_{10+i}$		-0.007			-0.022	
		(0.014)			(0.018)	
$Flood_{jt-2} \times elev_{<10i}$			0.046			0.045
			(0.016)			(0.016)
$Flood_{jt-2} \times elev_{10+i}$			0.007			0.007
			(0.011)			(0.011)
$ln(light_{t-1})$				Yes	Yes	Yes
Observations	1,379,280	1,372,088	1,375,024	1,351,484	1,345,342	1,339,592
No. of gridpoints	235,874	235,861	236,793	228,408	228,358	228,261

Notes: The results presented in this Table correspond to Equation 5 and use the sample of cities affected by at least one of the large flood events in our data, restricted to exclude cities that are entirely less than 10m above sea level.
The dependent variable in all regressions $ln(Y_{ijkt})$ is the natural log of mean light intensity at each gridpoint i (located in city j in country k) in year t.
$Flood_{jt+s}$ is a dummy indicating whether or not city j was hit by a large flood in year $t + s$.
$Elevation_h$ is a dummy for elevation band h, where h is either less than 10m above sea level, or 10m or more above sea level.
All regressions include year fixed effects and gridpoint fixed effects and country-specific trends.
Columns (4) to (6) corrected by the Arellano-Bond methodology using $ln(lights_{t-2})$ as an instrument for $ln(lights_{t-1})$.
Robust standard errors, clustered by country, in parentheses.

Table A9: Lights, GDP and floods for a panel of Indian districts

	(1) Ln light	(2) Ln light	(3) Ln GDP	(4) Ln GDP	(5) Ln light	(6) Ln GDP	(7) Ln light	(8) Ln light	(9) Ln GDP	(10) Ln GDP
Ln GDP	0.441	0.481			0.189					
	(0.039)	(0.027)			(0.051)					
Ln light			1.108	0.932		0.168				
			(0.117)	(0.104)		(0.121)				
Flood_t							-0.020		-0.005	
							(0.014)		(0.013)	
$\text{Flood}_t - 1$								-0.023		0.015
								(0.027)		(0.017)
Year 2003 only	Yes	Yes	Yes	Yes						
State f.e.		Yes		Yes						
District f.e.					Yes	Yes	Yes	Yes	Yes	Yes
Year f.e.					Yes	Yes	Yes	Yes	Yes	Yes
Observations	491	491	491	491	3,157	3,157	1,683	1,192	1,804	1,285

Notes: This table shows the correlation between log lights and log GDP in the year 2003 (Columns 1-4), and for an annual panel (2003-2008, columns 5-6) at the level of districts in India. The Table also reports results of regressions of log light and log GDP on a flood indicator using the district level panel data for India (Columns 7-10).
The light variable gives the mean light measure for each district. Flood is a dummy variable indicating if any gridpoint in the district was subject to one of the floods in our dataset in each year.
Robust standard errors are clustered at the level of state throughout the table.

Figure A1: This figure shows the coefficients from Columns 1-5 from Table 5, and their 95 percent confidence intervals. The figure includes additional years before the flood not shown in the table. Year 0 indicates the year of the flood.

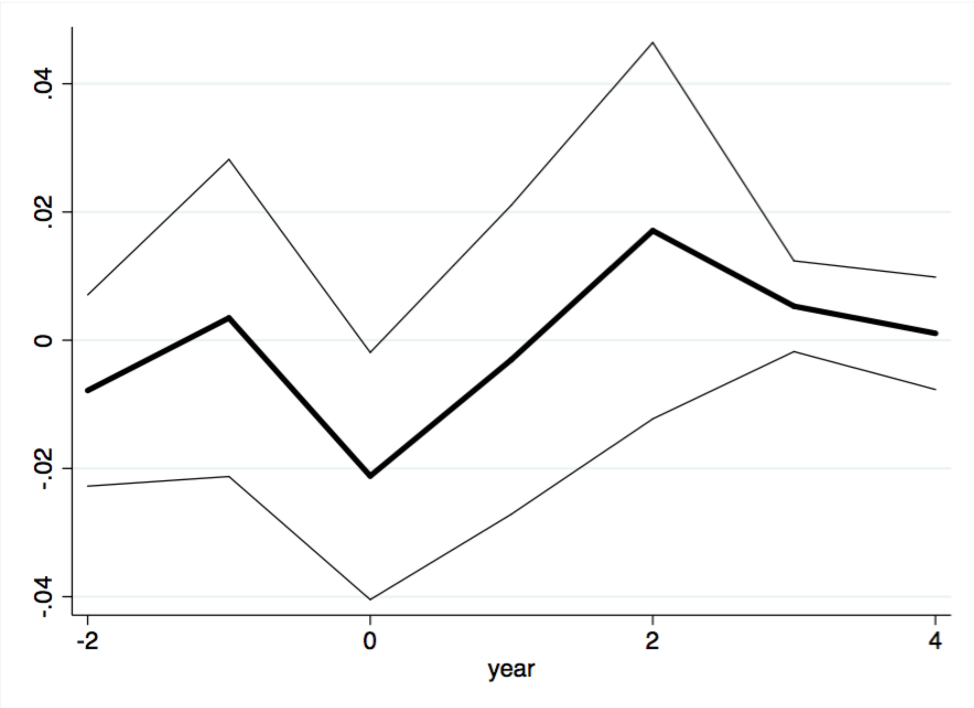

Figure A2: This figure shows the impact of a flood observed at time 0 separately for urban areas in our sample that we define as New, and poorly lit ("Poorlit") as follows: "New" areas are defined as having light $= 0$ in the year 1992. "Poorlit" is defined as areas with light > 0 and $<=$ some cutoff in the year 1992. In Panels A and B we chose arbitrary, round numbers as light cutoffs (10 and 20) for poorly lit areas. In Panel C we chose the cutoff such that the resulting number of observations is close to the number of observations in the "new" category. This turns out to be at $<= 8$. Median light intensity in our main dataset of urban areas is 14, the 25th percentile value is 7 and the 75th percentile 31. The regression specification to which these coefficients correspond is similar to the specification used for Table 7. The figure includes additional years before the flood not shown in the table.