I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14243

Cities and the Sea Level

Yatang Lin
Thomas K.J. McDermott
Guy Michaels

APRIL 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14243

Cities and the Sea Level

Yatang Lin
Hong Kong University of Science and Technology

Thomas K.J. McDermott
National University of Ireland Galway and LSE

Guy Michaels
LSE and IZA

APRIL 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Cities and the Sea Level*

Construction on low elevation coastal zones is risky for both residents and taxpayers who bail them out, especially when sea levels are rising. We study this construction using spatially disaggregated data on the US Atlantic and Gulf coasts. We document nine stylized facts, including a sizeable rise in the share of coastal housing built on flood-prone land from 1990-2010, which concentrated particularly in densely populated areas. To explain our findings, we develop a model of a monocentric coastal city, which we then use to explore the consequences of sea level rise and government policies.

JEL Classification: R11, Q54, R14

Keywords: cities, climate change, sea level rise

Corresponding author:
Guy Michaels
CEP, London School of Economics
Houghton Street
WC2A 2AE, London
United Kingdom

E-mail: g.michaels@lse.ac.uk

* We thank Emek Basker, Vernon Henderson, Alan Manning, Henry Overman, Steve Pischke, Eric Strobl, and John Van Reenen, and seminar/conference participants at Bank of Spain, ERSA 2018, IFAU, LSE, NUI Galway, University of Kent, University of Stockholm, UPF, and the US Census Bureau for their helpful comments. We are grateful to the Hong Kong University Grant Committee, Irish Research Council award no. GOIPD/2017/1147, and the ESRC's Centre for Economic Performance for their generous financial support.

1 Introduction

Where do people build houses in low elevation coastal zones (LECZ)? Some construction in LECZ puts residents at risk and exposes taxpayers who pay for bailouts, problems which sea level rise (SLR) exacerbates. We study the location of existing and new housing on LECZ in the US, using SLR maps and census data at a fine spatial scale. Using these data, we document nine stylized facts. One of our findings is that within 10km of the Atlantic and Gulf coasts, 12 percent of the 1990 housing stock was in flood-prone areas, but this share increased to 26 percent for net new construction from 1990-2010. We also show that this new construction in flood-prone areas concentrated particularly in densely populated areas. To understand why this happened, we develop a model of a monocentric coastal city, which explains our stylized facts. We then extend the model to explore the potential consequences of SLR in LECZ.

LECZ are often attractive places to live. Neumann et al. (2015) estimate that in 2000 they were home to around 10.2% of the world's population (625 million people), a figure expected to grow significantly by 2050.[1] But this growing attraction of LECZ poses problems, due to their susceptibility to floods. Over the past 36 years, floods worldwide killed more than 680 thousand people and displaced more than 650 million (Brakenridge 2021). And the conditions in LECZ are expected to worsen, as climate change warms the world's oceans and melts some of its glaciers. The UN's Intergovernmental Panel on Climate Change, IPCC (Pörtner et al. 2019), forecasts that mean global sea levels will rise by 43-84 centimeters by 2100. To compound the problem, climate change may also increase the severity of tropical storms (Berardelli 2019). All this is expected to rapidly raise global annual flood costs, which may exceed $1 trillion by 2050 (Hallegatte et al. 2013).

The problem of flooding – ranging from nuisance flooding to extreme flooding by tropical storms – is acutely felt in the US, where LECZ were home to 23.4 million people (8.2% of the population) in 2000 (Neumann et al. 2015). According to the National Oceanic and Atmospheric Administration (NOAA), since 2005, the US has suffered $1.24 trillion in economic losses from 173 major weather and climate disasters (NOAA 2021a). More than 83 percent of these losses were due to tropical storms, other severe storms, and flooding. And a recent government report estimates the expected annual losses from tropical storms in the US at 57 billion dollars at current conditions, including 19.4 billion dollars in public funds (US Congressional Budget Office 2019). Additional expenditures on social insurance in response to tropical storms may further increase the burden for taxpayers (Deryugina 2017). And to make matters worse, SLR on the US Atlantic and Gulf coasts is even more rapid than the global mean (Dahl et al. 2017).

Against this backdrop, it is important to understand where housing construction in coastal areas is taking place. To do so, we use high-resolution maps of sea level rise (NOAA 2021b), which identify locations that will be under water at high tide if sea levels rise by 1 foot (approximately 30.5 cm). Even without SLR, high tides and storms make these locations prone to flooding. To measure locational outcomes, we use data on housing units from the Census and Annual Communities Survey from

[1] Neumann et al. (2015) refer to LECZ as the contiguous and hydrologically connected zone of land along the coast and below 10 m of elevation. The coastal areas we study are generally low-elevation and close to the coast, as we define and discuss in Section 2.

1990-2010 at the finest spatial scale available - census blocks.[2] We complement these with similarly disaggregated data on built cover from 1996-2010 (NOAA 2021c).[3]

We use these data to document nine stylized facts, of which the first six are cross-sectional. The first three describe the shape of coastal locations in 1990. First, housing unit density peaks near – but not right at – the coast, and it declines more steeply on the coast side. Second, census-designated places near the coast are asymmetric – their Central Business District (CBD) is closer to their coast side edge – while places further inland are symmetric. Third, the asymmetry near the coast is more pronounced for large places.

Whereas the first three stylized facts tell us how housing concentrates near the coast, the next three begin to tell us why this happens. Fourth, we find that census blocks that are prone to SLR are much more sparsely built; but conditional on SLR-proneness, blocks closer to the coast are more densely built. Fifth, as we approach the coast, SLR-proneness rises steeply. Sixth, damages from flooding also rise steeply as we approach the coast. Together, these stylized facts suggest a tension between the amenity of coastal proximity, and the disamenity of flood-proneness, which increases steeply near the coast. These two forces balance at a bliss-point: close to the coast, but not too close.

While the first six stylized facts describe coastal housing around 1990, the last three describe how it has changed from 1990-2010. Our seventh stylized fact is that despite the considerations discussed above, much construction near the coast in recent decades took place in areas prone to SLR (and flooding). As mentioned above, net new construction from 1990-2010 was more than twice as prevalent in SLR-prone locations as in the 1990 stock of housing. Eighth, SLR-prone areas were more likely to be developed in dense census tracts, but not in sparse ones. Finally, our ninth stylized fact is that in the densest census tracts, new construction focused on medium-risk SLR-prone areas, still avoiding the riskiest ones.

We show that our nine stylized facts are robust to excluding census blocks, which were mostly shielded from private residential construction, because they are either protected areas, military bases, or parks. We also find evidence consistent with our stylized facts when we use data on built area, which cover all construction rather than just housing data.

To account for the nine stylized facts, we develop a model of a monocentric coastal city. In the model, coastal areas are characterized by both an amenity, which declines linearly in the distance to the coast, and a disamenity (flood-proneness), which declines convexly in distance to the coast. The city founder chooses a location that trades off these two factors – close to the coast, but not right at it.[4] This location becomes the city's focal point – the Central Business District (CBD). Residents then choose where to live, and they prefer locations close to the CBD, both because of their high net amenity value and because of the shorter commute. Housing density peaks around the CBD, but declines more steeply on the coast-side, because of the convex flood-proneness. The city expands over time into previously empty areas on both sides. On the coast side, this expansion involves

[2] We also use data at higher levels of aggregation on census tracts and census-designated places.

[3] The land cover data are derived from 30 x 30 meter Landsat satellite images, which we aggregate to 150 x 150 meter cells (the approximate size of the median census block) for computational reasons. The land cover data offer the advantage of a regular partition and coverage of non-residential construction, but their interpretation is less precise (e.g., they include roads and parking lots).

[4] Our model abstracts from productivity differences between locations that are, in any case, close to the coast, but adding productivity differences does not substantially change the picture.

building on increasingly flood-prone land.

After explaining how the model accounts for the nine stylized facts, we extend it in different ways. Our first extension allows for sea level rise. In the second, we allow for a finite number of high-elevation areas near the coast, which are safe from flooding. While most locations right by the coast are still flood-prone (and unpopulated), the handful of elevated locations there command high prices, consistent with what we see in the data. Third, we allow for costly and irreversible conversion of land to housing from alternative uses, which makes the developers' decisions dynamic rather than static. Finally, we examine government subsidies to flood-prone areas.

We then simulate our model to explore challenges that low-elevation coastal cities may face in the coming decades. These simulations point to four potential concerns for low-elevation coastal cities. First, the problem of housing in flood-prone locations looks set to worsen, either because cities expand towards the coast, or because of SLR, or because both happen simultaneously. This development threatens to increase flooding costs for both residents and taxpayers. Second, even if LECZ cities grow on aggregate, some neighborhoods within them may experience economic decline, as increased flood risk causes demand for housing to decline. This problem is exacerbated in the case of economically stagnant cities. Third, SLR further distorts the shape of LECZ cities, significantly lengthening the time costs of commuting to work. Finally, these cities face a potential crisis if their CBD comes under threat of being permanently submerged.

The main contributions of our paper are fourfold. First, we assemble a new dataset on the location of housing and flood risk, which covers thousands of kilometers of coast, spanning major urban centers, small towns, and rural areas. The data, which cover two decades, are at a highly disaggregated spatial scale. They include information on housing from the census and land cover from satellite imagery, as well as measures of SLR-proneness, flood damages, and regulatory restrictions. These data allow us to explore construction in areas where flood risks for residents and taxpayers are both high and rising, due to climate change. Second, we use these data to document how the existing housing stock and new construction vary by distance to the coast. The result is a novel and detailed picture of housing in LECZ, and its relationship to the vulnerability of different locations to flooding and SLR. Third, we develop a model, which provides a parsimonious explanation for our findings. The model answers questions such as: why does housing concentrate near, but not right at, the coast? Why are coastal cities asymmetric? Why is new housing in LECZ increasingly built on flood-prone areas, which were previously avoided? And why does this happen especially on the urban fringes? Finally, we extend our model and use it to study how SLR may reshape cities, and consider implications for rising costs of flooding and taxpayer subsidies, the economic decline of some neighborhoods, and lengthening commutes.

Our paper is related to the literature on the importance of urban amenities (Glaeser, Kolko and Saiz 2001) and the attraction of coastal areas (Rappaport and Sachs 2003). Though attractive, LECZ are also prone to flooding, and there are reasons to worry that they might be built over too densely. First, the flood-proneness of LECZ creates moral hazard, which results in overbuilding when taxpayers bear some of the costs of reconstruction following floods (Kydland and Prescott 1977) and of public construction of flood defences. Second, flood risk may be under-appreciated by residents, because official flood maps do not fully reflect current and future risks (US Department of Homeland

Security 2017), or because people are myopic (Burningham et al. 2008 and Pryce et al. 2011).[5] Our paper identifies a third reason why people build in flood-prone coastal areas: to reduce commuting costs to jobs in major city centers, which are often near the coast.

Our paper is also related to the literature on physical barriers to city growth. Building on Saiz (2010), who studies "hard" physical barriers to city growth, we characterize "soft" barriers, such as flood-prone areas. Soft barriers are locations that are not used for housing development in most circumstances, but are nevertheless built on as cities expand. Construction on soft barriers may involve risks not only to residents but also externalities (e.g., for taxpayers or the environment), which may necessitate policy intervention.[6] Also closely related is Harari (2020), who studies how physical barriers distort the shape of cities and lengthen commutes. Our paper differs in its geographic focus (the US as opposed to India), and more importantly in its study of flooding and SLR, which further distorts the shape of coastal cities.

Another related paper is Magontier, Solé-Ollé and Viladecans-Marsal (2019), who study political economy of coastal destruction in Spain. We differ in our focus on market forces (rather than the political economy), and in our study of the role of SLR.

Whereas Balboni (2020) studies exposure of Vietnam roads to SLR, we focus on internal city development (rather than on intercity roads), and how it evolves in the US. Other economic studies of the consequences of SLR include Hallegatte et (2013) and Desmet et al. (2021), who quantify its global costs. Our paper differs in its study of urban structure and how cities may expand even as sea levels rise.

Finally, our paper is also related to the literature on path dependence in city location (Bleakley and Lin 2012) and the adaptation of cities to large-scale environmental shocks, such as Hornbeck and Keniston (2017) and Kocornik-Mina et al. (2020). Our contribution here is to explore how coastal cities evolve and how SLR reshapes them.

2 Data

2.1 The area and units of analysis

In our analysis we focus on areas within 10km of the US Atlantic and Gulf coasts. This choice of area reflects a tradeoff between three considerations: a focus on flood-prone and SLR-prone LECZ; an analysis of fine spatial units; and computational constraints. First, we are interested in low-elevation coastal zones, and especially those that are prone to flooding and vulnerable to sea level rise. The area that we study spans the coastal edges of the Atlantic Coastal Plain and the Gulf Coastal Plain, both of which include many low-elevation coastal locations. The area that we study is highly prone to flooding: it held 1.7 percent of US housing units in 1990 (and about 2 percent in 2010), but accounted for 36 percent of the value of National Flood Insurance Program (NFIP) claims from 1973-2019. This area also experienced some of the fastest rates of local sea level rise in the world during the 20th century, a trend which is expected to continue and raise the frequency and severity of floods in these

[5] Ortega and Taşpınar (2018), Gibson and Mullins (2020), Hino and Burke (2020), and Keys and Mulder (2020) explore the updating of house prices following information on flooding and SLR.

[6] Construction in areas prone to wildfires is another example of soft barriers.

locations (Dahl et al. 2017).

Second, we analyze small spatial units, where the intersection of flood-proneness and construction can be pinpointed. Much of our analysis is at the level of census blocks – the smallest geographic units used by the US Census Bureau.[7] We also make use of more aggregated geographic units, including census designated places (to study the asymmetry of coastal places), and census tracts (the finest disaggregation for which we have data on damages from floods from the National Flood Insurance Program). The blocks and other geographical units that we use are from the 1990 census, with later data matched onto them, as detailed below and in the Data Appendix. All the census datasets that we use are sourced from the NHGIS data archive (Manson et al. 2019).

Finally, we face computational constraints in processing high-resolution data covering such a large area. Since, as we discuss below, SLR-prone land and NFIP damages are heavily concentrated within one or two km of the coast, we decided not to explore the area further inland than 10 km.[8]

A map of the area that we study is shown in Appendix Figure A1.[9] Since the coast is not straight but winding, the area within 0-1 km of the coast is larger than the area within 1-2 km of the coast, and so on. In our analysis we take this into account, as we explain below.

2.2 Main outcomes: housing and land cover data

Our housing data come from the US Census and Annual Communities Survey, observed in 1990 and 2010.[10] We harmonize all our census data to the geographical units (block boundaries) of the 1990 census, with 2010 data matched to 1990 in proportion to area shares, as described in detail in the Data Appendix. Our main dataset, composed of census blocks within 10km of the US Atlantic and Gulf coasts, includes some 544,065 observations, covering a total area of 128,757 sq km. The median area of blocks in our data is 0.021 sq km (like a square with 145 meters on each side), and the median number of housing units per block in 1990 is 12. At the level of blocks, we observe the number of housing units and the median price of owner-occupied dwellings (housing units).[11]

As an alternative measure of the extent and intensity of development in coastal areas, we use land cover data based on Landsat Thematic Mapper (TM) satellite imagery (NOAA 2021). The Landsat data come in the form of a raster dataset where each 30m x 30m observation (or pixel) has been assigned to one of 25 land cover categories.[12] In our analysis, we focus on the four developed categories, which represent different extents of constructed surfaces (including buildings, roads, and parking lots).[13] For computational reasons, we aggregate the Landsat data to 150m x 150m cells (the

[7] We complement our census block data with a gridded dataset of 150m x 150m cells, which is the approximate size of the median census block. More details on this alternative dataset are included below and in the Data Appendix.

[8] The only exceptions where we show areas further inland than 10 km are in a few illustrative examples in our appendix, as discussed below. Our economic analysis consistently focuses on the area within 10 km of the coast.

[9] We use the Database of Global Administrative Boundaries (GADM 2018) to define the coast. This shapefile includes sections of major rivers, such as the Charles in Boston, East River and the Hudson River in New York City, and the Potomac in Washington, DC, as part of the coastline. But lakes and upstream sections of rivers are typically excluded from the coast shapefile, and consequently Philadelphia, New Orleans, and Houston, are largely outside our dataset. Overall, the area we study consists of parts of 18 states and the District of Columbia, as listed in the Data Appendix.

[10] What we refer to as "2010" is more precisely data for 2006 - 2010 from the Annual Communities Survey.

[11] Just over a fifth (22.4%) of blocks in our sample were empty - i.e. had zero housing units - in 1990.

[12] These categories include various classifications of open water, wetlands, agricultural land, forest etc., as detailed here: https://coast.noaa.gov/digitalcoast/training/ccap-land-cover-classifications.html.

[13] The four developed categories in the Landsat data are: "developed - high intensity" where constructed materials

approximate size of the median census block in our data), taking the midpoint values of the four developed categories to arrive at a measure of the fraction of each cell's land area that is developed (i.e., covered in constructed materials). We observe this variable in 1996 and 2010, the earliest and latest years for which we have complete Landsat data.[14]

2.3 SLR data

Our data on sea level rise come from detailed maps of areas anticipated to be inundated for various future sea level rise scenarios, which we obtained from NOAA's digital coast platform (Marcy et al. 2011). These maps show inland extent of inundation for scenarios of sea level rise from 0 to 6 feet. Importantly, the mapping process also takes account of major federal leveed areas, which are assumed, for the purposes of creating these inundation maps, to be high enough and strong enough to prevent inundation, regardless of the SLR scenario.[15]

In our analysis we focus on the share of an area (e.g., a block), which would be under water at high tide if SLR is 1 foot (approx. 30.5cm). Information on sea level rise was added to the blocks (and cells) by intersecting the shapefiles for blocks (cells) with shapefiles of areas expected to be inundated for 1ft of sea level rise using GIS software. We then calculate the share of each census block (or cell) that is exposed to 1ft of sea level rise, which we refer to as $SLR1ft$. We further define low-risk areas as blocks (or cells) in our data where $SLR1ft = 0$; medium-risk, where $SLR1ft \in (0, 0.5]$; and high-risk, where $SLR1ft \in (0.5, 1]$. Of the blocks in our sample, 86% are low-risk. Overall, the mean share 1ft SLR for the entire sample of blocks is 0.045, and the area-weighted mean share is around 0.18.[16]

2.4 Building restrictions

Our dataset also includes information on areas where building construction may be restricted. While such restrictions may be an endogenous response by governments at different levels to the danger of building close to the coast, we nevertheless examine the role that such regulations may have in our setting. Here we discuss three different types of regulations: restricted areas, where housing development may be particularly constrained; state "setback lines" close to the coast, beyond which construction may be more regulated; and local government regulations on building density.

We begin by gathering data on what we refer to as restricted development areas – i.e. areas where development is likely to be prohibited or highly constrained, for one of three reasons: due to a protected areas designation, including for conservation, natural resource management or recreation (e.g. the Everglades in Florida); because of the presence of a city park (e.g. Central Park in New York City or Boston Common), or where land is owned by the military (e.g. Norfolk naval base in

account for 80 to 100 percent of the total cover at that location; "developed medium intensity" (50-79 percent constructed material); "developed - low intensity" (21-49 percent); and "developed - open space", where constructed material accounts for less than 20 percent of land cover.

[14] Our cells dataset includes over 6 million observations, or cells within 10km of the Atlantic and Gulf coasts. The mean share developed in 1996 is 0.066. More than 70% of cells in our data have share developed $= 0$ in 1996. By construction, the share developed measure is top-coded at 0.9, but there are fewer than 7,000 cells in our data with this value for developed share in 1996.

[15] The SLR maps assume that New Orleans is safe even from 6-foot of SLR.

[16] A recent study has suggested that many areas on US Atlantic and Gulf coasts could experience 1ft SLR as early as 2045 (Dahl et al. 2017).

Virginia). As a robustness check, we replicate all our main findings excluding blocks (or cells) where 50% or more of the land area is accounted for by the sum of any of these three restricted development areas. While blocks that meet our restricted development area criterion make up just 3.5% of the observations (18,863 blocks), they account for a sizeable share of the total land area in our sample (over 19%). The vast majority (over 88%) of the land area in our sample that we identify as restricted for development is accounted for by areas that have been designated as protected. New development in restricted areas is quite limited – some 90,000 new housing units were added to blocks with at least 50% land area restricted for development between 1990 and 2010, compared with 3.24 million new housing units across all blocks in our sample. Full details of the data sources and definitions used to identify restricted development areas are included in the Data Appendix.

Another factor which may influence construction is the existence of "setback lines", which are designed to protect fragile environments close to the coast. Unlike the EU and other countries, the US does not have a federal "setback line", which prohibits construction within a fixed distance of the shoreline (Simpson et al. 2012). Instead, there are setback lines in some states.[17] The geographic location of those lines differs both between and within states, but in many cases they appear to be drawn within a few tens of meters of the coast. Construction is not necessarily banned even beyond setback lines, but it may be more regulated. While we have no shapefiles showing the areas covered by these lines, we have examined the lines themselves in one state of particular importance – Florida.[18] Much of Florida's line (Coastal Construction Control Line Program 2021) runs near the sea-facing edge of its barrier islands, leaving many SLR-prone areas open to construction. And visual inspection shows that there are buildings even beyond that line. Nevertheless, it is likely that the existence of "setback lines" contributes to the low housing density in the immediate vicinity of the coast.

Moving from state regulations to the more local level, we consider housing market land restrictions on building density. Specifically, we use Density Restriction Index (DRI) from the Wharton Residential Land Use Regulation Index (Gyourko et al. 2019). Since these data are at a different level of aggregation than the one we use, we aggregate them to the county level, and match them to our block-level data.

2.5 Additional data sources

Besides the datasets discussed above, other data used in our analysis include information on census designated places (both polygons and points), sourced from the NHGIS data archive (Manson et al. 2019), which we use to define place extents and the location of the CBD of each place. NHGIS assigns place points based on the Geographic Names Information System (GNIS) coordinates of each place's historical or functional center (typically the central business district). Other characteristics of places – e.g. their size and asymmetry – are calculated based on information from blocks aggregated to places. Specifically, we define place size in two ways: one is the linear distance $|x_L - x_R|$, where x_L is approximated as the minimum distance to coast from the centroid of any block in a place, and

[17] Of the states in our dataset, Alabama, Delaware, Florida, Georgia, Maine, New Hampshire, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, South Carolina, Texas, and Virginia have setback lines (Simpson et al. 2012).

[18] Even in the case of Florida we do not have the shapefile of the restricted area, but only of the line itself.

x_R the maximum distance to coast from the centroid of any block in the same place; the other is the sum of the area of blocks in the place. Place asymmetry is defined as the ratio of distances $\frac{|x_R - x_0|}{|x_R - x_L|}$, where x_0 is distance to the coast from the CBD.[19]

Data on historical damages from coastal flooding are taken from the National Flood Insurance Program (NFIP), operated by FEMA, which subsidizes flood insurance provision. In particular, we use data on insured losses from coastal floods, available at the census tract level, from 1973 to 2019. There are two points to note about NFIP. First, NFIP includes an implicit subsidy component. For example, the Congressional Budget Office (CBO 2017) notes that in 2016, "the overall shortfall of $1.4 billion is attributable largely to premiums' falling short of expected costs in coastal counties, which constitute roughly 10 percent of all counties with NFIP policies but account for three-quarters of all NFIP policies nationwide ... the net short- fall measured over all coastal counties is $1.5 billion, whereas the net surplus measured over all inland counties is $200 million." A recent analysis concluded that while NFIP's shortfalls cannot be attributed to any single incident, it borrowed significantly following Hurricanes Katrina in 2005 and Sandy in 2012. In 2017, as NFIP reached its borrowing cap of $30.5 billion, Congress canceled $16 billion of its liabilities, to allow NFIP to borrow more in response to Hurricanes Harvey, Irma, and Maria (Peterson Foundation 2020). It is also noteworthy that by our estimates, claims made to NFIP grew at a rate of around 4-5 percent in real terms from 1978-2019.[20]

While claims made under the NFIP by no means capture the totality of economic losses from coastal floods (or in fact the totality of residential losses from flooding, as some damage is uninsured), the NFIP data have the advantage of being available at a relatively fine level of geographic disaggregation – the census tract level – which makes these data well suited to our task of estimating how damages from flooding vary with distance from the coast. We convert these claims data to 2020 US dollars and aggregate the damages data across the entire period available (1973-2019).[21] We also normalize the figures by dividing the damages by the number of housing units in each tract from 2014-2018 (Manson et al. 2019).

Information on public spending associated with coastal flooding, which we use to calculate the share of damages subsidized by the taxpayer, was largely sourced from a recent Congressional Budget Office report (CBO 2019). This report estimates that $19.4 billion of taxpayer money is spent annually on mitigation of and relief from the damages caused by hurricanes.

Additional data sources used for our model simulation are detailed in Appendix Table A5, and we discuss the parameter estimates themselves in Section 4.5.1.

[19] While asymmetry is generally on the interval $(0, 1)$, space is two-dimensional, so in principle it may exceed 1 in some cases. In practice there are fewer than a handful of such cases, and excluding or Winsorizing them at 1 makes no appreciable difference to the results.

[20] NFIP claims were significantly lower before 1978. When we use all the data from 1973-2019, we find an even higher growth rate of around 10-12 percent.

[21] All dollar values used in our analysis, except where otherwise stated, are normalized to 2020 US dollars, using a GDP deflator (Federal Reserve Bank of St. Louis 2020).

3 Empirical findings

This section documents nine stylized facts about the location of housing and its exposure to flood risk and sea level rise, focusing on the area that lies within 10 km of the US Atlantic and Gulf coasts, as discussed in Section 2. This section consists of three parts. First, we discuss three stylized facts that characterize housing and places near the coast in the cross-section; second, we discuss three stylized facts that begin to reveal why housing near the coast follows the patterns that we document; and finally, we show three stylized facts on the development of coastal housing over time.

3.1 Stylized facts on the cross-section of coastal housing

The first stylized fact we document is that **housing unit density peaks near – but not right at – the coast**. To show this, we calculate the number of housing units in each 150 meter distance bin from the coast, assigning the housing units in each census block to the bin where its centroid falls. We then normalize the total number of housing units in each bin by the area of that bin, which we approximate using the cells.[22] The results, in Panel (a) of Figure 1, show that the logarithm of housing unit density peaks around 2.475 km from the coast, and declines asymmetrically, falling more rapidly on the coast side.[23] Specifically, housing density declines steeply as we approach the coast (falling about 0.55 log points over less than 2.5 km) and more slowly on the inland side (falling about 0.85 log points over about 7.5 km). A similar pattern can be seen in Panel (b) of Figure 1, which restricts the analysis to census blocks with housing units, and reports point estimates and 95 percent confidence intervals from estimating the regression:

$$\ln\left(hdensity_i\right) = \beta_{11} + \beta_{12}\mathbf{Bin}_i + \epsilon_{1i}, \tag{1}$$

where $hdensity_i$ is the number of housing units per square km in census block i, $\mathbf{Bin}_i$ is a vector of indicators for 50 meter distance bins from the coast, and ϵ_{1i} is an error term, which is clustered by state here and in all the spatial regressions we report below.[24] The figure peaks around 3km from the coast, and declines on both sides of the peak, again with a steeper decline on the coast side. As we discuss below, the steep decline near the coast side of Panel (b) understates the sparseness of housing density near the coast, since there are more empty blocks in the immediate vicinity of the coast; for that reason, we prefer the specification in Panel (a). We repeat the analysis of the two panels above in Panels (a) and (b) of Appendix Figure A2, this time excluding restricted areas (as discussed in the

[22]Census block centroids provide a good approximation of housing location, since areas with dense housing are partitioned into small blocks. But block centroids are less precise when it comes to measuring area, because areas with sparse housing (or no housing) tend to be in large census blocks. Using cell data to approximate land area in each distance bin is therefore more reliable, since the cells are by construction evenly distributed, and of equal size.

[23]Housing density is quite similar to the peak in other nearby distance bins, and it displays some geographic variation. Specifically, in the US South, housing density peaks closer to the coast, consistent with a higher amenity value of the beach. But in every case we examined, density falls steeply very close to the coast and more gradually further away from it.

[24]This approach follows Donaldson and Hornbeck (2016) and others. We explored using spatial clustering following Bester, Conley, and Hansen et al. (2011), using 1 x 1 degree clusters. This gave slightly smaller standard errors than those we report. Using Conley (1999) standard errors is more technically challenging in our setting, due to the large number of observations.

Data Section).[25] The results are largely unchanged.[26]

Panel (c) of Figure A2 repeats the analysis of panel (a) of Figure 1 but using only block-level data, for area as well as housing units. Here the decline in density near the coast is even steeper.[27] Finally, Panel (d) of Appendix Figure A2 repeats the analysis of panel (a) of Figure 1 using cell-level data on built area instead of housing units. Using the built area data allows us to examine the extent not only of residential housing, but also of commercial and industrial areas, as well as roads and other artificial structures. Here the distribution peaks around 2km from the coast, and once again the decline on either side is asymmetric and similar in magnitude to that in Figure 1.

In interpreting the above-mentioned housing distribution, it is worth noting several additional empirical regularities. Commuting remains an important aspect of cities, and the vast majority of housing units that we consider are primary residences, where people work throughout most of the year.[28] Specifically, only around 1 percent of the housing units in our sample are second homes.[29] Since we do not have fine-grained data on business activity, we assume in the discussion below that peak housing density corresponds to the location of the Central Business District (CBD). We note that given the limitations of our data we cannot explore multiple employment centers within the city, although we discuss this possibility below.[30]

Our second stylized fact is related to the first, namely that **census-designated places close to the coast are asymmetric**.[31] To show this, we use data on places and their CBDs to estimate regressions of the form:

$$asymmetry_j = \beta_{21} + \beta_{22}\mathbf{Bin}_j + \epsilon_{2j}.$$

Here $asymmetry_j$ is the ratio $\frac{|x_R - x_0|}{|x_R - x_L|}$, where the numerator is the distance from each place's furthest point from the coast to its CBD, and the denominator is the distance from each place's furthest point from the coast to its nearest point to the coast; $\mathbf{Bin}_j$ is a vector of indicators for 1 km distance bins from the coast; and ϵ_{2j} is an error term.[32] As Table 1 shows, places whose centroids are within 4km from the coast are asymmetric: the distance from their CBD to their inland edge is roughly double the distance from their CBD to the coast side edge. In contrast, places around 4-10km from the coast are roughly symmetric. An example of this can be seen in Appendix Figure A3, which shows places

[25] As discussed in Section 2, we have no data on the location of all setback areas where construction is more regulated. Their existence may contribute to the steep fall in housing density within around 150 meters from the coast, but is unlikely to drive the overall pattern where housing density peaks around 2-3 km from the coast.

[26] Similarly, controlling for the Density Restriction Index (DRI) has little impact on the patterns shown in Panel (b) of Figure 1 (results available on request).

[27] As discussed above, this may be related to a less precise measurement of area using the block-level data, whose sizes are uneven.

[28] At the same time, the very recent rise in the popularity of telecommuting may suggest that there will be less physical commuting in the future.

[29] While the share of second homes rises in the immediate vicinity of the coast, it is still less than 7 percent even there. We also note that mobile homes make up only around 5 percent of our sample.

[30] The equivalent figures to Panels (a) and (b) for 2010 reveal a very similar picture; the peak of the housing density moves 300 m inland in the 2010 equivalent of Panel (a) but stays constant in the equivalent of Panel (b). In Section 4.5.2 we consider cases where the CBD moves over time.

[31] Census-designated places are numerous and we have data on their CBD location. Many of them are contained within the area of our study, even if metropolitan areas as a whole extend further inland.

[32] Our asymmetry measure, $\frac{|X_R - X_0|}{|X_R - X_L|}$, is for the most part, bounded on the interval [0,1]. There is a small minority of cases where the measure exceeds 1, since in reality X_R, X_0, and X_L are not all on one line. Nevertheless, excluding these few cases does not substantively affect our estimates.

in the Greater Boston area: those close to the coast are asymmetric, while those further away are more symmetric.

The third stylized fact is that **the asymmetry near the coast is more pronounced for large places**. We show this by using the place-level data to estimate regressions of the form:

$$asymmetry_j = \beta_{31} + \beta_{32} size_j + \beta_{33} \ln\left(dist_coast_j\right) + \beta_{34}\left[size_j \cdot \ln\left(dist_coast_j\right)\right] + \epsilon_{3j}, \tag{2}$$

where $size_j$ measures the size of place j, either as $\ln\left(area_j\right)$ where the area is in square kilometers or as the distance $|x_R - x_L|$ in km; $\mathbf{dist_coast}_j$ is the mean distance from each place's blocks to the coast; and ϵ_{3i} is an error term. The estimates in columns (1) and (2) of Table 2, which add the restriction $\beta_{33} = \beta_{34} = 0$, show that on average, larger places (using either of the above measures) are more asymmetric. Columns (3) and (4), which are unrestricted, show that the asymmetry is more pronounced for large places when their CBDs are closer to the coast.

3.2 Stylized facts on mechanisms that shape the coastal housing distribution

Whereas the three stylized facts above tell us how economic activity concentrates near the coast, the next three tell us something about why this is the case. Our fourth stylized fact is that **blocks that are highly prone to sea level rise (SLR) are less densely built, but conditional on SLR-proneness, blocks closer to the coast are more densely built**. To show this, we focus on the share of the area of each census block, which will be under water at high tide if sea level rise (SLR) were 1 foot, or 0.305 meters (we refer to this share as $SLR1ft$). In interpreting $SLR1ft$ we note that it matters not only for a future with higher sea levels, but also for the present: areas with high $SLR1ft$ are more prone to both frequent low-intensity "nuisance flooding" and to flooding from impactful events, such as tropical storms (Dahl et al. 2017). Therefore, all else equal, living in areas with high $SLR1ft$ likely involves costs (a point which we revisit below), and can be viewed as a disamenity. To examine how much these areas are avoided, we split the census blocks into three groups: high-risk medium-risk, and low-risk, as discussed in Section 2. We then repeat the analysis in Panel (a) of Figure 1 separately for each of the three groups of blocks.[33] The results in Figure 2 show that at every distance bin from the coast, low-risk census blocks are about two to three times more densely built than medium-risk blocks, while the medium risk ones are, at most distance bins, several times denser than the high-risk ones. These results are confirmed in robustness checks that we report in Figure A4, where we repeat the analysis in Figure 2 excluding the restricted areas (Panel (a)) and then using the fraction of cell area that is built, based on our gridded data (Panel (b)). When we look within each of the three risk groups, housing density tends to increase as we approach the coast. In other words, conditional on the level of risk, proximity to the coast seems like an amenity.[34]

As our fifth stylized fact shows, however, **as we approach the coast, SLR-proneness rises steeply.** We show this in two ways in Figure 3. As Panel (a) of Figure 3 shows, the fraction of blocks that are low risk is fairly stable at well over 90% in the area 3-10km from the coast. In the three km closest to the coast, however, this share declines, slowly at first and then rapidly, reaching less than 20% as

[33] Except this time we use area data from the blocks.

[34] Since we restrict our analysis to areas within 10 km of the coast, we abstract from differences in productivity across locations. We revisit this point in our discussion of the model.

we get very close to the coast. Over the same range the share of medium risk rises from less than 10% to almost 50%. Meanwhile, the share of high risk, which is less than 10% even a few hundred meters from the coast rises even more sharply to almost 40% very close to the coast. Panel (b) of Figure 3 reports the mean $SLR1ft$ by distance to the coast. This share is lower than 5% in the area 1-10km from the coast, but increases steeply to almost 45% as we get very close to the coast. Appendix Figure A5 shows that these results are again robust to excluding restricted areas. Together, this evidence suggests that the amenity of proximity to the coast, which increases gradually (as we saw in the fourth stylized fact), is offset by a convex disamenity due to flood risk as we near the coast.

To see why this matters, we turn to the sixth stylized fact: **damages from flooding rise steeply as we approach the coast**. To show this, Figure 4 reports point estimates and 95% confidence intervals from the regression:

$$\ln(damage_k) = \beta_{41} + \beta_{42}\mathbf{Bin}_k + \epsilon_{4k}, \tag{3}$$

where $damage_k$ is the total dollar sum of NFIP claims from 1973-2019 (in 2020 USD), normalized by an estimate of the number of housing units from 2014-2018 in census tract k; $\mathbf{Bin}_k$ is a vector of indicators for 150 meter distance bins from the coast; and ϵ_{4k} is an error term.[35] As the figure shows, claims in the distance bin closest to the coast are about 2.5 to 3 log points (or about 12-20 times) higher than in the areas around 4-10km from the coast. While NFIP claims represent only a fraction of the total costs of flooding over the past few decades, this figure indicates that flood costs rise convexly as we approach the coast.[36]

Having characterized the first six stylized facts, we now examine the distribution of prices near the coast. Panel (a) of Figure A6 reports estimates using the same specification as Panel (c) of Figure 1, except plotting the fraction of blocks in each 50-meter distance bin from the coast, for which median house prices are missing. Median prices are missing if blocks are empty or very sparsely populated, so that disclosing moments from the price distribution would reveal information about individual housing units. The figure shows that median house prices are missing for about 30 percent of the census blocks from around 1-10 km from the coast. In the 1 km closest to the coast, however, the fraction missing rises steeply, to almost 67 percent in the blocks closest to the coast. Panel (b) shows that where median house prices are available, they are also fairly flat around 1-10 km from the coast, rising steeply in the 1km closest to the coast. Interpreting this pattern is not straightforward, because of the missing blocks; the coverage within blocks (only 64.1% of housing units in 1990 were owner-occupied); differences in housing characteristics within locations and across them; and the use of the median. Nevertheless, at first glance, the findings we document may seem surprising: blocks near the coast are flood-prone and much sparser than others, and this sparseness is not driven by restricted areas, as Panels (a) and (b) of Figure 1 show; yet where house prices are recorded there, they are high. We explain this apparent puzzle in Section 4.4.3, by noting that while locations in blocks close to the coast are generally flood-prone and therefore in low demand, there may be small

[35] The use of the recent housing units measure mitigates the risk that NFIP claims per housing unit will appear large near the coast because housing expanded there, as we discuss below. The patterns we document are, however, robust to using 1990 housing units in the denominator.

[36] As we discuss in the Data Appendix, NFIP costs cover only a fraction of total damages from flooding. We use it here because it affords spatially disaggregated data at the level of census tracts. Since census tracts are considerably larger than blocks and focus on built areas, there is no point in excluding restricted areas at this level of the analysis.

higher-elevation areas within these blocks, where flooding is much less of a problem, and where prices are high.

Returning to our first six stylized facts, we note that stylized facts 4-6 help explain stylized facts 1-3: conditional on risk, people seem to prefer to live as close as possible to the coast, but as we approach the coast risks increase steeply. This gives rise to the distribution of housing density, which peaks near the coast and declines asymmetrically, falling more steeply on the coast side than on the inland side.

3.3 Stylized facts on changes in coastal housing over time

Whereas the first six stylized facts describe coastal area housing at a point in time, mostly around 1990, the last three stylized facts describe how they changed from 1990-2010. The seventh stylized fact is that despite the risks discussed above, **much construction near the coast in recent decades took place in areas with SLR risk**. This is shown in Table 3. In 1990, areas with medium or high SLR risk accounted for about 12% of the housing units in our area of study, and this fraction increased to around 14% in 2010. This came about because 26% of the net increase in housing units in the area we study from 1990-2010 took place in medium or high-risk blocks.

The eighth stylized fact tells us where the risky new developments took place: **SLR-prone areas were developed in dense census tracts, but not in sparse ones**. Table 4 reports regression estimates using the specification as in (1), except that the dependent variable is the change in housing units in each census block from 1990-2010 and the regressor is $SLR1ft$. These regressions are estimated separately for four groups of census blocks, grouped by the housing density of the census tracts that contain them (where this density excludes the own block's density). As the table shows, in sparse census tracts, the growth in housing units is negatively associated with SLR. But in dense census tracts, new construction is positively associated with SLR proneness (although these estimates are only marginally significant). All this suggests that where there is plenty of space to build, SLR-prone areas are avoided, in line with the evidence discussed above; SLR-prone areas are, however, built on in dense areas, presumably because no other local alternatives exist. We repeat this analysis excluding the restricted areas (Appendix Table A1) and the results are largely unchanged. We then repeat the analysis again using the cell data on built area, where this time "neighborhoods" are larger (1 square km) areas, whose fraction built we calculate excluding the own cell. In this case the estimates in sparse "neighborhoods" are again negative, while in the densest category the estimates are positive but imprecise. Appendix Figure A7 shows four case studies, which illustrate development in the fringes of dense tracts: Revere and Chelsea in Greater Boston, Massachusetts; Jamaica Bay and Rockaway Peninsula in the borough of Queens, New York City, New York; Miami Beach and Miami, Florida; and Clearwater and Largo, Tampa Bay area, Florida.

Finally, our ninth stylized fact is that **in the densest census tracts, new construction focused on medium-risk rather than high-risk areas**. To show this, Table 5 reports estimates from two regressions, which restrict the analysis to the densest group of census tracts discussed above. Column (1) uses specification (1), but with the change in housing units in each census block from 1990-2010 as the dependent variable and an exhaustive set of bins for different percent SLR in each census block. Column (2) is the same, except that the regressors are an indicator $I_{SLR1ft>0}$ (that is, medium or high

risk) and a continuous measure ($SLR1ft$). Both specifications tell a similar story: new construction took place in medium-risk areas more than in low-risk areas, but the highest risk areas were still generally avoided. Appendix Table A3 repeats this analysis, excluding restricted areas, and the results are largely unchanged. Appendix Table A4 repeats the analysis, this time using the cells instead of blocks (as discussed above), and the results are again similar to those in Table 5.

4 Model

4.1 Baseline assumptions

In this section, we present a model on coastal development that helps to reconcile the nine stylized facts that we documented above. The model is in discrete time, and periods are denoted by t. Spatially, we extend the monocentric city model (Alonso 1964, Mills 1967, Muth 1969), by placing it in the context of a coast, proximity to which offers both benefits and costs.[37] The key geographic locations of the city are the CBD, denoted by x_0; the coast-side and inland edges of the city, denoted by x_{Lt} and x_{Rt}; and the coast itself, whose initial location is normalized to 0.[38] Initially, the CBD location is chosen by a historical city founder, and then the city persists for T periods (decades). In each period, developers choose where to build, taking into account the preferences of residents, who choose where to locate.

The city founder is assumed to be myopic, and chooses a location x to maximize their locational utility

$$U^F(x) = -\theta_1 x - \theta_2 x^{-\sigma}. \tag{4}$$

We assume that $\theta_1 > 0$,reflecting our observation that proximity to the coast has an amenity value (air, views, bathing), which we assume is linear.[39] We also assume $\theta_2 > 0$ and $\sigma > 0$, reflecting a convex disamenity (higher risk of flooding).[40] As we discuss further below, housing density increases and then decreases as we head inland from the coast, which is consistent with a demand-based explanation. We note that for simplicity, the model is deterministic, and the risk of flooding is captured by the last term of the utility function. We assume that the founder's chosen location becomes the city's CBD, x_0.[41]

There is a continuum $[0, \overline{x}]$ of competitive and forward-looking developers, each of whom owns

[37]The monocentric city model is commonly used in urban economics. While it abstracts for the multiplicity of employment locations within cities, it has useful comparative statics and strong empirical support (e.g., Duranton and Puga 2014). Moreover, when adapted as we describe below, it neatly and parsimoniously explains our findings. Nevertheless, the existence of multiple employment centers within cities may mitigate some of the distortions that we document.

[38]Later, when we explore SLR, we relax this assumption by allowing x_{ct} to grow over time. All our analysis focuses on inland locations ($x > x_{ct}$).

[39]It is possible that across wider areas than the coastal band that we study, the amenity component of the utility function also declines convexly in distance to the coast. But the key assumption is that it is less convex than the disamenity term, so for simplicity we assume a linear amenity term in the vicinity of the coast, which is the area we focus on. As we discuss below, this assumption is motivated by the convex increase in flood risk as we near the coast.

[40]We focus on the negative consequences of flooding, but the disamenity modelled here may include indirect effects of flooding on the soil's suitability for housing, as well as the effects of wind gusts.

[41]The location of many cities on the US Atlantic and Gulf coasts was established more than a century ago, so for simplicity we assume that their location choices were myopic. We also ignore any productivity component in the city founder's locational choice, although adding this would not make much difference to the model overall.

a plot of land of measure 1 in location x.[42] Each period, each developer can allocate their plot to housing, which yields a period price of $p_t(x)$, or to agriculture, which has a period price p_A.[43] The developers' time preference is captured by $\delta \in (0,1)$, and in each period every developer maximizes their present-discounted stream of future prices.

Finally, there is a continuum of perfectly mobile residents. In every period $t = 1,..,T$, each resident may live in the city or outside it.[44] If they live in the city, they inelastically supply one unit of labor, receive a wage, and spend their income on consumption and housing, in which case their utility is:

$$U(c_t, h_t, x) = c_t^{\alpha} h_t^{1-\alpha} - \theta_1 x - \theta_2 x^{-\sigma} \tag{5}$$

where c_t denotes private consumption goods and h_t denotes housing in period t and $\alpha \in (0,1)$ is the consumption share of income.[45] We assume that residents' preferences satisfy standard assumptions $(U_c > 0, U_h > 0, U_{cc} < 0, U_{hh} < 0)$. The residents' locational preferences are the same as those of the city founder. The budget constraint of each resident in period t is:

$$p_t(x)h_t + c_t = w_t - |x - x_0| \tag{6}$$

where the price of consumption is normalized to 1; w_t is wage, and $|x - x_0|$ reflects the time cost of commuting. Each resident also has an outside option of living outside the city, with utility $\bar{U} > 0$. We initially consider a city whose attractiveness to residents and developers increases (at least weakly) relative to the outside option, or in other words that w_t increases (weakly) in t.

We solve the model as a Nash equilibrium, where developers take into account the expected maximization of other developers and of the residents.

4.2 Equilibrium

Here we summarize the equilibrium conditions of the model, a visual illustration of which is discussed in Section 4.5.2.

City founder: maximization of the city founder's decision implies, using the first-order condition, that

$$x_0 = \left(\frac{\sigma\theta_2}{\theta_1}\right)^{\frac{1}{\sigma+1}}. \tag{7}$$

Residents decide where to live and the share of consumption goods and housing in their consumption bundle. In equilibrium they are indifferent between all city locations, including the city endpoints, and their outside option $\bar{U}$. Residents' indifference between locations then determines the price function, $p_t(x)$, for each period.

Developers decide which locations should be part of the city, taking into account the present discounted stream of future prices. Since the baseline setup of the model is static, developers will

[42] We assume that $\overline{x} > 0$ is sufficiently high not to constrain the coast side development.

[43] We follow the literature by labelling non-housing use as agriculture, although in practice there may be other alternative uses of land. In the baseline model we assume that agricultural prices are fixed across time and space, and that there is no cost of converting land across uses. We relax the latter assumption in an extension in Section 4.4.2. One caveat that we do not consider is salinity, which may affect some forms of agriculture, but not others (e.g., fishing).

[44] We assume that there are more than enough residents to populate the city at each point in time.

[45] There are no savings in the model.

build in all locations such that

$$p_t(x) \geq p_A. \tag{8}$$

Since (as we show below) prices decrease monotonically as we move away from the CBD, the boundaries of the city x_{Lt} and x_{Rt} are pinned down by the equations:

$$p_t(x_{Rt}) = p_t(x_{Lt}) = p_A. \tag{9}$$

Note that because of the assumptions discussed above, developers can repurpose land costlessly in every period, severing any dynamic link between periods. Below we discuss an extension where housing construction is costly and irreversible, which introduces dynamic considerations.

4.3 Relating the model to the stylized facts

We now discuss how the model rationalizes the nine stylized facts that we observe. We begin with Stylized fact 4, that flood-prone areas are more sparsely built, but holding flood risk constant proximity to the coast is seen as an amenity. This motivates our assumption that $\theta_1 > 0$. At the same time, Stylized facts 5 and 6 show that the cost of flooding rises convexly with proximity to the coast, motivating our assumptions that $\theta_2 > 0$ and $\sigma > 0$.

Next, we turn to Stylized fact 1, showing that housing density is single-peaked and decreases on both sides of the CBD.

Proposition 1 *Define housing density* $dens_t(x) \equiv \frac{1}{h_t(x)}$, *we get the following result:*

$$\textit{For each period } t = 1, ..., T\textit{: if } x < x_0 \textit{ then } \frac{\partial \ln(dens_t(x))}{\partial x} > 0\textit{; if } x > x_0 \textit{ then } \frac{\partial \ln(dens_t(x))}{\partial x} < 0. \tag{10}$$

Proof. See appendix. ■

At this point we revisit the house price profile shown in Figure A6, which was quite flat from 1-10 km from the coast. This pattern is largely consistent with our model, as long as commuting costs account for a small share of income, which is what we find in Section 4.5.1 below. In the model, each resident spends a share $1 - \alpha$ of their income on housing, and this corresponds to the price of their "housing unit". Those near the CBD get a small amount of high-value land, while those near the city edge get a larger amount of cheaper land. And indeed, as our first stylized fact suggests, locations further from the CBD have fewer housing units per square km, or in other words more area per housing unit, consistent with the model. As discussed above, however, this still leaves the matter of explaining the higher prices on the sparse areas from 0-1 km from the coast. We revisit this in the extension in Section 4.4.3.

Another characteristic of places near the coast is reflected in Stylized fact 2, which shows that places near the coast are asymmetric: their CBD is closer to their coast-side edge than to their inland edge. We show that this is the case in the model.

Proposition 2 *The city develops asymmetrically around the CBD:* $|x_{Rt} - x_0| > |x_0 - x_{Lt}|$.

Proof. See appendix. ■

Our final static empirical result, Stylized fact 3, is that the asymmetry near the coast is more pronounced for large cities. And in the model, the city's asymmetry goes away if it is very small.

Proposition 3 $\lim_{x_{Lt} \to x_0, x_0 \leftarrow x_{Rt}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$

Proof. See appendix. ■

Turning to our empirical findings on changes over time, we consider what happens when the city expands over time. To do so, we focus on the case where w_t strictly increases over time, making the city more attractive relative to the outside options for both the residents (living elsewhere) and the developers (collecting the agricultural price). Stylized fact 7 is that much of the new construction takes place in flood-prone areas. In the model, if the city expands, it will do so on both edges of the city, and the expansion on the coast side will take place in areas that are at high risk of flooding. This leads to a different notion of the geographic constraints of the city than Saiz (2010) and Harari (2020), where the city expands until it reaches "hard" edges. In contrast, our model allows for "soft" edges, which residents and developers would like to avoid, but which are developed regardless as the city expands.

Stylized fact 8 shows that expansion into flood-prone areas takes place in dense neighborhoods, but not in sparse ones. In the model, if a city is small (and therefore sparse), it will expand into areas that are further from the coast, and therefore safer; but large (and dense) cities will expand into flood-prone areas.

Finally, Stylized fact 9 is that while dense neighborhoods expand into flood-prone areas, they avoid the highest-risk areas. In the model, prices decline monotonically from the CBD (see above), and therefore development will take place at the city's edge, avoiding the most flood-prone areas right by the coast.

4.4 Extensions

Here we consider extensions of the baseline model, including: sea level rise; irreversible housing construction; a limited number of elevated locations near the coast; and government subsidies to offset the disamenity of proximity to the coast.

4.4.1 Sea level rise

We model sea level rise as a change in the location of the coast, x_{ct}. In this case, each resident's utility is

$$U(c_t, h_t, x) = c_t^{\alpha} h_t^{1-\alpha} - \theta_1 (x - x_{ct}) - \theta_2 (x - x_{ct})^{-\sigma}. \tag{11}$$

For simplicity, we focus on the case where sea levels rise linearly, in a city which slopes linearly from the coast to the CBD, although the model can be adapted to nonlinear SLR. We assume here that once a location submerges due to SLR it becomes uninhabitable. Moreover, SLR changes the relative attractiveness of locations that are not submerged. Locations on the coast-side of the CBD become less attractive, because in this range the costs of the disamenity rise faster than the benefits from the amenity, as is evident from the city founder's problem (above). But locations further inland may

actually benefit from SLR, at least for a while, since the amenity value of being closer to the coast may dominate the increased disamenity from flooding. Because we assume that the CBD is fixed, this may lead to even more asymmetric coastal cities, and further raise average commuting costs. The problem of "misshapen cities" (Harari 2020) may therefore be further aggravated by SLR.

4.4.2 Irreversible housing construction

The analysis so far is simplified by the assumption that developers can costlessly switch land use from agriculture to housing and then back again, so each period is independent of the rest. To relax this assumption, we add a time cost of switching; we refer to this version of the model as "dynamic", to distinguish it from the baseline static model. For tractability, we assume that conversion from agriculture to housing costs one period's price, and is irreversible (so conversion from housing to agriculture is infinitely costly). In this case, the developers' problem can be solved by comparing the present discounted value of prices from all time periods until T (or with SLR, until the plot submerges). So the developer's problem in city location x and period $t = 1, ..., T$ is:

$$Max \left\{ \sum_{s=t}^{T} \delta^s p_A, \underset{\widehat{s}=t,...,T}{Max} \left[\sum_{t \leq s < \widehat{s}} \delta^s p_A + \sum_{T \geq s > \widehat{s}} \delta^s p_s(x) \right] \right\}. \tag{12}$$

This condition replaces expression (9). The introduction of these costs dampens the incentives to expand the city, both due to the opportunity cost of receiving the agricultural price for a period (instead of developing) and because of the option value of developing later.

We note that the modelling assumption above relates only to extensive margin changes (whether land has housing or not), and not to intensive margin ones (how many units of housing it has). Adding frictions on the number of housing units is more analytically involved (see for example Henderson, Regan, and Venables 2021). We conjecture that barriers to increased housing density as the city expands exacerbate the extensive-margin expansion towards the coast and the distortion caused by SLR.

4.4.3 Limited number of elevated coastal locations

The baseline model assumes that flood-proneness increases monotonically as we approach the coast. But in reality, there are, even in LECZ, a limited number of elevated locations that are not prone to flooding. To model this, we assume a finite set of elevated locations $\{x_{E1}, ..., x_{EN,}\}$, where for the residents, $\theta_2 = 0$.[46] In elevated locations close to the coast, developers may build expensive housing, even outside the city boundaries, as long as commuting costs are not too high. Since the elevated locations are a finite set, they do not affect the overall distribution of housing density, nor the gradual expansion of the city into riskier areas.[47] But as a result, housing prices near the coast (measured in the locations that have housing) may be higher than elsewhere.[48]

[46]Since each of those locations is isolated, we assume that they are not suitable for the city's founder, whose choice remains unchanged. If they are included in the city founder's choice set, the CBD may be even closer to the coast.

[47]The elevated areas are generally built on early in the city's development, assuming that commuting costs are not too high.

[48]While this is outside the scope of our model, being in an elevated location with views on empty surrounding areas may even enhance the coastal amenity. Such locations may also appeal to high-income individuals, although again our

4.4.4 Government subsidies

As we discuss in the data section, there is evidence that the government subsidizes coastal development, for example through flood relief. We explored allocating government subsidy in proportion to the losses we document in Stylized Fact 6 (in proportion to the fitted values from equation (13) below). The typical reason for the subsidy is that it mitigates the effects of catastrophic shocks, such as those from major storms, whose main impact is on residents living close to the coast. Such large-scale shocks may be difficult to credibly insure against, and as a result the government often steps in. In the context of our model, however, such subsidies increase prices and construction near the coast. This may be particularly problematic when subsidies interact with SLR. Sufficiently large subsidies may even shift peak prices closer to the coast than the CBD location (without subsidy), which in principle may help to explain why house prices peak very close to the coast. In practice, however, the subsidies that we explored were insufficiently large to generate this effect. Our preferred explanation is therefore the one discussed above in Section 4.4.3, which can also account for the missing values, as discussed in Section 3.

4.5 Simulations

4.5.1 Parameter estimation

In this section we study a synthetic low-elevation coastal city, and explore its evolution under different assumptions and scenarios. Coastal cities vary, of course, in size and location, depending on local conditions and history. Our model therefore illustrates the conditions that may prevail in a city whose characteristics are similar to those we find when averaging across distances from the coast across the area we study.

We begin by explaining our choice of parameters for the simulated model, where we study the city's evolution in decennial intervals. Since our parameter estimates mostly rely on 1990 data, we begin our simulation in the year 1990 ($t = 1$). We restrict ourselves to cases where the CBD is not yet submerged by period T even in the case with faster SLR (see below). In practice, we chose the year 2200 ($t = 22$) as the final period, and the simulation estimates that we report are for 1990-2100. This allows for a long horizon (one century) even when decisions for 2100 are made. Spatially, the simulation divides the area from 0-10 km from the 1990 coastline into 10,000 equal one-meter bins.[49]

As we discuss in detail in Appendix Table A5, we use various approaches to obtain our parameters. First, we take from the literature parameter values that have been carefully estimated in relevant settings, including the consumption share, α (following Davis and Ortalo-Magne 2011); the elevation of the CBD, for which we use the Population-weighted median elevation of Miami (Hallegatte et al. 2013); and travel speed, which affects $I(x)$ (where we follow Couture et al. 2018).

Second, we calculate some of the parameters directly from the data sources, including the CBD location (x_0), the distance from the coast to the coast-side edge of the city (x_L), the wage (w), and the discount rate (δ), as discussed in Appendix Table A5.[50]

model does not account for heterogeneity in income.

[49] We verify that income is positive throughout the area we simulate. We also restrict the inland edge of the city, x_{Rt}, so that it does not extend beyond the area of our simulation; in practice this does not happen from 1990-2100.

[50] The CBD location in the model is closer to the coast (around 1.3 km) than the peak of the empirical density distribution

We also estimate some of the parameters. To obtain an estimate for $I(x)$ we begin with earnings (which equal the wage in our model), and then subtract the earnings loss from commuting, $|x - x_0|$. To get this last term we convert 1km of distance from the CBD into 500km of annual commute (assuming 250 working days per year with return commutes); then assuming the commuting speed, we convert each 1km distance into a measure of annual hours of commute; we then divide this term by mean annual hours worked to approximate the fraction of earnings lost for each 1km from the CBD. Finally, we convert this into decennial earnings by multiplying by 10.

Estimating utility parameters, such as σ, is more challenging. We cannot measure utility directly, so instead we use our data on economic losses, as discussed in Stylized Fact 6, and assume that the utility loss is roughly proportional to these losses. Specifically, we use the binned census tract data from Figure 4 and the nonlinear regression:

$$damage_k = \psi x_k^{-\sigma} + \varepsilon_{k5}, \tag{13}$$

to estimate σ.

Next, we use (17) and the data from Panel (a) of Figure 1 to estimate θ_1, θ_2, and $\bar{U}$ using the regression:

$$\ln\left(dens_t\left(x_k\right)\right) = \frac{\alpha}{1-\alpha}\ln\left(\alpha\right) + \frac{\alpha}{1-\alpha}\ln\left(I_t\left(x_k\right)\right) - \frac{1}{1-\alpha}\ln\left(\bar{U} + \theta_1 x_k + \theta_2 x_k^{-\sigma}\right) + \varepsilon_{k6}. \tag{14}$$

We report the regression estimates in Appendix Table A5, and in Appendix Figure A8 we compare the fit of our estimates to the data. As the figure shows, the parsimonious specification captures well the single-peaked and asymmetric distribution of housing density.

To fix the boundaries of our simulated city, we also need to estimate the agricultural price p_A, which represents the opportunity cost of urban development. Following the static model's assumptions, we equate it to the housing price at x_L in 1990. For the dynamic version of the model (with irreversible housing construction), we invert out the agricultural price as the constant price, which leaves developers indifferent between converting the land from agricultural to residential use and not converting at the left edge of the city in the first period (1990) of the simulation, under no expectations of future SLR or subsidies.[51] In other words, we calculate $\sum_{t\leq s<\widehat{s}}\delta^s p_A + \sum_{T\geq s>\widehat{s}}\delta^s p_s(x)$ for all $1 \leq \widehat{s} \leq T$, and adjust the value of p_A until $Max\left[\sum_{t\leq s<\widehat{s}}\delta^s p_A + \sum_{T\geq s>\widehat{s}}\delta^s p_s(x)\right] = \sum_{T\geq s>1}\delta^s p_s(x)$ and $\sum_{T\geq s>1}\delta^s p_s(x) \geq \sum_{T\geq s\geq 1}\delta^s p_A$.[52]

In general, the estimated agricultural prices in the static and dynamic settings are different. We allow this flexibility to ensure that the urban expansion on the coastal side started from the same original point in 1990.

Wage growth in the model serves as a reduced-form mechanism for driving urban growth, since we hold the outside option for residents ($\bar{U}$) and developers (p_A) constant. In the simulations, we consider the case of 1 percent wage growth per decade, which leads to a decennial growth of roughly

(around 2.5 km). This may be in part because the density distribution peak is not very steep.

[51] We assume that urban development decisions before 1990 had paid relatively little attention to future SLR risks.

[52] If we had carried over the agricultural price from the static version to the dynamic version when wage growth is positive, the city would have expanded rapidly, as developers would have weighed the growing price of housing against the price of agricultural land, which we assume is constant.

10 percent in housing units, similar to what we find in Table 3. We also consider an alternative case where the wage does not grow.

For simulations that involve SLR, we consider three alternatives. First, a case without SLR. Second, our baseline SLR case of a 63.5 cm vertical increase in sea level from 1990-2100, which is halfway between the two main IPCC scenarios of 43 and 84 cm (Pörtner et al. 2019). In this case, the 5.77-cm vertical increment of sea level every decade, translates into a horizontal movement of coastline at the rate of 34.9 meters per decade.[53] Third, we consider a faster SLR scenario, which corresponds either to the case where the area of our study experiences local SLR that is faster than the global average (Dahl et al. 2017), or to a city whose CBD elevation is 2/3 of the one in our baseline scenario.

Finally, calculating the subsidy per housing unit involves the following steps. We begin with the annual public expenditures on tropical storm damage, which are estimated at around $19.4 billion per year as of 2019 (CBO 2019).[54] We assume that 36 percent of these costs are allocated to the area of our study, as in the case of the NFIP share (see Section 2). Next, we adjust these to the 1990, using the roughly 3.8 percent long-run annual growth of NFIP expenditure.[55] We then allocate the costs by distance to the coast, using the estimates from regression (13).

4.5.2 Simulation estimates

We summarize some aspects of the simulated model for 1990 in Appendix Figure A9. This figure shows the linear decline in coastal amenity and the convex decline in flooding disamenity as we move away from the coast, with the marginal effect of both equating at the CBD. The figure also shows commuting costs rising in distance to the CBD. Finally, the bottom panel shows the housing density and the city boundaries.

In Figure 5 we report some of our findings from the simulations, focusing on the extensive margins of city expansion, corresponding to our initial question: where do people build on LECZ?

We initially consider the baseline simulation, with no sea level rise. This scenario, which is described in panel (a) of Figure 5, illustrates some of the stylized facts that we discuss above: in 1990 the city is relatively small, and hence only slightly asymmetric around its CBD. As the city becomes bigger, it also becomes more asymmetric. The city expands on both sides, and the expansion on the coast-side is towards increasingly flood-prone areas, taking in the least-bad locations that are still unbuilt. Panel (b) shows estimates from the dynamic equivalent of this scenario, and the results are largely unchanged.

Next, we add baseline sea level rise in panel (c) of Figure 5, using the midway point between the two main scenarios in Pörtner (2019), for a city whose elevation is similar to Miami's, with a CBD 2 meters above sea level. Now we can see that the city's advance on the coast side is slower, even without dynamic considerations, because locations close to the coast become increasingly flood-prone even before they are submerged. Nevertheless, the city expands towards the coast, even taking in locations that are later submerged. In this case switching land use and even an abandonment of part

[53] We assume the elevation of CBD to be 2 meters. The horizontal movement is calculated as 1210/2*0.0577 = 34.9, where 1210m is the model-implied distance $x_0 = \left(\frac{\sigma\theta_2}{\theta_1}\right)^{\frac{1}{\sigma+1}}$ between CBD and the left boundary of the city assuming no SLR.

[54] NOAA (2021a) shows that tropical storms are, by some margin, the largest single contributor to damages from large-scale natural disasters in the US.

[55] These are based on NFIP data from 1978-2019.

of the city is (by assumption) not directly costly, but SLR still distorts the city, by making it more asymmetric. Another aspect of this distortion is the more rapid expansion of the city on the inland side, where the marginal benefit of the approaching coast is, at least for a while, positive, as can be seen from the city founder's problem. The combined effect of the slower expansion on the coast side and the more rapid expansion on the inland side further distorts the city's shape and lengthens typical commutes.[56] Finally, we observe a small area of the city where the number of housing units declines by more than 10 percent relative to the peak density across all previous periods.[57] In the model this does not cause problems to anyone other than the developer. But in reality, neighborhoods with declining demand may lead to a host of economic and social problems, although these lie outside the scope of our model.

In the dynamic version that corresponds to this scenario (Panel (d)), the city expands less on the coast side, as the cost of development deters some of the expansion in the face of SLR. On the inland side, however, the expansion is very similar to the static model with moderate SLR.

In Panel (e) we consider the case of rapid SLR – 1.5 times the speed of the baseline SLR. This faster speed may represent one of the three factors: faster local SLR on the US Atlantic and Gulf coast than the global mean, as discussed above; a city with a lower elevation CBD, of about 1.33 meters; or moderately faster global SLR than currently anticipated. Even in this case the city expands towards the coast as the coast moves closer to the city, resulting in higher costs of flooding. The coast-side expansion is, however, slower in this case, and stops altogether in the dynamic version of this scenario (Panel f). In this case, especially in the dynamic model where urban land cannot be reconverted into agricultural land, there are even more declining neighborhoods. At the same time, in both Panels (e) and (f) the city expands even more on the inland side, because the faster-moving coastline brings the inland locations closer to the coast, increasing their amenity value (net of flood costs), as least for a while. This results in a further distortion of the city's shape, and even longer commutes. Finally, the fast SLR scenario highlights the problem that the city ultimately faces: to survive SLR in the long run, it needs to move its CBD, which could be very costly, and again lies outside the scope of the model.

In Appendix Figure A10 we consider additional scenarios. The first two panels show a city without rising wages, but with rapid SLR. Here the city shrinks due to SLR, with the coast-side contraction more sizeable than the inland-side expansion. In the static case (Panel (a)), urban land is converted to agriculture on the coast-side as the coast approaches, while in the dynamic case (Panel (b)) those neighborhoods go into decline.

The bottom two panels of Figure A9 consider a government subsidy to offset some of the flood costs with baseline SLR. In the static case (Panel (c)), this leads to faster expansion on the coast side in the face of SLR, and even in the dynamic case (d), we see rapid expansion towards the coast, in contrast to the case without subsidy, as discussed above.

In addition to the scenarios above, we also considered the case where the CBD location is not chosen just once by the city founder, but is selected in each period to trade off the amenity and

[56] The existence of multiple employment centers within the city, which we do not model, may mitigate some of the distortion caused by longer commutes.

[57] Using the 10 percent threshold allows us to visualize economically declining locations. The demand-driven declines, caused by rising flood risk, represent a larger fall in period prices than 10 percent, and with SLR these locations eventually become uninhabitable.

disamenity of coastal proximity (results available on request). In this case, the CBD and housing gradually shift inland as sea levels rise. This allows cities to overcome the problem of SLR in the static case, or mitigate them considerably in the dynamic case, where left-behind declining neighborhoods are now also further from the CBD, and new extensive-margin developments are costly. This version of the model, however, does not account for the costs of building new buildings to replace existing ones, and the coordination costs involved in moving a CBD.[58]

In summary, these simulations highlight four problems of low-elevation cities. First, the problem of flooding worsens over time, either because cities expand towards the coast, or because of SLR, or because both happen simultaneously. This development threatens to increase flooding costs for both residents and taxpayers. The costs could be exacerbated by government subsidies to flood-prone areas. Second, even if LECZ cities grow on aggregate, some neighborhoods decline, as increased flood risk causes prices and population to decline. This problem is worse for cities that are economically stagnant. Third, SLR further distorts the shape of LECZ cities, significantly lengthening the time costs of commuting to work. Finally, LECZ cities face a potential crisis if their CBD comes under threat of being permanently submerged, so it is important to consider how moveable this center of economic activity is, and at what cost.

4.5.3 Policy implications

Governments could enact various policies to mitigate the problems discussed above, especially the rising exposure of taxpayers. First, governments could consider taxing new developments in flood-prone areas, if there are viable alternative uses to the land (such as agriculture in our model), which are not taxed. The difference between the dynamic scenarios (where extensive margin adjustments are costly) and the static scenarios indicate that with SLR, raising the costs of extensive margin development restricts it to some extent.[59] The limiting case is an outright ban on extensive margin developments, although enacting and enforcing such a ban might be difficult.

Second, governments could offer the subsidy only to existing housing. One such policy is the UK government's Flood Re, which provides subsidized flood insurance only to "grandfathered" housing, built before 2009.[60] Comparing the outcomes in Panels (c) and (d) of Figure 5 (without a subsidy) with Panels (c) and (d) of Figure A9 (with a subsidy), we see that the subsidy led to more coast-side expansion, so withdrawing it could help limit government exposure.

Moving beyond our current model, governments could attach further conditions to their subsidy. These conditions could include stricter building standards, such as construction on stilts imposed by the US Federal government when compensating the victims of Hurricane Sandy.[61] Or governments could restrict the number of times a given property is bailed out, or offer other incentives to move

[58] Alternatively, we also explored the case where the CBD is immobile but located at the empirical peak of the housing density instead of the model-predicted location. This leads at least in the short run, to cities skewed around their CBD in the opposite way to what we observe in the data, since locational fundamentals and CBD location attract population to two different locations. We therefore do not consider this case to be of much empirical relevance.

[59] The dynamic simulations also restrict switching land use from housing to agriculture, but the argument still holds even if that were not the case (results available on request).

[60] See https://www.floodre.co.uk/can-flood-re-help-me/eligibility-criteria/

[61] e.g., https://www.ft.com/content/f95aa4e2-b3e6-11e7-aa26-bb002965bce8

instead of rebuilding, as Canada has recently done.[62] With SLR proceeding at pace, the costs to taxpayers of fixing neighborhoods or even cities may at some point become prohibitive. An example of how far things have deteriorated in another part of the world can be seen in Indonesia, whose government is investing heavily in moving its capital from flood-prone Jakarta.[63]

Ultimately, of course, slowing down climate change and SLR could also reduce the costs, especially those associated with large-scale urban moves. This remains a central policy challenge.

5 Conclusions

This paper contributes to our understanding of housing construction in LECZ. We begin by documenting nine stylized facts. These reveal the distribution of housing stock density, which peaks near the coast. They also show the asymmetry of the housing density distribution and of places near the coast, an asymmetry which is particularly pronounced for large places. We relate these findings to the tradeoff between the amenity value of proximity to the coast, conditional on flood-proneness, and the convex increase in flood risk and damages from floods close to the coast. We show how new construction in recent decades avoided flood-prone areas in sparse locations, but in dense locations new construction took place on the least-bad flood-prone areas.

We then develop a simple model of a monocentric city, which combines the amenity value of proximity to the coast with a convex cost of building very close to the coast. This model allows us to explain the patterns that we see, and answer questions such as: why does population concentrate near (but not right at) the coast? Why are coastal places asymmetric? Why does this asymmetry vary by place size? And why does construction take place in flood-prone urban fringes?

Finally, we extend our model and use it to study how SLR may reshape cities. This allows us to explore the evolution of future flood costs, as cities expand towards the coast even as the coast moves towards them; the economic decline of areas even within expanding cities, as SLR reduces demand for locations that become increasingly flood-prone; the lengthening of commutes, as cities' asymmetry around their historical CBDs grows; and the threats to coastal cities that depend on low-elevation CBDs.

By combining empirical evidence with a simple and highly adaptable model, our paper offers a path for researchers and policy makers to consider the implications of a range of interventions in low-elevation coastal cities, in an era when climate change poses increasingly important challenges.

References

Alonso, William. "Location and land use. Toward a general theory of land rent". Publication of the Joint Center for Urban Studies. Cambridge, MA: Harvard University Press (1964).

Bester, C. Alan, Timothy G. Conley and Christian B. Hansen. "Inference with Dependent Data Using Cluster Covariance Estimators". Journal of Econometrics, 165, no.2 (2011): pp. 137-151.

[62] https://www.nytimes.com/2019/09/10/climate/canada-flood-homes-buyout.html
[63] e.g. https://www.ft.com/content/5a463614-c7e4-11e9-af46-b09e8bfe60c0

Balboni, Clare. "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities" Revise & Resubmit, American Economic Review (2020).

Berardelli, Jeff. "How climate change is making hurricanes more dangerous". Yale Climate Connections. New Haven: CT (2019).
Accessed July 2020 at https://yaleclimateconnections.org/2019/07/how-climate-change-is-making-hurricanes-more-dangerous/

Brakenridge, Robert. "Global Active Archive of Large Flood Events". Dartmouth Flood Observatory. Boulder, CO: University of Colorado.
Accessed May 2020 at http://floodobservatory.colorado.edu/Archives/index.html

Burningham, Kate, Jane Fielding and Diana Thrush. "'It'll never happen to me': understanding public awareness of local flood risk". Disasters, 32, no. 2 (2018): pp. 216-238.

Coastal Construction Control Line Program (2021). Accessed March 2021 at
https://ca.dep.state.fl.us/mapdirect/?webmap=a8c9e92fbad5446d987a8dd4ee5dc5cc

Conley, Timothy G. "GMM estimation with cross sectional dependence". Journal of Econometrics, 92, no.1 (1999): pp. 1-45.

Couture, Victor, Gilles Duranton and Matthew A. Turner. "Speed". The Review of Economics and Statistics. 100, no. 4 (2018), pp. 725-739.

Dahl, Kristina A., Melanie F. Fitzpatrick and Erika Spanger-Siegfried. "Sea Level Rise Drives Increased Tidal Flooding Frequency at Tide Gauges along the U.S. East and Gulf Coasts: Projections for 2030 and 2045". PloS One, 12, no. 2 (2017).

Database of Global Administrative Boundaries (GADM). "USA Country Outline (Version 3.6)". Davis, CA: GADM (2018). Accessed September 2020 at
https://gadm.org/download_country_v3.html

Davis, Morris A. and François Ortalo-Magné. "Household Expenditures, Wages, Rents". Review of Economic Dynamics. 14, no. 2 (2011): pp. 248-261.

Deryugina, Tatyana. "The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." American Economic Journal: Economic Policy (2017), 9(3): 168–198

Desmet, Klaus, Robert E. Kopp, Scott A. Kulp, Dávid K. Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg and Benjamin H. Strauss. "Evaluating the Economic Cost of Coastal Flooding". American Economic Journal: Marcoeconomics, forthcoming (2021).

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A "Market Access" Approach", Quarterly Journal of Economics, Volume 131, Issue 2, May 2016, Pages 799–858

Duranton, Gilles and Diego Puga, Chapter 5 - The Growth of Cities, in Handbook of Economic Growth (Philippe Aghion and Steven N. Durlauf, eds.), Elsevier, Volume 2, 2014, Pages 781-853,

Federal Emergency Management Agency (FEMA). "National Flood Insurance Program Redacted Claims – v1". Washington DC: FEMA (2021). Accessed July 2020 at
https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims-v1

Federal Reserve Bank of St. Louis. "Gross Domestic Product: Implicit Price Deflator (GDPDEF)". St. Louis, MO: Federal Reserve Bank of St. Louis (2020). Accessed January 2021 at
https://fred.stlouisfed.org/series/GDPDEF

Gibson, Matthew and Jamie T. Mullins. “Climate Risk and Beliefs in New York Floodplains”. Journal of the Association of Environmental and Resource Economists, 7, no. 6 (2020): pp. 1069-1111.

Glaeser, Edward L., Jed Kolko and Albert Saiz. “Consumer city”. Journal of Economic Geography, Oxford University Press, 1, no. 1 (2001): pp. 27-50.

Hallegatte, Stephane, Colin Green, Robert J. Nicholls and Jan Corfee-Morlot. “Future flood losses in major coastal cities,” Nature Climate Change. 3, no. 9 (2013): pp. 802-806.

Harari, Mariaflavia. “Cities in Bad Shape: Urban Geometry in India”. American Economic Review, 110, no. 8 (2020): pp. 2377-2421.

Henderson, J Vernon, Tanner Regan and Anthony J. Venable. “Building the City: from Slums to a Modern Metropolis”. The Review of Economic Studies, forthcoming (2020).

Hino, M. and Marshall Burke. “Does information about climate risk affect property values?” NBER Working Paper, no. 26807 (2020).

Hornbeck, Richard and Daniel Keniston. “Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872”. American Economic Review, American Economic Association, 107, no. 6 (2017): pp. 1365-1398.

Keys, Benjamin J. and Philip Mulder. “Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise”. NBER Working Paper, no. 27930 (2020).

Kocornik-Mina, Adriana, Thomas McDermott, Guy Michaels and Ferdinand Rauch. "Flooded Cities". American Economic Journal: Applied Economics, 12, no. 2 (2020): pp. 35-66.

Kydland, Finn E. and Edward C. Prescott. “Rules rather than discretion: The inconsistency of optimal plans”. The Journal of Political Economy, 85, no. 3 (1977): pp.473-491.

Magontier, Pierre, Albert Solé-Ollé and Elisabet Viladecans-Marsal. “The political economy of coastal destruction”. Paper presented at the Centre for Economic Performance – Urban and Regional Economics Seminars (February 2019).

Manson, Steven, Jonathan Schroeder, David Van Riper and Steven Ruggles. “IPUMS National Historical Geographic Information System: Version 14.0” [Database]. Minneapolis, MN: IPUMS (2019). Accessed February 2021 at: http://doi.org/10.18128/D050.V14.0

Marcy, Doug, Nate Herold, Nate Herold, Kirk Waters, William Brooks, Brian Hadley, Matt Pendleton, Keil Schmid, Mike Sutherland, Kyle Dragonov, John McCombs and Sean Ryan. “New Mapping Tool and Techniques For Visualizing Sea Level Rise And Coastal Flooding Impacts”. Charleston, SC: NOAA Coastal Services Center (2011). Accessed September 2020 at
https://coast.noaa.gov/data/digitalcoast/pdf/slr-new-mapping-tool.pdf , originally published in the Proceedings of the 2011 Solutions to Coastal Disasters Conference, American Society of Civil Engineers (ASCE), and reprinted with permission of ASCE.

Mills, Edwin S. “An Aggregative Model of Resource Allocation in a Metropolitan Area”. American Economic Review Papers and Proceedings 57, no. 2 (1967): pp. 197-210.

Muth, Richard F. Cities and housing. Chicago: University of Chicago Press (1969).

National Oceanic and Atmospheric Administration (NOAA) (2021a). “Billion-Dollar Weather and Climate Disasters: Summary Stats”. Washington DC: NOAA (2021). Accessed February 2021 at https://www.ncdc.noaa.gov/billions/summary-stats/US/2005-2020

National Oceanic and Atmospheric Administration (NOAA) (2021b). "Sea Level Rise Viewer". Washington DC: NOAA (2021). Accessed June 2020 at
https://coast.noaa.gov/digitalcoast/tools/slr.html

National Oceanic and Atmospheric Administration (NOAA) (2021c). "Digital Coast". Washington DC: NOAA (2021). Accessed June 2020 at https://coast.noaa.gov/digitalcoast/data/

Neumann, Barbara, Athanasios T. Vafeidis, Juliane Zimmermann and Robert J. Nicholls. "Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment". PloS one, 10, no. 3 (2015).

Ortega, F. and Süleyman Taşpınar. "Rising sea levels and sinking property values: Hurricane Sandy and New York's housing market". Journal of Urban Economics, 106 (2018), pp. 81-100.

Peter G. Peterson Foundation (2020). Budget Basics: The National Flood Insurance Program. Accessed March 2021 at:
https://www.pgpf.org/budget-basics/the-national-flood-insurance-program

Pörtner, Hans O., Debra C. Roberts, Valérie Masson-Delmotte, Panmao Zhai, Melinda Tignor, Elvira Poloczanska, Katja Mintenbeck, Andrés Alegría, Maike Nicolai, Andrew Okem, Jan Petzold, Bardhyl Rama and Nora M. Weyer. "IPCC Special Report on the Ocean and Cryosphere in a Changing Climate". Geneva, Switzerland: United National Inter-Governmental Panel on Climate Change (2019).

Pryce, Gwilym, Yu Chen & George Galster. "The Impact of Floods on House Prices: An Imperfect Information Approach with Myopia and Amnesia", Housing Studies, 26, no. 2 (2011): pp. 259-279.

Rappaport, J. and Jeffrey D. Sachs. "The United States as a Coastal Nation". Journal of Economic Growth, 8, no. 1 (2003): pp. 5-46.

Saiz, Albert. "The Geographic Determinants of Housing Supply". Quarterly Journal of Economics, 125, no. 3 (2010): pp. 1253–1296.

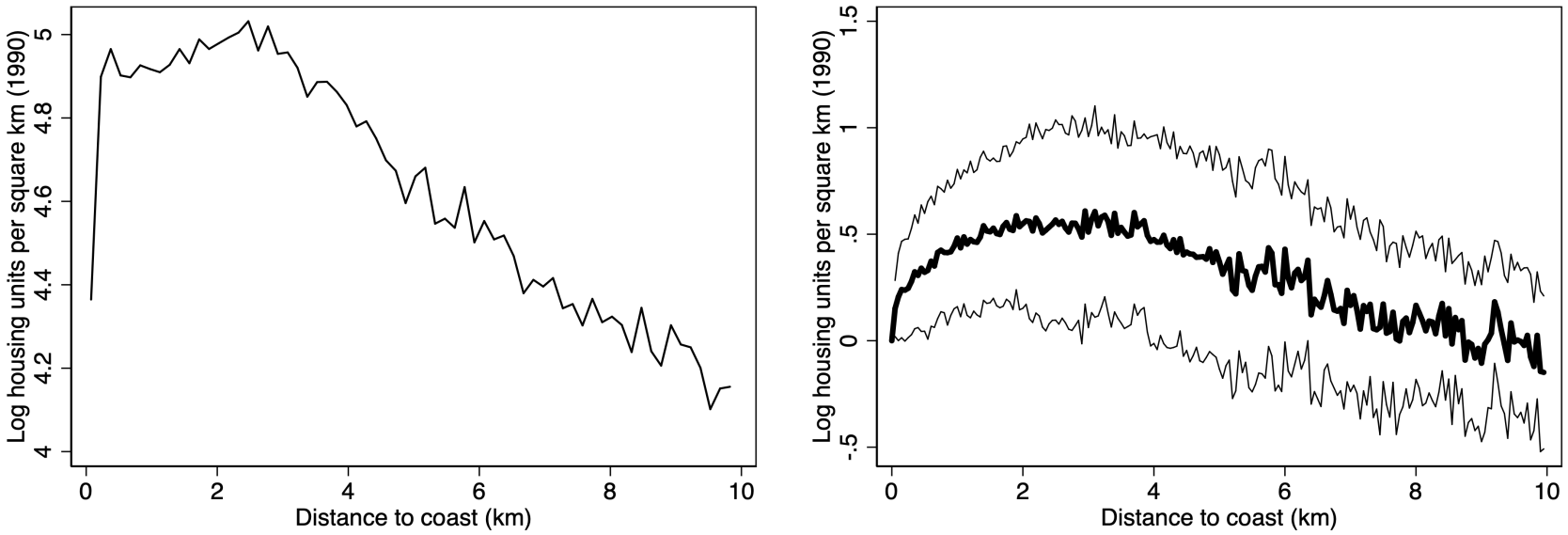
Simpson, Murray C., Colleen S.L. Mercer Clarke, John D. Clarke, Daniel Scott, and Alexander J. Clarke. Coastal Setbacks in Latin America and the Caribbean: A Study of Emerging Issues and Trends that Inform Guidelines for Coastal Planning and Development. Inter-American Development Bank VPS/ESG TECHNICAL NOTE No. IDB - TN - 476 (October 2012).

Statistica. "Homeownership rate in the U.S. 1990-2020." Accessed February 2021 at
https://www.statista.com/statistics/184902

US Congressional Budget Office (CBO). "The National Flood Insurance Program: Financial Soundness and Affordability", September 2017, available online from www.cbo.gov/publication/53028

US Congressional Budget Office (CBO). "Expected Costs of Damage from Hurricane Winds and Storm Related Flooding". Washington DC: Congressional Budget Office (2019). Accessed July 2020 at:
https://www.cbo.gov/system/files/2019-04/55019-ExpectedCostsFromWindStorm.pdf

US Department of Homeland Security, Office of Inspector General. "FEMA Needs to Improve Management of Its Flood Mapping Programs". Report for Office of the Inspector General, OIG-17-110 (2017).



(a) (b)

Figure 1: (Stylized Fact 1) Population concentrates near – but not right at – the coast. Panel (a) shows ln(housing units in 1990 per square kilometer) by distance to the coast for 150m bins. The figure in Panel (a) is based on our full sample of blocks whose centroids are within 10km of the US Gulf and Atlantic coasts. However, given that we partition the 10km from the coast into 150m bins, and take integers, the last observation, which is truncated (from 9.9-10km from the coast) is excluded from the figure. The figure in Panel (a) therefore uses information from 542,246 census blocks. In Panel (a) housing units are taken from census data at the block level, while land area in each distance bin from the coast is taken from the Landsat gridded data. The gridded data are used for calculating area because for sparsely populated blocks or empty blocks, which often cover a large area, the centroid is inadequate for capturing their distance distribution from the coast, as described in more detail in the Data Appendix. Panel (b) shows regression coefficients and 95% confidence intervals from a regression of ln(housing units in 1990 per square kilometer) on 50m distance bins. The standard errors are clustered by State. Density in Panel (b) is calculated using housing units and land area from census data at the block level. The issue noted above is less relevant here, since the regression restricts the sample to blocks with non-zero housing units (422,311 blocks).

Table 1: (Stylized fact 2) Places near the coast are asymmetric, with CBD closer to the coast

VARIABLES	(1) Asymmetry
Distance to coast bins (km) $\in$ [0,1)	0.14
	(0.03)
Distance to coast bins (km) $\in$ [1,2)	0.14
	(0.03)
Distance to coast bins (km) $\in$ [2,3)	0.08
	(0.03)
Distance to coast bins (km) $\in$ [3,4)	0.07
	(0.04)
Distance to coast bins (km) $\in$ [4,5)	-0.01
	(0.04)
Distance to coast bins (km) $\in$ [5,6)	-0.00
	(0.07)
Distance to coast bins (km) $\in$ [6,7)	-0.04
	(0.03)
Distance to coast bins (km) $\in$ [7,8)	-0.02
	(0.03)
Distance to coast bins (km) $\in$ [8,9)	-0.05
	(0.03)
Constant	0.54
	(0.03)
Observations	1,583

Notes: This table reports results from a regression of place asymmetry on 1 km distance bins from the coast. Standard errors in parentheses are clustered by state. Place asymmetry is defined as the ratio of the distance $|X_R - X_0|$ to the distance $|X_R - X_L|$. The sample here is restricted to places for which the mean distance to coast from the centroids of blocks within that place is less than 10km. The omitted category is [9,10).

Table 2: (Stylized fact 3) Bigger places are more asymmetric, conditional on being near the coast

	(1) Asymmetry	(2) Asymmetry	(3) Asymmetry	(4) Asymmetry
Ln($\lvert X_R - X_L \rvert$)	0.015		0.332	
	(0.007)		(0.071)	
Ln(area)		0.028		0.178
		(0.005)		(0.042)
Ln(Distance to coast)			0.114	0.225
			(0.077)	(0.093)
Ln($\lvert X_R - X_L \rvert$)*Ln(Distance to coast)			-0.030	
			(0.009)	
Ln(area)*Ln(Distance to coast)				-0.018
				(0.006)
Constant	0.497	0.166	-1.102	-1.749
	(0.061)	(0.068)	(0.571)	(0.693)
Observations	1,583	1,583	1,583	1,583

Notes: This table reports results from regressions of place asymmetry on measures of place size, place distance to the coast, and their interaction, all in logs. Place asymmetry is defined as the ratio of the distance $|X_R - X_0|$ to the distance $|X_R - X_L|$. The area variable is the sum of the area of blocks that are within a place's boundaries. The sample here is restricted to places for which the mean distance to coast from the centroids of blocks within that place is less than 10km. Standard errors in parentheses are clustered by state.

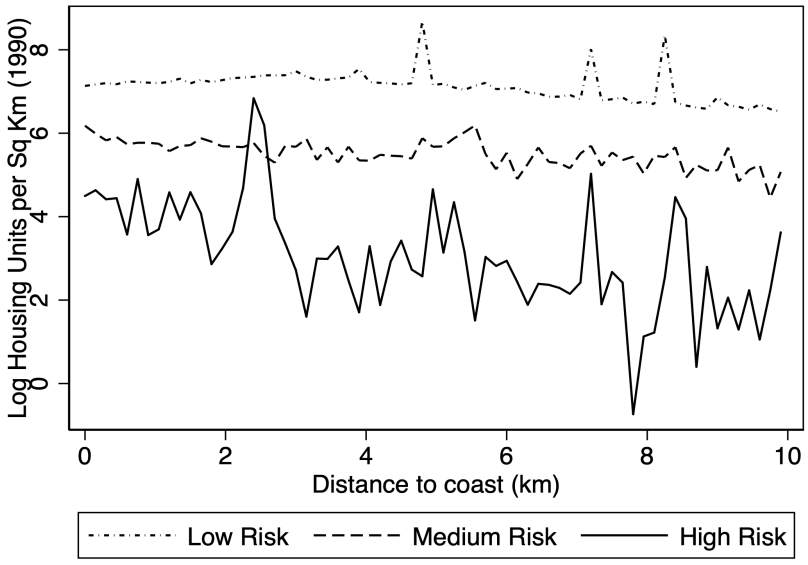
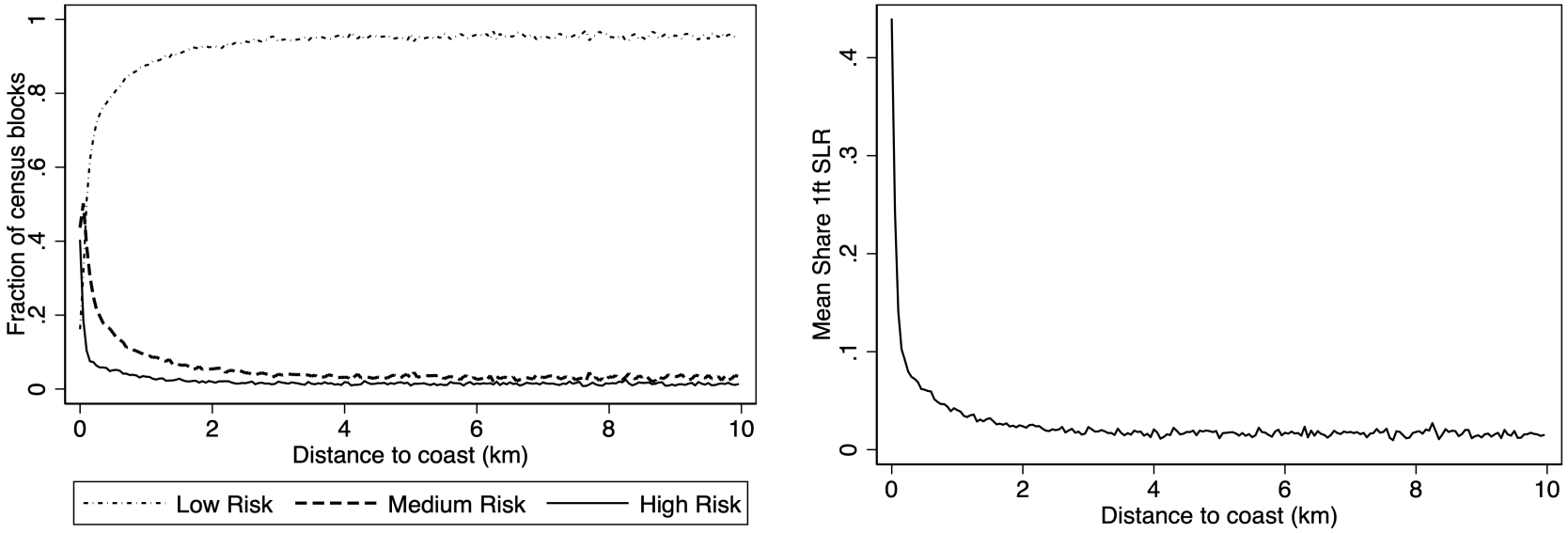


Figure 2: (Stylized fact 4) Areas that are highly prone to sea level rise (SLR) are less built. The figure shows the log of average block level density (housing units per square kilometer in 1990, based on census data) by distance to the coast in 150m bins, and share of area under water with 1 foot of sea level rise. The three risk categories are defined by the share of each census block that will be under water at high tide if sea levels rise by 1 foot (30.4cm): we label blocks as high risk if the share of 1ft SLR is > 0.5, as medium risk where $0 <$ share 1ft SLR $<= 0.5$, and as low risk where share of 1ft SLR $= 0$. This risk reflects odds of flooding even today, without any SLR. As the figure shows, at each distance from the coast, the riskier areas are more sparsely built.



(a) (b)

Figure 3: (Stylized fact 5): Flood risk helps explain why people do not build right on the coast. Panel (a) shows the fraction of census blocks in each risk category by distance to the coast in 50m bins. The three risk categories are defined by the share of each census block that will be under water at high tide if sea levels rise by 1 foot (30.4cm): we label blocks as high risk if the share of 1ft SLR is > 0.5, as medium risk where $0 <$ share 1ft SLR $<= 0.5$, and as low risk where share of 1ft SLR $= 0$. This risk reflects odds of flooding even today, without any SLR. Panel (b) shows the mean share of block area that is subject to 1ft SLR, by distance to the coast in 50m bins.

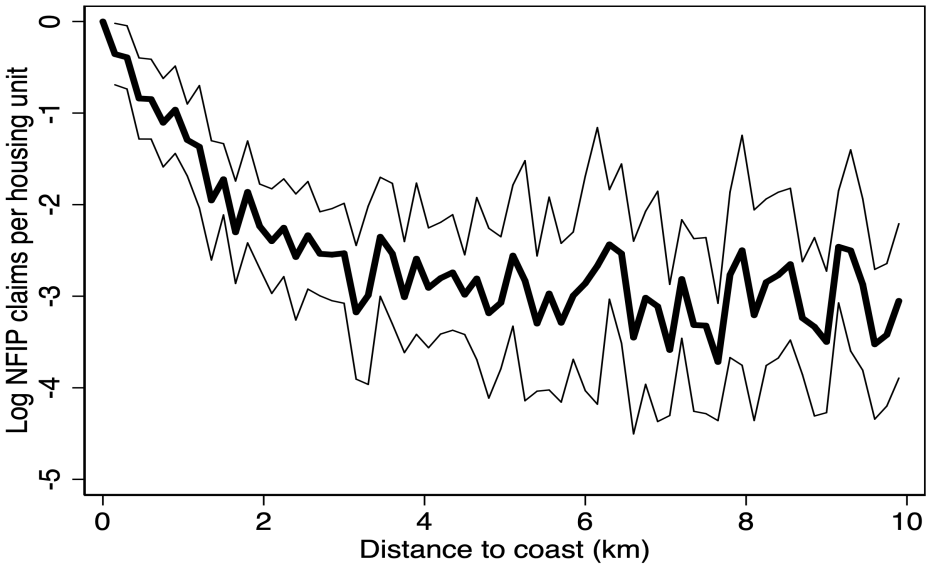


Figure 4: (Stylized fact 6) Damages from flooding decline rapidly with distance from the coast. The figure shows the estimated coefficients and 95% confidence interval from a regressions of NFIP claims per housing unit, on 150m distance bins from the coast. Standard errors are clustered by state. The NFIP claims data are observed at the census tract level, and have been aggregated across all available years (1973-2019), and converted to 2020 US dollars. Housing units are also observed at the census tract level, and are taken from 2014-2018 estimates. Full details of the data sources used are included in the Data Appendix.

Table 3: (Stylized fact 7) Much construction near the coast took place in areas with SLR risk

	(1) Housing units in risky blocks (millions)	(2) Housing units in all blocks (millions)	(3) Fraction of housing units in risky blocks (%)
1990	1.75	14.87	12%
2010	2.58	18.11	14%
Change (1990-2010)	0.83	3.24	26%

Notes: Column (1) reports numbers of housing units in census blocks whose centroids are within 10km of the coast and where at least some portion of the census block will be under water at high tide if sea levels rise by 1 foot (30.4cm). Column (2) reports numbers of housing units in all blocks whose centroids are within 10km of the coast. Column (3) reports the fraction of housing units in census blocks whose centroids are within 10km of the coast that are also in blocks where at least some portion of the census block will be under water at high tide if sea levels rise by 1 foot (30.4cm).

Table 4: (Stylized fact 8) SLR-prone areas were developed in dense tracts but not in sparse ones

	(1)	(2)	(3)	(4)
Housing units per sq km in containing census tracts	$\leq$10	$(10, 100]$	$(100, 1000]$	>1000
Share 1ft SLR	-3.07	-3.13	2.35	6.76
	(1.72)	(1.22)	(1.21)	(3.55)
Constant	4.14	7.58	5.72	4.62
	(1.12)	(1.23)	(0.78)	(1.08)
Observations	24,920	149,463	283,207	86,471

Notes: The outcome in each case is the change in housing units at the block level from 1990-2010. Columns divide the data by levels of housing units per square km in census tracts (1990), excluding own block. Standard errors in parentheses are clustered by state. Urban population density: at least 386 people (not housing units) per square km. Results robust to controlling for log distance to the coast.

Table 5: (Stylized fact 9) SLR-prone areas developed in dense tracts were those with least SLR

	(1)	(2)
Share 1ft SLR $\in$ (0.0,0.1)	13.53	
	(8.15)	
Share 1ft SLR $\in$ [0.1,0.2)	15.91	
	(5.61)	
Share 1ft SLR $\in$ [0.2,0.3)	8.26	
	(3.94)	
Share 1ft SLR $\in$ [0.3,0.4)	16.53	
	(6.74)	
Share 1ft SLR $\in$ [0.4,0.5)	10.40	
	(6.34)	
Share 1ft SLR $\in$ [0.5,0.6)	4.54	
	(3.42)	
Share 1ft SLR $\in$ [0.6,0.7)	-0.33	
	(2.63)	
Share 1ft SLR $\in$ [0.7,0.8)	-2.60	
	(3.29)	
Share 1ft SLR $\in$ [0.8,0.9)	-2.85	
	(1.63)	
Share 1ft SLR $\in$ [0.9,1.0)	-2.79	
	(2.02)	
Share 1ft SLR = 1	-3.88	
	(1.02)	
Some SLR		15.83
		(6.25)
Share 1ft SLR		-19.40
		(7.97)
Constant	4.24	4.24
	(1.19)	(1.19)
Observations	86,471	86,471

Notes: The outcome in each case is the change in housing units at the block level from 1990-2010. The sample here is restricted to blocks with tract housing density >1000 housing units per square km. The omitted category, 0 SLR, accounts for 95.5% of the blocks with this level of tract housing density in 1990. Standard errors in parentheses are clustered by state. In column (2), the variable Some SLR is an indicator for blocks that have share 1ft SLR >0, while Share 1ft SLR is a continuous measure of the share of each block prone to 1ft SLR. Results robust to controlling for log distance to the coast.

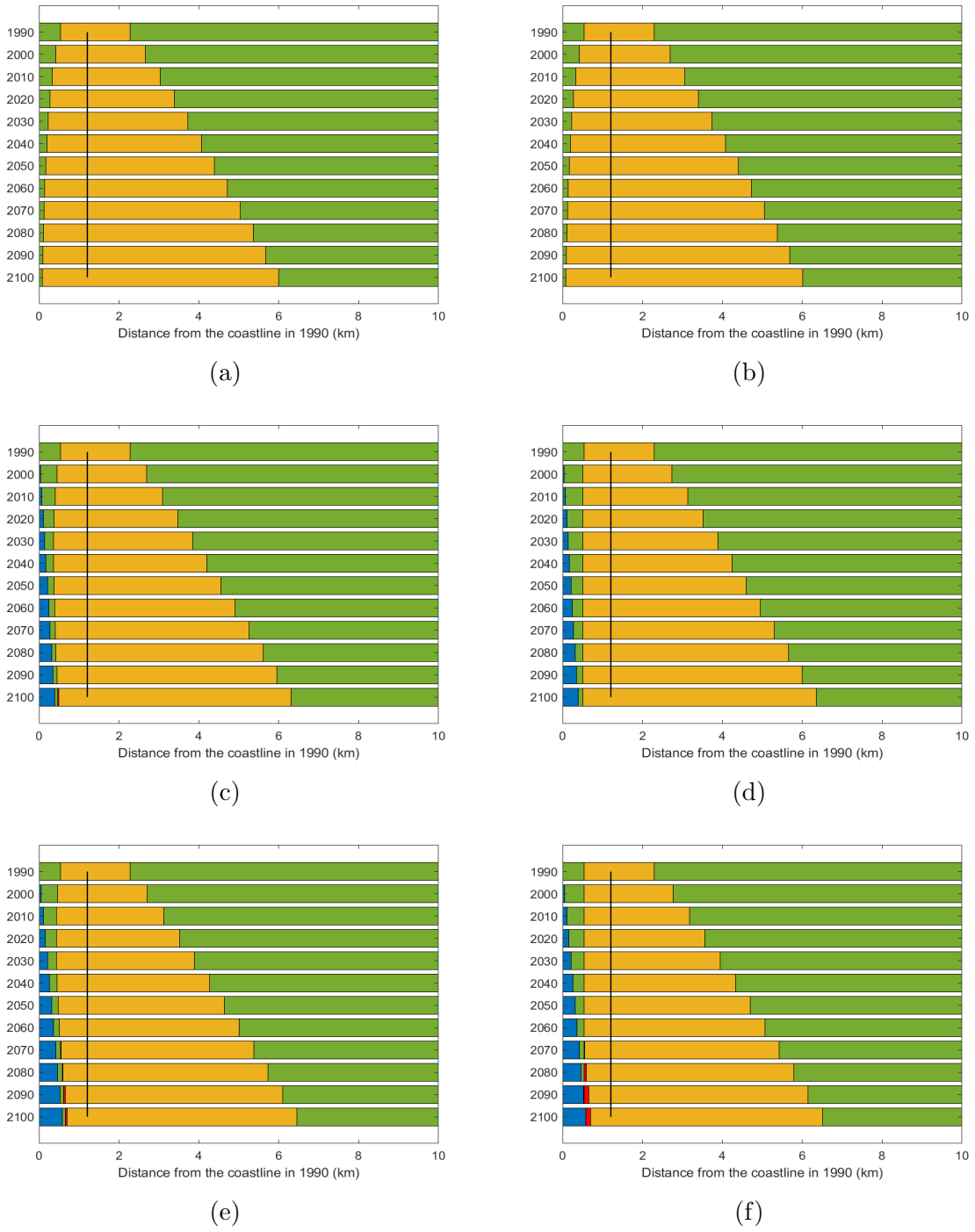


Figure 5: Simulation results. The figure shows results of simulations, as described in the main text, and based on parameter values detailed in Table A5. Panel (a) shows results from the static model with no SLR. Panel (b) is the same but for the dynamic version of the model. Panel (c) shows results from the static model with baseline SLR (0.0577m per decade, and CBD elevation at 2m). Panel (d) shows the same for the dynamic model. Finally, panel (e) shows results from the static model with faster SLR (equivalent to a city with CBD at 1.33m elevation), and Panel (f) the equivalent results for the dynamic model. In each panel of the figure yellow denotes the city, green denotes agricultural land, blue denotes the sea, and red denotes areas whose housing density declined at least 10 percent from the maximum level. The time period is on the vertical axis in 10 year time-steps, and the horizontal axis shows distance to the coast in km.

For Online Publication - Appendices for:

Cities and the Sea Level

Yatang Lin (Hong Kong University of Science and Technology)
Thomas K.J. McDermott (National University of Ireland Galway and LSE)
Guy Michaels (London School of Economics)

Online Appendix A: Model

Proposition 1 *Define housing density* $dens_t(x) \equiv \frac{1}{h_t(x)}$*, we get the following result:*

$$\text{For each period } t = 1, ..., T\text{: if } x < x_0 \text{ then } \frac{\partial \ln(dens_t(x))}{\partial x} > 0\text{; if } x > x_0 \text{ then } \frac{\partial \ln(dens_t(x))}{\partial x} < 0. \quad (15)$$

Proof. Define the net income: $I_t(x) \equiv w_t - |x - x_0|$, and consider only locations where $I_t(x) > 0$ as potential city locations.[64] With Cobb-Douglas preferences $c_t(x) = \alpha I(x), h_t(x) = (1-\alpha)\frac{I_t(x)}{p_t(x)}$. Plugging these into the residents' indifference condition $U(c_t, h_t, x) = \bar{U}$, solving for $p_t(x)$, and taking logarithms we get

$$\ln(p_t(x)) = \frac{1}{1-\alpha}\ln\left(\alpha^{\alpha}(1-\alpha)^{1-\alpha}\right) + \frac{1}{1-\alpha}\ln(I_t(x)) - \frac{1}{1-\alpha}\ln\left(\bar{U} + \theta_1 x + \theta_2 x^{-\sigma}\right). \quad (16)$$

Due to the preferences $h_t(x) = (1-\alpha)\frac{I_t(x)}{p_t(x)}$, and using (16) we get

$$\ln(dens_t(x)) = \frac{\alpha}{1-\alpha}\ln(\alpha) + \frac{\alpha}{1-\alpha}\ln(I_t(x)) - \frac{1}{1-\alpha}\ln\left(\bar{U} + \theta_1 x + \theta_2 x^{-\sigma}\right). \quad (17)$$

and using the condition for the city founder we get the sign of $\frac{\partial \ln(dens_t(x))}{\partial x}$ ■

Proposition 2 *With constant wages, the city will be developed asymmetrically around the CBD, so:* $|x_{Rt} - x_0| > |x_0 - x_{Lt}|$.

Proof. Define the indirect utility function on housing and consumption $u(c(x), h(x), x)$ as $v(p(x), x)$, define the location specific part of utility $U_{Loc}(x) \equiv -\theta_1 x - \theta_2 x^{-\sigma}$

$$\begin{aligned}
\forall \hat{x} > x_0, \quad U_{Loc}(\hat{x}) &= -\theta_1 \hat{x} - \theta_2 \hat{x}^{-\sigma} \\
=> U'_{Loc}(\hat{x}) &= -\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)} \\
=> U_{Loc}(2x_0 - x_{Lt}) &= U_{Loc}(x_0) + \int_{x_0}^{2x_0 - x_{Lt}} U'_{Loc}(\hat{x}) d\hat{x} \\
&= U_{Loc}(x_0) + \int_{x_0}^{2x_0 - x_{Lt}} (-\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)}) d\hat{x}
\end{aligned}$$

[64] Any location x, which is too far from the CBD, such that $I(x) \leq 0$, remains unpopulated.

$$\forall \hat{x} < x_0, \quad U_{Loc}(\hat{x}) = -\theta_1 \hat{x} - \theta_2 \hat{x}^{-\sigma}$$
$$=> U_{Loc}(x_{Lt}) = U_{Loc}(x_0) + \int_{x_{Lt}}^{x_0} U'_{Loc}(\hat{x}) d\hat{x}$$
$$= U_{Loc}(x_0) + \int_{x_{Lt}}^{x_0} (-\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)}) d\hat{x}$$

$$U_{Loc}(2x_0 - x_{Lt}) - U_{Loc}(x_{Lt}) = \int_{x_0}^{2x_0 - x_{Lt}} (-\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)}) d\hat{x} + \int_{x_{Lt}}^{x_0} (-\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)}) d\hat{x}$$
$$= \int_{x_{Lt}}^{2x_0 - x_{Lt}} (-\theta_1 + \theta_2 \sigma \hat{x}^{-(\sigma+1)}) d\hat{x}$$

Now with $x_0 = \left(\frac{\sigma\theta_2}{\theta_1}\right)^{\frac{1}{\sigma+1}}$, we have $\theta_1 = \sigma\theta_2 x_0^{-(\sigma+1)}$

$$U_{Loc}(2x_0 - x_{Lt}) - U_{Loc}(x_{Lt}) = \int_{x_{Lt}}^{2x_0 - x_{Lt}} (\theta_2 \sigma \hat{x}^{-(\sigma+1)} - \theta_2 \sigma x_0^{-(\sigma+1)}) d\hat{x} > 0$$

We observe that $U_{Loc}(2x_0 - x_{Lt}) - U_{Loc}(x_0) > 0$ by Jensen's Inequality, since $\frac{\partial^2 x^{-(\sigma+1)}}{\partial x^2} > 0$ or in other words, the function is convex.

The utility equalization condition stipulates that $U(2x_0 - x_{Lt})) = U(x_{Lt}) = \bar{U}$. So we should have $v(p(2x_0 - x_{Lt}), 2x_0 - x_{Lt}) < v(p(x_{Lt}), x_{Lt})$. Since $v(p(x), x)$ is a function of price $p(x)$ and commuting time cost ($|x - x_0|$) only, and $|2x_0 - x_{Lt} - x_0| = |x_{Lt} - x_0|$, we have $p(2x_0 - x_{Lt}) > p(x_{Lt})$.

Meanwhile, given constant construction cost within the city, the zero-profit condition for the developers yield $E\left[\sum_{s=t}^{T} \delta^t p_t(x_{Lt})\right] = E\left[\sum_{s=t}^{T} \delta^t p_t(x_{Rt})\right] = c_s$

From the developers' indifference, we get $\forall t, \quad p_t(x_{Rt}) = p_t(x_{Lt})$

Applying the Lemma below, we have $2x_0 - x_{Lt} < x_{Rt}$, so that $x_{Rt} - x_0 < x_0 - x_{Lt}$, which is equivalent to $|x_{Rt} - x_0| > |x_0 - x_{Lt}|$.

■

Lemma 1 *When $x > x_0$, $\frac{\partial p(x)}{\partial x} < 0$*

Proof. With utility equalization $U(c(x), h(x), x) = \bar{U}$, we have $V(p(x), x) - \bar{U} = 0$.

$$\frac{\partial p(x)}{\partial x} = -\frac{\frac{V(p(x),x)}{p(x)}}{\frac{V(p(x),x)}{x}}$$

Since we have $\frac{\partial V(p(x),x)}{\partial p(x)} < 0$ and $\frac{\partial V(p(x),x)}{\partial x} < 0$ when $x > x_0$, so $\frac{\partial p(x)}{\partial x} < 0$

Note that here $\partial V(p(x), x)/\partial x < 0$ is obtained based on the fact that x enters the utility function directly through $|x - x_0|$ and $-\theta_1 x - \theta_2 x^{-\sigma}$, both of which are decreasing in x when $x > x_0$. ■

Proposition 4 $\lim_{x_{Lt} \to x_0, x_0 \leftarrow x_{Rt}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$

Proof. As $x_{Lt} \to x_0$, by continuity

$$U_{Loc}(2x_0 - x_{Lt}) - U_{Loc}(x_{Lt}) = \int_{x_{Lt}}^{2x_0 - x_{Lt}} (\theta_2 \sigma \hat{x}^{-(\sigma+1)} - \theta_2 \sigma x_0^{-(\sigma+1)}) d\hat{x} \to 0$$

Following the same logic in the proof in proposition above, we should have $v(p(2x_0 - x_{Lt}), 2x_0 - x_{Lt}) \to v(p(x_{Lt}), x_{Lt})$, and as a result $p(2x_0 - x_{Lt}) \to p(x_{Lt}) = p(x_{Rt})$

$p(x)$ is a monotonic and continuous function of x when $x > x_0$. Therefore, $2x_0 - x_{Lt} \to x_{Rt}$, we have $\lim_{x_{Lt} \to x_0, x_0 \leftarrow x_{Rt}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$. ■

Online Appendix B: Data

Defining the Sample

In our analysis we focus on US Atlantic and Gulf coasts. These areas have experienced some of the fastest rates of local sea level rise in the world during the 20th century, and this trend is expected to continue and raise the frequency and severity of floods (Dahl et al. 2017). There are parts of 18 states and one district included in our data. These are: Alabama, Connecticut, Delaware, District of Columbia, Florida, Georgia, Louisiana, Maine, Maryland, Massachusetts, Mississippi, New Hampshire, New Jersey, New York, North Carolina, Rhode Island, South Carolina, Texas, and Virginia.

Our focus is on areas within 10km of the coast. Specifically, we restrict our analysis to census blocks, or census tracts, or cells, whose centroid is within 10km of the coast.[65] For the purposes of distance calculations, we define the coast using administrative boundary shapefiles for the US from Database of Global Administrative Boundaries (GADM 2018).[66] This includes some rivers, such as the Charles in Boston, East River in New York City, and the Potomac in Washington, DC, as part of the coastline. But lakes and upstream sections of rivers are typically excluded, and consequently Philadelphia, New Orleans, and Houston, are largely outside our dataset.

Data on places are taken from the NHGIS archive (Manson et al. 2019) based on census data and definitions.[67] Specifically, the place level data that we use include two sets of shapefiles: polygons that we use to define place boundaries; and points that we use to represent the CBD of each place.[68] These shapefiles are matched to our block-level data, using GIS, and we define place level characteristics using data on blocks that intersect each place shapefile.

Data on Sea Level Rise

[65] For census designated places, we restrict our attention to places for which the mean distance to coast from the centroids of blocks within that place is less than 10km. One exception where we observe places more than 10km from the coast is in the appendix figure that illustrates the relative asymmetry of coastal places (Figure A2), which shows the CBD of the city of Boston, and a number of census designated places whose own CBD is within 25km of Boston's.

[66] The GADM administrative boundary data have the advantage of providing a consistent definition of the coastline, at a finer resolution than the outline of the census blocks. The shapefiles we use are from the GADM database www.gadm.org, version 2.8, November 2015 (link last accessed September 2020).

[67] The definition of places used by the Census Bureau is detailed here:
https://www.census.gov/content/dam/Census/data/developers/understandingplace.pdf.

[68] Where possible, NHGIS assigns place points based on the GNIS coordinates of each place's historical or functional center (typically the central business district). In cases where these coordinates are not available, the point denotes the geographical centroid of the place. More detail is available here: https://www.nhgis.org/documentation/gis-data/place-points.

Detailed maps of areas anticipated to be inundated for various future sea level rise scenarios were obtained from NOAA's digital coast platform.[69] The maps show inland extent (and relative depth) of inundation for scenarios of sea level rise from 0 to 6 feet above mean higher high water (MHHW). The inundation maps are created by NOAA using a modified "bathtub" model, whereby each of 7 different SLR scenarios (0-6ft, in 1ft increments) are added to Mean Higher High Water (MHHW) – the base elevation.[70] This surface is then subtracted from a digital elevation model (based on lidar data).

We focus on the extent of inland inundation (ignoring relative depths), and consider only areas that are "hydrologically connected" (according to the digital elevation model used by NOAA). Separate maps are also available for low-lying areas, which are considered hydrologically "unconnected" areas that may flood (in addition to the locations identified as "hydrologically connected").[71] Importantly, the mapping process also takes account of major federal leveed areas, which are assumed, for the purposes of creating these inundation maps, to be high enough and strong enough to prevent inundation, regardless of the SLR scenario assumed. In other words, the maps show areas that are expected to be inundated for a given SLR scenario, and which are not currently protected by major federal levees.[72]

In our analysis we take the somewhat conservative approach of focusing only on the 1ft sea level rise scenarios. These identify areas would be under water at high tide if SLR is 1 foot (approx. 30.5cm), which could be a reality for many locations within just a few decades. Many areas on US Atlantic and Gulf coasts are projected to experience 1ft SLR by 2045 (see Dahl et al. 2017).

Information on sea level rise was added to the blocks (and cells) by intersecting the shapefiles for blocks (cells) with shapefiles of areas expected to be inundated for 1ft of sea level rise. Given the large size of the SLR data, this was done on a state-by-state basis, with the resulting state-level datasets exported to Stata and combined into a single dataset (one each for blocks and for cells).

Given the sheer scale of the data, and the amount of geo-processing involved, the merging of SLR data into our main block-level and gridcell datasets was scrupulously checked and double checked - including visual inspection of every individual intersection, and the inspection of individual output datasets from GIS, including checking of summary statistics on a state-by-state basis.

For many states, the SLR information from NOAA (described above) comes in the form of multiple shapefiles, for different segments of the state's coastline. In general these were merged together and then dissolved to form a single feature (in order to remove any potential for within-state overlaps), before intersecting with the blocks (or cells). However, there were three states in our sample - Florida, Maryland and North Carolina (FL, MD and NC) - where this merging and dissolving process did not work, due to the size of the individual SLR layers. The potential for (slight) overlaps between individual SLR layers for these three states was handled as follows:

For NC and MD, the individual SLR layers were intersected separately with the grid, with the resulting outputs merged to the gridded dataset one-by-one, such that where there are overlaps (i.e.

[69] The data are available from https://coast.noaa.gov/slrdata/ (last accessed September 2020).

[70] The base elevation is NOAA VDATUM MHHW.

[71] More details on the NOAA SLR maps are available at https://coast.noaa.gov/data/digitalcoast/pdf/slr-faq.pdf (last accessed September 2020).

[72] As an example of this, New Orleans is assumed not to be inundated in any of the SLR scenarios.

cells that overlap with more than one SLR layer), the SLR values for that gridcell are overwritten, as opposed to summed. This could lead to some very minor understating of the true SLR area for a small number of cells, but we considered it preferable to potentially double counting some areas of SLR. For similar reasons, it is also possible that we understate true SLR for some cells near state boundaries (where a gridcell is transected by a state boundary). This is not an issue for the blocks data, as block boundaries line up exactly with state boundaries.

For FL, an additional layer based on the overlaps between individual SLR layers was created and these areas of overlap were subtracted from the calculated SLR area for the relevant cells.

In merging SLR information to the blocks data, the procedure of over-writing SLR areas for blocks that appear in multiple SLR layers might not be appropriate, given that (some) blocks are relatively large geographically compared to grid cells. For each of FL, MD and NC an additional layer based on overlaps between SLR layers was created. Areas of overlap are then subtracted from the original estimates of SLR area for the relevant blocks.

This issue of overlapping SLR layers for FL, MD and NC, in practice only affects around 2,000 cells (out of a total sample of over 6 million cells) and around 240 blocks (out of a total sample of over 544,000 blocks). So in either dataset it is only 0.04% of the sample that is affected.

A total of six blocks (one in FL and five in NC) with what looked like anomalously large share 1ft SLR values (greater than 1) were dropped from the blocks dataset. One block (in NC) with a negative area value was also dropped from the dataset. For the remaining blocks in our data, the share 1ft SLR variable was winsorized to 1 (this affected 1,830 blocks, with a maximum share 1ft SLR of 1.003).

Housing Data

Our housing data (prices and quantities) come from US Census housing data, at the block level. For 1990 these come from the census; what we refer to as "2010" are actually data for 2006 - 2010 from the Annual Communities Survey. The data are obtained from the National Historical Geographic Information System (NHGIS) data archive (Manson et al. 2019).[73]

Because block boundaries change from one census to another, any analysis comparing housing data from different censuses requires harmonization, or interpolated, into a common geography. We opt to harmonize the data backwardsï¿½ï¿½ to the earlier of the years we are comparing (1990). This is because many more census blocks split than merge over time. Using the earlier year therefore more often results in the simpler procedure of adding two block groups together, rather than attempting to allocate data from one block to two new blocks. Additional discussion of the various pitfalls of harmonizing census data over time is also available from Logan, Xu & Stults (2014).

We use the geographical crosswalk files produced by NHGIS. Such files provide, for every census block in the 'source' census geography, the approximate portion of its area lying within the boundaries of another, 'target' geography's block. Specifically, the information available in the NHGIS mapping is the fraction of the source geography s block i intersecting with each target geography t block j (block pairings with 0 intersection are not included in the file), denoted as p_{areaij} :

$$p_{areaij} \equiv \frac{(B_i^s \cap B_j^t)}{(B_i^s)} \tag{18}$$

[73]Data are available to download from: https://data2.nhgis.org/

Where B_i^s is the area of block B_i as per the source census geography, s, and B_j^t is the area of block B_j as per the target census geography, t.

The NHGIS crosswalk files, however, only harmonise forwards. The target geography is the latest available (2010 Census geography), while the source geographies available are either the 1990 or 2000 US Census boundaries. Harmonising backwards from 2010 to 1990 therefore requires that the intersect value contained within the crosswalk files, p_{areaij} be amended as follows:

$$p_{areaji} = \frac{(B_i^{90} \cap B_j^{10})}{(B_j^{10})} = \frac{(B_i^{90} \cap B_j^{10})}{(B_i^{90})} * \frac{(B_i^{90})}{(B_j^{10})} \tag{19}$$

In other words, the area of the intersect we are interested in, p_{areaji}, is obtained by taking the forward-looking intersects provided in the crosswalk files, p_{areaij}, and multiplying by the ratio of block i area in 1990 to block j area in 2010 .

Subsequently, we allocate housing units from each 2010 block to each year t block as:

$$H_{ji}^{90,10} = p_{ji} * H_j^{10} \tag{20}$$

Where H_j^{10} denotes the number of housing units in 2010 block j, and $H_{ji}^{90,10}$ is the number of housing units in 2010 allocated to 1990 block i.

Data on Land Cover

As an alternative measure of the extent and intensity of development in coastal areas, we use land cover[74] data from NOAA's Coastal Change Analysis Program (C-CAP).[75] Through the C-CAP program, NOAA produces nationally standardized land cover and land change data for the coastal regions of the U.S.[76] The C-CAP program is part of the National Land Cover Database (NLCD), and is considered the coastal expression of this national database. The use of standardized data and consistent methods allows for comparison over time and across different regions. C-CAP data are developed, primarily, from Landsat Thematic Mapper (TM) satellite imagery. The Landsat data have a 30m pixel size. The minimum inland extent of C-CAP's mapping boundary is based on state-designated Coastal Zone Management (CZM) boundaries, NOAA's Coastal Assessment Framework (which includes definitions of estuarine and coastal drainage areas, or EDAs and CDAs), and designation of coastal counties (counties that are at least 15% within the EDA and CDA).[77]

The C-CAP data involve the classification of each 30m x 30m pixel into one of 25 categories - including various classifications of open water, wetlands, agricultural land, forest etc. The C-CAP data are produced to meet an 85% accuracy specification, i.e. that 85 times out of 100 the C-CAP classification is correct, reflecting the majority land cover call on the ground for the same area.

Of particular interest for our purposes are the four "developed land" categories. Each represents

[74]Land cover, which captures the physical state of land resources, is distinct from land use, which denotes how the land is being used. The C-CAP data represent land cover, so we stick to this terminology throughout.

[75]The data are available from
https://coast.noaa.gov/digitalcoast/data/.

[76]The C-CAP homepage is
https://coast.noaa.gov/digitalcoast/data/ccapregiona.

[77]For more details see here:
https://coast.noaa.gov/data/digitalcoast/pdf/ccap-faq-regional.pdf.

a different extent of constructed surfaces – from "developed - high intensity" where constructed materials account for 80 to 100 percent of the total cover in the pixel, to the "developed - open space" category, where constructed material accounts for less than 20 percent of land cover, with two intermediate classes; "developed medium intensity" (50-79 percent constructed material), and "developed - low intensity" (21-49 percent).

The C-CAP data are an imperfect measure of development intensity, given that we cannot distinguish different use types; what is observed in the data is the land cover type, which in the case of developed areas simply corresponds to the extent of "constructed materials" in a given area. For example, according to the C-CAP documentation, the "developed - high intensity' 'classification "includes heavily built-up urban centers and large constructed surfaces in suburban and rural areas with a variety of land uses". Similarly, the "developed - medium intensity" class "commonly includes multi- and single-family housing areas, especially in suburban neighborhoods, but may include all types of land use".[78] It is therefore not clear what exactly land cover represents (e.g. developed areas may include roads and parking lots) and there is also no price or quality information included in these data.

Balanced against these constraints are some advantages for our purposes: The C-CAP data represent an objective measure of development (captured via satellite imagery) that is consistent over time and covers the entire coastal area of the contiguous US. They also represent an even partition of land (unlike blocks or block groups, which vary in geographical size according to population density), and land cover data include non-residential construction, which would be not be captured by our housing data.

For some states (for example, those near the Great Lakes) there are image files available from 1975. For others the data start in 1992. However, for the complete US coast, the data start in 1996, which we take as the start of our sample period. The C-CAP data are generally updated every 5 years, and at the time of compiling our dataset, the latest available year was 2010, so we have four observations for every location in our data: 1996, 2001, 2006, and 2010. The updates involve mapping only those areas that have changed in the interval. Typically, according to the C-CAP documentation, about 20% of areas have changed in any 5 year period.[79]

For use in our analysis, we take land cover observations in 1996 and 2010 (and calculate changes from 1996-2010). We aggregate the 30m x 30m pixels in the C-CAP data to 150 x 150 meter cells.[80] In aggregating the underlying data, we take the midpoints of the ranges for each "developed land" category and assign that value to each pixel, averaging across a maximum of 25 pixels to create our gridded data. Pixels not classified in any of the "developed land" categories take a value of 0 for these purposes.

Data on Restricted Development Areas

[78]For more details on C-CAP classifications see
https://coast.noaa.gov/digitalcoast/training/ccap-land-cover-classifications.html. Comparison with other land classification systems is included in the C-CAP FAQ, here:
https://coast.noaa.gov/data/digitalcoast/pdf/ccap-faq-regional.pd.

[79]See the FAQ:
https://coast.noaa.gov/data/digitalcoast/pdf/ccap-faq-regional.pdf.

[80]In some cases the cells in our grid are smaller than this, given that they are cut by the coast on one side and by a boundary 10km from the coast on the other.

In certain coastal areas, development may not occur because this has been restricted by governments. For example, developers often cannot build on coastal areas designated as protected, whether for conservation purposes, natural resource management or recreation (e.g. the Everglades in Florida); on smaller city parks (e.g. Central Park in New York City or Boston Common), or on land owned by the military for naval or air force bases (e.g. Norfolk naval base in Virginia). Therefore, for every US state, these three types of areas were identified and labelled as such in the main analysis.

Information on areas classed as protected, as of 2017, was sourced from the Protected Area Database for the United States (PAD-US), which is compiled by the USGS.[81] An area that is protected can fall into one of four categories, known as the protection "Status". Status 1 constitutes the highest level of protection from development, Status 2 and 3 confer successively less (but still significant) levels of protection, while Status 4 denotes areas where some conversion of the natural land cover may be permitted.[82] We consider as areas where development is restricted those which are designated Status 1-3.

Information on local parks, gardens and woodland in urban coastal areas, as of 2019, is taken from ESRI.[83] Finally, information on the location of US military bases along the US coastline is taken from the US Department of Transport.[84]

The shapefiles for each of these were combined using ArcGIS. In some cases, the areas mapped by each of these datasets overlapped. Certain parts of the coast could, for example, be listed as both a local park and part of PAD-US designated protected area. To avoid double-counting, therefore, overlaps are assigned to just one of the three restriction types, with priority given to military-owned land first and PAD-US protected areas second. The resulting maps were intersected with both the census blocks and grid cells of interest (those within 10km of the coast), and the percentage of each block or cell covered by PAD-US protected areas, local parks or military-owned land was calculated. For each individual census block and grid cell, these percentages were added to obtain a value for the percentage of land where development is restricted. As a robustness exercise any blocks where this exceeded 50% were dropped and the relevant stylized facts replicated - the results (included as additional figures and tables in the appendix) remain qualitatively and quantitatively unchanged by this restriction on the data.

Damages from Flooding

The data used to evaluate the economic damages caused by coastal flooding, and how this varies with distance from the coast, are taken from the National Flood Insurance Program (NFIP). This program, operated by FEMA, offers building and contents insurance to home-owners, renters or businesses who are at risk of incurring damages from flooding. Information on the amounts reimbursed through the NFIP in different places, therefore, can provide an insight into the geographical patterns of damages from coastal flooding.

[81] Data available here:
https://www.sciencebase.gov/catalog/item/5963ea3fe4b0d1f9f059d955..

[82] A full list of Status definitions can be found here:
https://www.sciencebase.gov/catalog/item/56bba50ce4b08d617f657956.

[83] Data available here:
https://www.arcgis.com/home/item.html?id=578968f975774d3fab79fe56c8c90941.

[84] Data available here:
http://osav-usdot.opendata.arcgis.com/datasets/d163fcde26de4d21aa06aa141ce3a662_0.

While claims made under the NFIP by no means capture the totality of economic losses from coastal floods (or in fact the totality of residential losses from flooding, as much damage may be uninsured), the NFIP data have the advantage of being available at a relatively fine level of geographic disaggregation – the census tract level. This contrasts with data on other measures of flooding damages, for example FEMA's Public and Individual Assistance programs, which are only broken down to the county level. Moreover, our analysis focuses on housing, so that data on insured residential losses from coastal floods, at a relatively fine spatial scale, represent an appropriate way to estimate how damages from flooding vary with distance from the coast.

We use data on every NFIP policy where a claim was made for reimbursement of eligible losses, and where at least $100 was subsequently paid out.[85] For all such claims made, the census tract of the property or business with respect to which the payment was made is recorded. Of 31,466 census tracts in our 19 states of interest, 25,128 (80%) recorded at least one claim for reimbursement through the NFIP in the period spanning 1973 through to 2019.

The NFIP dataset specifies, for each claim, the amounts paid out in building insurance, contents insurance and the increased cost of compliance insurance.[86] The maximum amount that can be insured (and subsequently paid out) under each of these three elements is capped, with the applicable limits varying by residence type or building use.[87] For about 1% of both contents and building claims, the amounts paid out were equal to the applicable coverage limits. To account for this top-coding of the damages data, we assume the true damage values would follow a Pareto distribution, and pay-outs equal to the coverage limits were multiplied by 1.5.

For each individual claim, the resulting amounts for the three types of cover were summed and standardized to 2020 US dollars.[88]

The claim-level data was then collapsed by census tract and year, before being merged with a variable indicating each census tract's distance from the coast. This was calculated as the straight-line distance from each tract's geographical centroid to the nearest stretch of coastline (defined as elsewhere in our data), using ArcGIS. The resulting dataset therefore shows how total and annual average NFIP pay-outs vary at the census tract level, by distance to the coast.

Public subsidies to flood risk

Data on the amount of public spending associated with coastal flooding, which we use to calculate the share of damages subsidized by the taxpayer, was largely sourced from the Congressional Budget Office (CBO)'s 2019 report: Expected Costs of Damage from Hurricane Winds and Storm Related Flooding.[89]

The report's calculations are projections based on previous years' coastal storm-related spend-

[85] Data is available to download from OpenFEMA:
https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims.

[86] This covers the extra costs some may face, in terms of specific building requirements imposed by local authorities to reduce future flood risk.

[87] The specific coverage limits can be checked here:
https://emilms.fema.gov/IS1101b/groups/74.html.

[88] Standardisation to 2020 US dollars was done by applying deflator (Federal Reserve Bank of St. Louis, 2020), here:
https://fred.stlouisfed.org/series/GDPDEF.

[89] Report available here:
https://www.cbo.gov/system/files/2019-04/55019-ExpectedCostsFromWindStorm.pd.

ing, plus current conditions for climate, sea levels, and property development in places at risk of such storms. It breaks down the expected costs across three key sectors (public, residential, and commercial), as well as whether such costs were incurred specifically by flooding or hurricane winds. In terms of costs shouldered by the taxpayer, the report counts both ex-ante public spending (e.g. spending on preventative or mitigation measures such as building levees), and ex-post spending (e.g. FEMA's Public and Individual Assistance disaster relief programs), from across a range of US Government Departments.

In total, $19.4bn (USD 2020) of tax-payer money is spent on mitigation of and relief from the damages caused by hurricane winds and storm-related flooding. Over $12.7bn of this is spent on programs which assist the public sector, for example FEMA's Public Assistance program which finances the reconstruction and repair of public buildings, roads and other infrastructures. A further $5.5bn is spent on the residential sector, largely individual households requiring financial assistance from the state to rebuild their homes or businesses. Finally, smaller amounts correspond to the commercial sector or to administration. The vast majority of such spending, $18bn, comes from the federal budget.

Including the costs that are borne by the private sector, the total annual cost of storm-related flooding and winds is calculated at $57bn. This suggests that just over one third (19.4/57) of costs are effectively subsided by the state. Within the residential sector specifically, where losses are estimated at $36bn/annum, the $5.5bn of costs that are reimbursed by the state constitute a subsidy of just over 15%. This reflects the fact that a significant chunk of residential losses will be reimbursed by insurers, including the National Flood Insurance Program (NFIP), or are simply not eligible for state support and are therefore shouldered by residents.

References

Bleakley, Hoyt and Jeffrey Lin "Portage and Path Dependence" Quarterly Journal of Economics, Volume 127, Issue 2, May 2012, Pages 587–644

Dahl, Kristina A., Melanie F. Fitzpatrick and Erika Spanger-Siegfried. "Sea Level Rise Drives Increased Tidal Flooding Frequency at Tide Gauges along the U.S. East and Gulf Coasts: Projections for 2030 and 2045". PloS One, 12, no. 2 (2017).

Database of Global Administrative Boundaries (GADM). "USA Country Outline (Version 3.6)." Davis, CA: GADM (2018). Accessed September 2020 at
https://gadm.org/download_country_v3.html

ESRI (2021). "USA Parks Layer Package". Redlands, CA: ESRI. Accessed June 2020 at
https://www.arcgis.com/home/item.html?id=578968f975774d3fab79fe56c8c90941

Federal Emergency Management Agency (FEMA). "National Flood Insurance Program Redacted Claims – v1". Washington DC: FEMA (2021). Accessed July 2020 at
https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims-v1

Federal Reserve Bank of St. Louis Economic Data. "Gross Domestic Product: Implicit Price Deflator (GDPDEF)." St. Louis, MO: Federal Reserve Bank of St. Louis (2020). Accessed January 2021 at https://fred.stlouisfed.org/series/GDPDEF

Gyourko, J., Hartley, J., & Krimmel, J. (2019). The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index. (No. w26573). National Bureau of Economic Research.

Logan, John R., Zengwang Xu and Brian J. Stults. "Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database." The Professional Geographer, 66. no. 3 (2014): pp. 412-20.

Manson, Steven, Jonathan Schroeder, David Van Riper and Steven Ruggles. "IPUMS National Historical Geographic Information System: Version 14.0" [Database]. Minneapolis, MN: IPUMS (2019). Accessed February 2021 at: http://doi.org/10.18128/D050.V14.0

National Oceanic and Atmospheric Administration (NOAA) (2021b). "Sea Level Rise Viewer". Washington DC: NOAA (2021). Accessed June 2020 at
https://coast.noaa.gov/digitalcoast/tools/slr.html

National Oceanic and Atmospheric Administration (NOAA) (2021c). "Digital Coast". Washington DC: NOAA (2021). Accessed June 2020 at https://coast.noaa.gov/digitalcoast/data/

National Oceanic and Atmospheric Administration (NOAA) (2021d). "Coastal Change Analysis Program (C-CAP) Regional Land Cover". Charleston, SC: NOAA Office for Coastal Management (2021). Accessed September 2020 at

www.coast.noaa.gov/htdata/raster1/landcover/bulkdownload/30m_lc/

US Census Bureau. "Understanding "Place" in Census Bureau Data Products". Washington DC: US Census Bureau (2021). Accessed June 2020 at

https://www.census.gov/content/dam/Census/data/developers/understandingplace.pdf

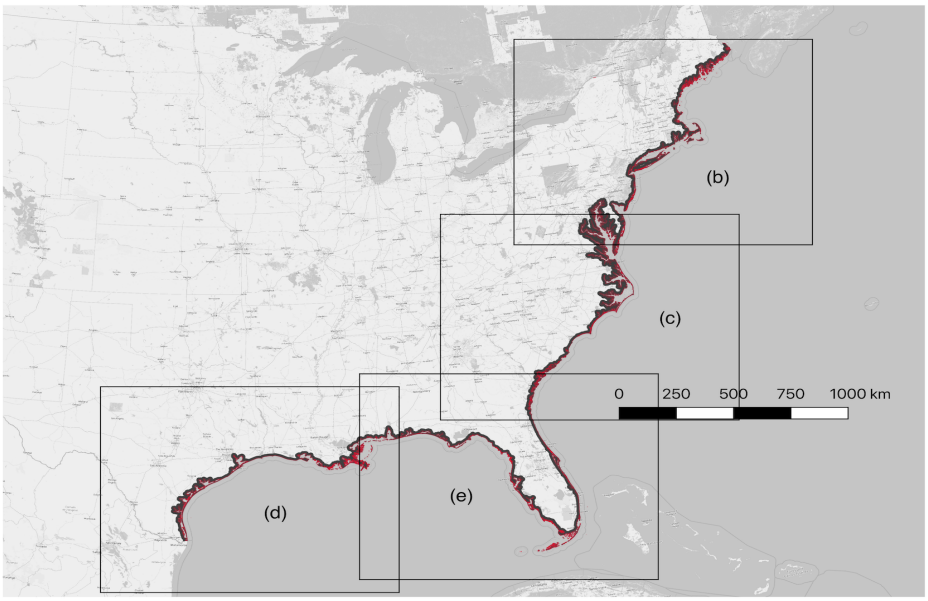
US Congressional Budget Office (CBO). "Expected Costs of Damage from Hurricane Winds and Storm Related Flooding". Washington DC: Congressional Budget Office (2019). Accessed July 2020 at:

https://www.cbo.gov/system/files/2019-04/55019-ExpectedCostsFromWindStorm.pdf

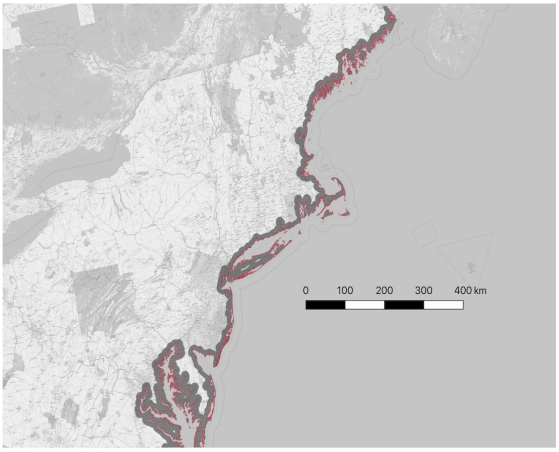
US Department of Transport (USDOT) – Bureau of Transportation Statistics (BTS). "Military Bases Dataset". Washington DC: USDOT. Accessed June 2020 at

http://osav-usdot.opendata.arcgis.com/datasets/d163fcde26de4d21aa06aa141ce3a662_0

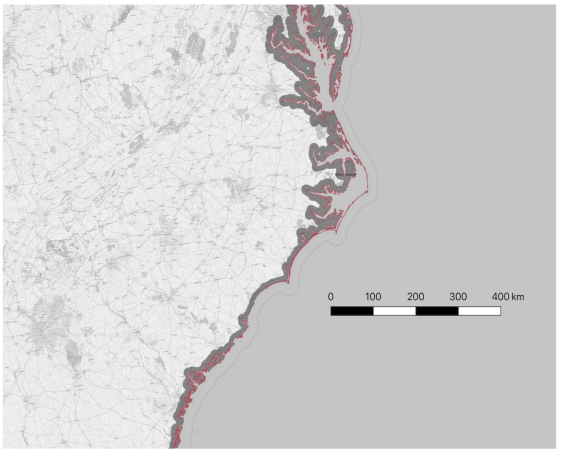
US Geological Survey (USGS) – Science Base. "Protected Areas Database of the United States (PAD-US)". Reston, VA: USGS (2017). Accessed June 2020 at
https://www.sciencebase.gov/catalog/item/5963ea3fe4b0d1f9f059d955.



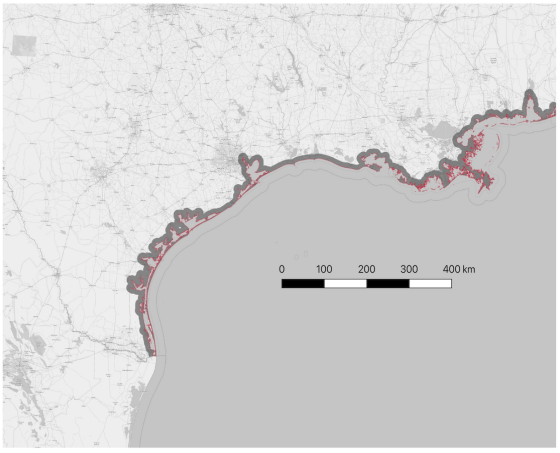
(a)



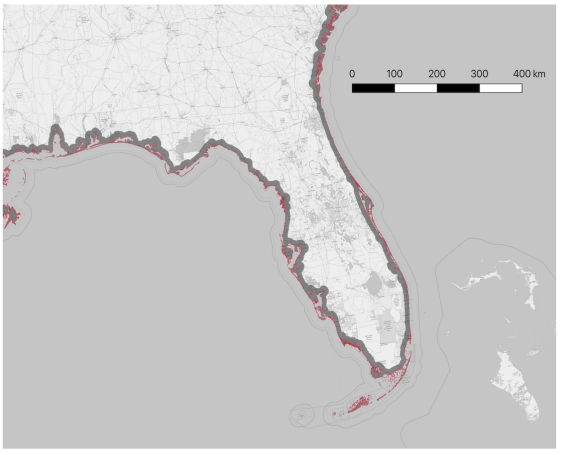
(b)



(c)



(d)



(e)

Figure A1: The figure shows the extent of our sample area, for the full US Atlantic and Gulf coasts in Panel (a), and split into sections in Panels (b) to (e). The coastline, as defined by the Database of Global Administrative Boundaries (GADM 2018) shapefiles, and used throughout our analysis, is highlighted in red, while the area included in our analysis, which extends 10km inland from the coast, is shaded in grey. The basemap here is from OpenStreetMaps, and is included in this figure for illustrative purposes, but not used elsewhere in our analysis. All data are projected using the Albers Equal Area projection.

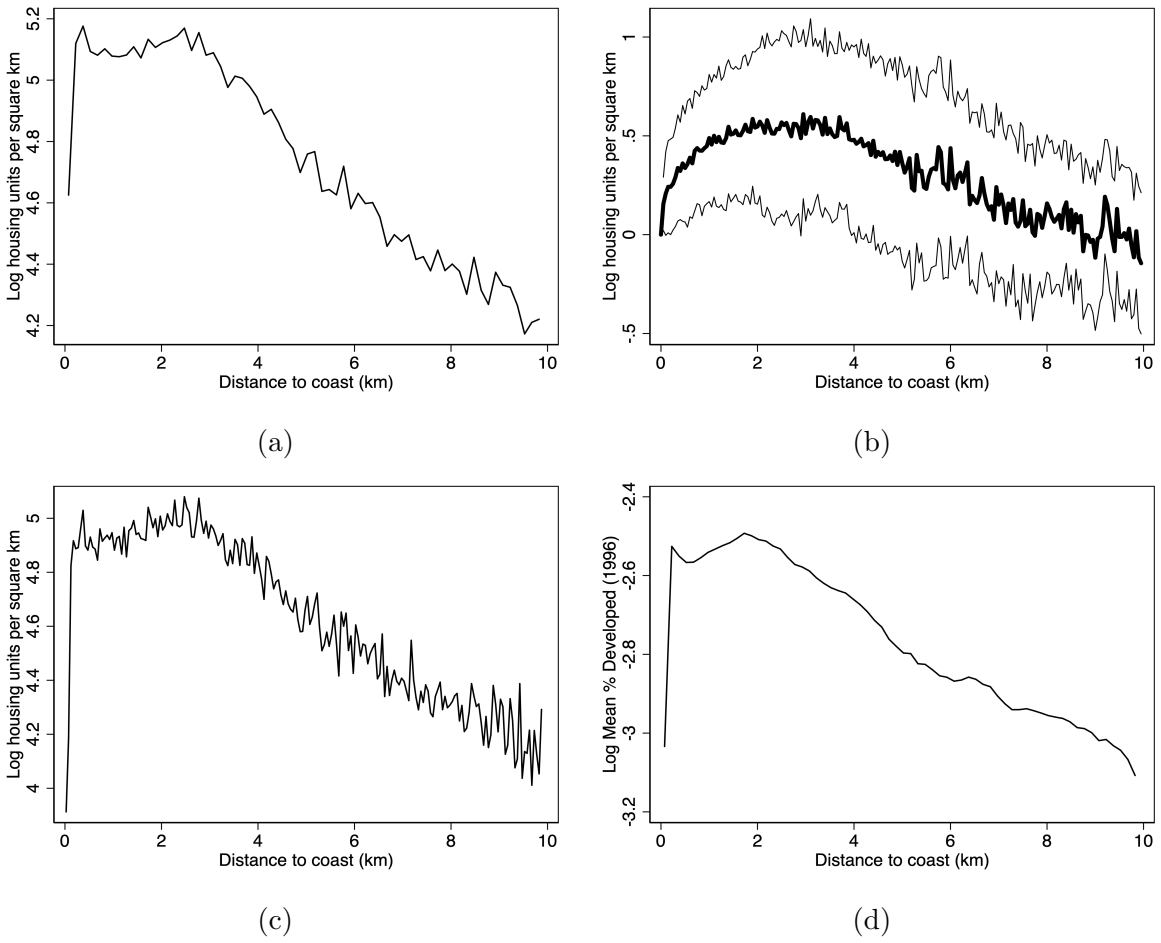


Figure A2: (Robustness to Stylized Fact 1) Population concentrates near – but not right at – the coast. Panel (a) shows log housing density (average number of housing units in 1990 per square kilometer) by distance to the coast for 150m bins, based on census data at the block level. Panel (b) shows regression coefficients and 95% confidence intervals from a regression of log housing density (number of housing units in 1990 per square kilometer) on 50m distance bins. Panels (a) and (b) here are exactly as per Figure 1 in the main text, with the exception that here the data exclude blocks (and gridcells) where 50% or more of block (or gridcell) area is restricted for development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. 18,863 out of 544,065 blocks and 841,719 out of 6,070,597 gridcells are excluded on the basis of this criterion. Exact definitions and data sources for restricted areas are detailed further in the Data Appendix. Panel (c) is the same as Fig1a in the main paper, except using 50m bins instead of 150m. Specifically, it shows log housing density (average number of housing units in 1990 per square kilometer) by distance to the coast for 50m bins, based on census data at the block level. Panel (d) shows average % gridcell developed (in logs) based on the Landsat data, as observed in 1996 - the earliest year for which we have complete Landsat data.

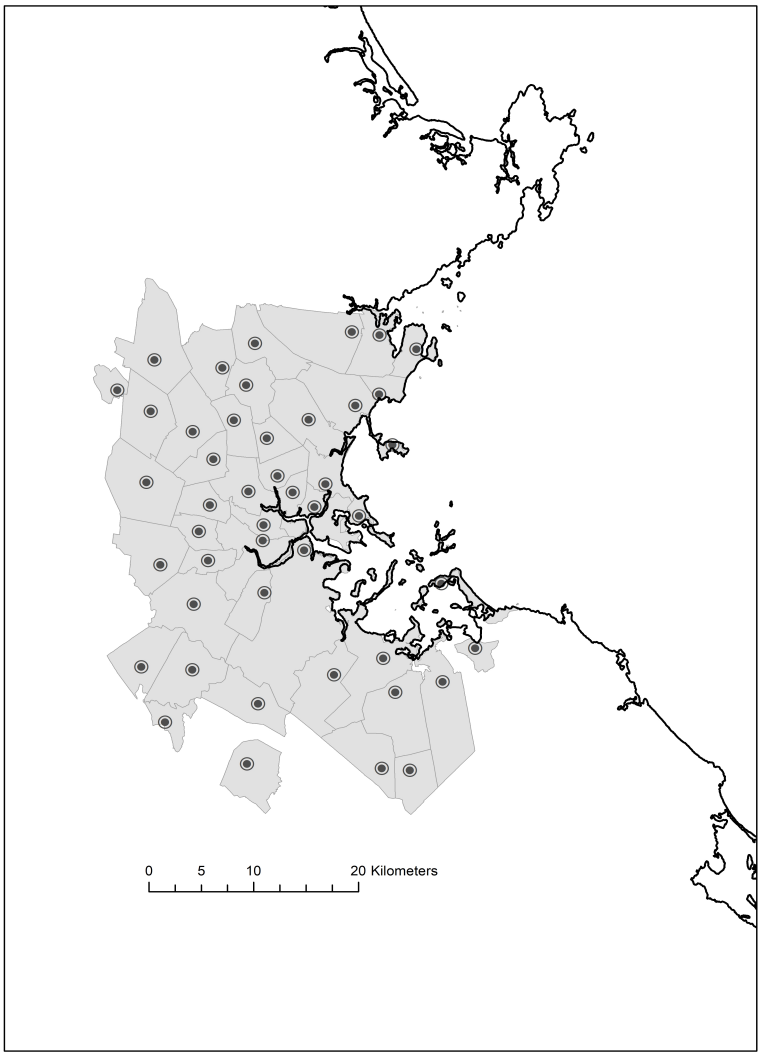
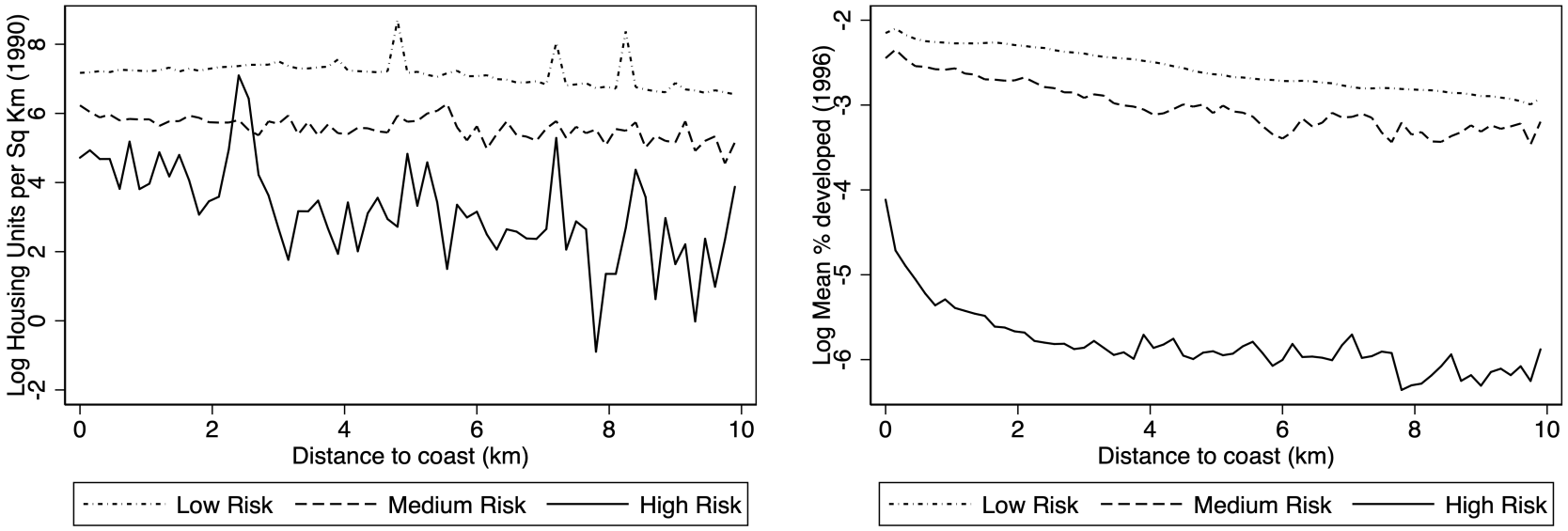
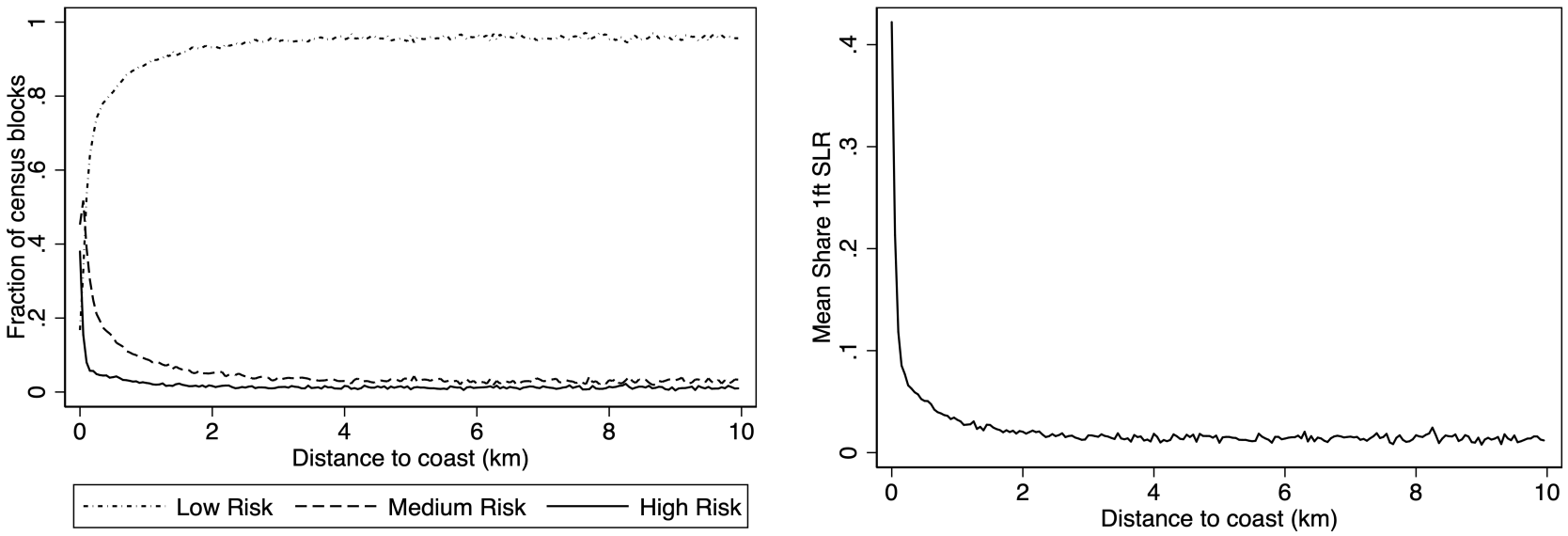


Figure A3: (Example of Stylized Fact 2) Places near the coast are asymmetric. The figure shows the CBD of the city of Boston plus a number of Census Designated Places (CDPs) whose own CBD is within 25km of Boston's. The polygons of each CDP, as well as the points denoting the location of their CBD, are taken from NHGIS (more detail on how the NHGIS determines each place's CBD are provided in the Data Appendix). The coastline, as per the GADM shapefiles used throughout our paper, is also shown.



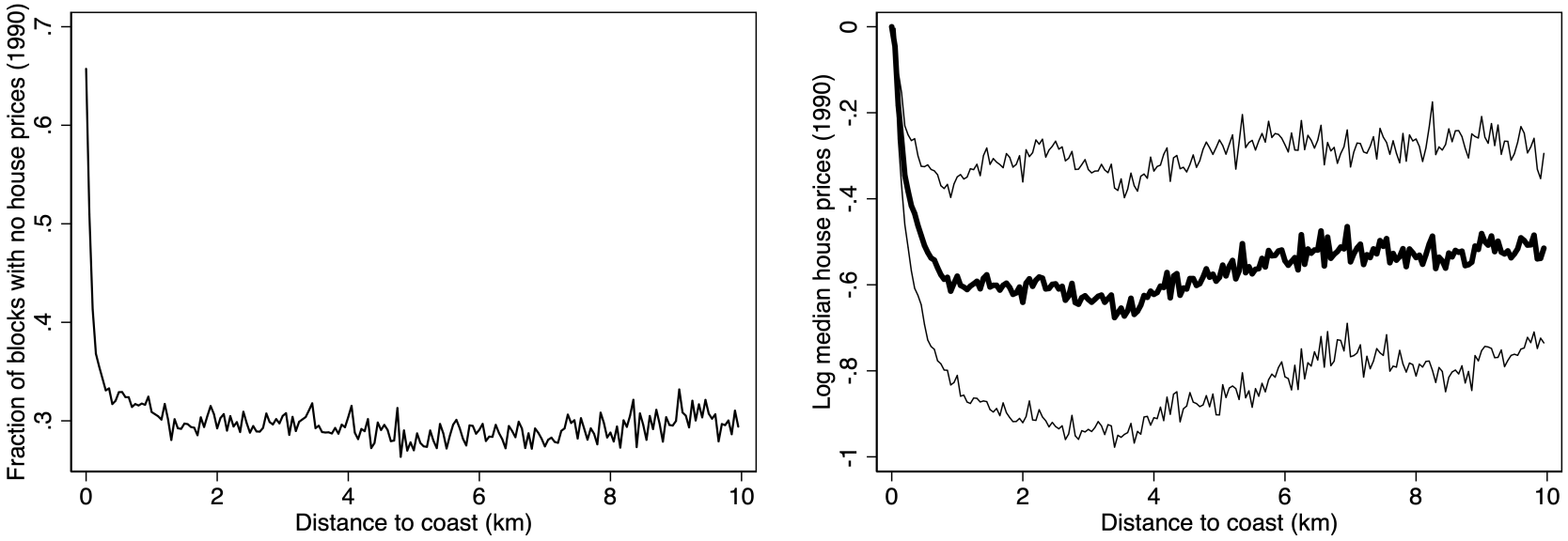
(a) (b)

Figure A4: (Robustness to Stylized fact 4 excluding restricted areas) Areas that are highly prone to sea level rise (SLR) are less built. Panel (a) replicates Figure 2 in the main paper, and shows housing units per square kilometer in 1990 (based on census data at the block level) by distance to the coast in kilometers and share of area under water with 1 foot of sea level rise. The three risk categories are defined by the share of each census block that will be under water at high tide if sea levels rise by 1 foot (30.4cm): we label blocks as if the share of 1ft SLR is > 0.5, as where $0 <$ share 1ft SLR $<= 0.5$, and as where share of 1ft SLR $= 0$. This reflects odds of flooding even today, without any SLR. As panel (a) shows, at each distance from the coast, the riskier areas are more sparsely built. In this case the data exclude blocks where 50% or more of block area is restricted for development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. 18,863 out of 544,065 blocks are excluded on the basis of this criterion. Exact definitions and data sources for restricted areas are detailed further in the Data Appendix. Panel (b) shows the average % of gridcells classified as developed in 1996 (in logs), based on Landsat data, using the same risk and distance categories as in (a).



(a) (b)

Figure A5: (Robustness to Stylized fact 5 excluding restricted areas) Flood risk helps explain why people do not build right on the coast. Panel (a) replicates Figure 3a in the main paper, and shows the fraction of census blocks in each risk category by distance to the coast in 50m bins. The three risk categories are defined by the share of each census block that will be under water at high tide if sea levels rise by 1 foot (30.4cm): we label blocks as high risk if the share of 1ft SLR is > 0.5, as medium risk where $0 <$ share 1ft SLR $<= 0.5$, and as low risk where share of 1ft SLR $= 0$. This risk reflects odds of flooding even today, without any SLR. Panel (b) replicates Figure 3b in the main paper and shows the mean share of block area that is subject to 1ft SLR, by distance to the coast in 50m bins. In each case the data used here exclude blocks where 50% or more of block area is restricted for development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. 18,863 out of 544,065 blocks are excluded on the basis of this criterion. Exact definitions and data sources for restricted areas are detailed further in the Data Appendix.



(a) (b)

Figure A6: (Analysis of house prices in proximity to the coast) Panel (a) shows the fraction of blocks in our data with no information on house prices in 1990, by 50m distance bins to the coast. Panel (b) shows cofficients and 95% confidence intervals from a regression of log median house price at the block level on indicators for 50m distance bins from the coast, for blocks where 1990 house price information was available.

Table A1: Robustness to stylized fact 8 excluding restricted areas

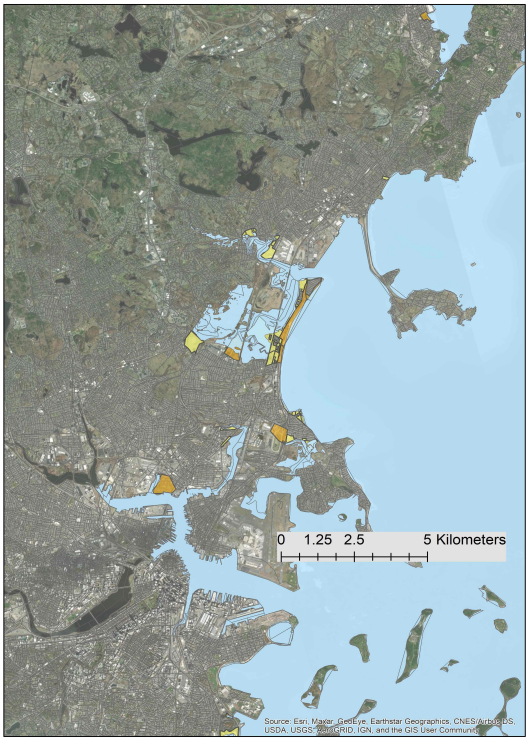
	(1)	(2)	(3)	(4)
Housing units per sq km in containing census tracts	<=10	(10, 100]	(100, 1000]	>1000
Share 1ft SLR	-2.95	-3.33	2.00	7.69
	(2.33)	(1.57)	(1.10)	(3.90)
Constant	3.97	7.79	5.77	4.53
	(1.08)	(1.31)	(0.77)	(1.05)
Observations	15,811	130,510	286,525	92,354

Notes: The outcome in each case is the change in housing units at the block level from 1990-2010. Columns divide the data by levels of housing units per square km in census tracts (1990), excluding own block. Standard errors in parentheses are clustered by state. Urban population density: at least 386 people (not housing units) per square km. Results robust to controlling for log distance to the coast. In each case the data exclude blocks where 50% or more of block area is restricted for development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. Exact definitions and data sources for protected areas are detailed further in the Data Appendix.

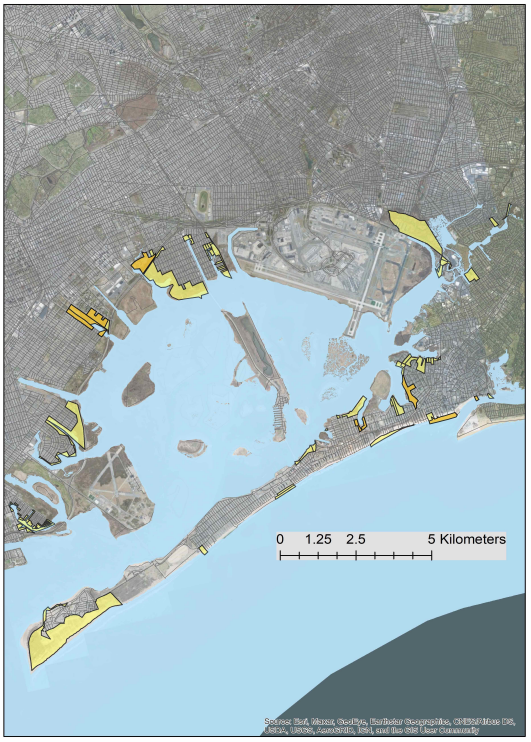
Table A2: Robustness to stylized fact 8 using the Landsat data

	(1)	(2)	(3)	(4)	(5)
Average % of gridcell developed in 1996 in 1km square neighbourhoods	<6.25	$[6.25, 12.5)$	$[12.5, 25)$	$[25, 50)$	>=50
Share 1ft SLR	-5.16	-26.61	-26.93	-15.43	2.44
	(1.42)	(4.64)	(3.91)	(3.95)	(7.14)
Constant	4.89	24.65	26.73	20.01	8.87
	(1.33)	(4.25)	(3.67)	(2.97)	(2.50)
Observations	4,608,714	382,674	428,990	436,677	156,070

Notes: The outcome in each case is the change in the % gridcell classified as developed from 1996-2010, based on Landsat data. Columns divide the data by levels of average % gridcell developed (1996) in 1km square neighbourhoods, excluding own gridcell. Standard errors in parentheses are clustered by state.



(a) Revere and Chelsea, MA



(b) Jamaica Bay and Rockaway, NY

(c) Miami Beach and Miami, FL

(d) Clearwater and Largo, FL

Figure A7: This figure shows four case studies, which illustrate development on the fringes of dense tracts. Specifically, the figures highlight census blocks with share 1ft SLR> 0, and with net development from 1990 to 2010 of up to 50 housing units (in yellow) and greater than 50 housing units (in orange).

Table A3: Robustness to stylized fact 9 excluding restricted areas

	(1)	(2)
Share 1ft SLR ∈(0.0,0.1)	13.99	
	(9.15)	
Share 1ft SLR ∈ [0.1,0.2)	14.97	
	(5.38)	
Share 1ft SLR ∈ [0.2,0.3)	8.45	
	(3.31)	
Share 1ft SLR ∈ [0.3,0.4)	15.14	
	(5.93)	
Share 1ft SLR ∈ [0.4,0.5)	9.92	
	(5.92)	
Share 1ft SLR ∈ [0.5,0.6)	7.35	
	(3.69)	
Share 1ft SLR ∈ [0.6,0.7)	-0.32	
	(2.43)	
Share 1ft SLR ∈ [0.7,0.8)	0.99	
	(3.21)	
Share 1ft SLR ∈ [0.8,0.9)	-3.10	
	(1.64)	
Share 1ft SLR ∈ [0.9,1.0)	-2.72	
	(2.02)	
Share 1ft SLR = 1	-3.83	
	(1.04)	
Some SLR		15.77
		(6.69)
Share 1ft SLR		-18.72
		(8.57)
Constant	4.14	4.14
	(1.17)	(1.17)
Observations	92,354	92,354

Notes: The outcome in each case is the change in housing units at the block level from 1990-2010. The sample here is restricted to blocks with tract housing density >1000 housing units per square km. The omitted category, Share 1ft SLR =0, accounts for 95.5% of the blocks with this level of tract housing density in 1990. Standard errors in parentheses are clustered by state. In column (2), the variable Some SLR is an indicator for blocks that have share 1ft SLR >0, while Share 1ft SLR is a continuous measure of the share of each block prone to 1ft SLR. The data here exclude blocks where 50% or more of block area is restricted for development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. Exact definitions and data sources for restricted areas are detailed further in the Data Appendix.

Table A4: Robustness to stylized fact 9 using the Landsat data

	(1)	(2)
Share 1ft SLR $\in$ (0.0,0.1)	8.74	
	(2.28)	
Share 1ft SLR $\in$ [0.1,0.2)	3.32	
	(2.82)	
Share 1ft SLR $\in$ [0.2,0.3)	2.79	
	(4.05)	
Share 1ft SLR $\in$ [0.3,0.4)	4.58	
	(3.61)	
Share 1ft SLR $\in$ [0.4,0.5)	0.93	
	(5.59)	
Share 1ft SLR $\in$ [0.5,0.6)	3.11	
	(6.00)	
Share 1ft SLR $\in$ [0.6,0.7)	4.45	
	(7.44)	
Share 1ft SLR $\in$ [0.7,0.8)	1.09	
	(5.94)	
Share 1ft SLR $\in$ [0.8,0.9)	9.66	
	(8.14)	
Share 1ft SLR $\in$ [0.9,1.0)	-0.97	
	(4.61)	
Share 1ft SLR = 1	-4.26	
	(3.59)	
Some SLR		6.98
		(2.45)
Share 1ft SLR		-7.84
		(4.02)
Constant	8.59	8.59
	(2.55)	(2.55)
Observations	156,070	156,070

Notes: The outcome in each case is the is the change in the % gridcell classified as developed from 1996-2010, based on Landsat data. The sample here is restricted to gridcells in 1 km square neighbourhoods with average % gridcell developed of at least 50%. Standard errors in parentheses are clustered by state. In column (2), the variable Some SLR is an indicator for blocks that have share 1ft SLR >0, while Share 1ft SLR is a continuous measure of the share of each block prone to 1ft SLR.

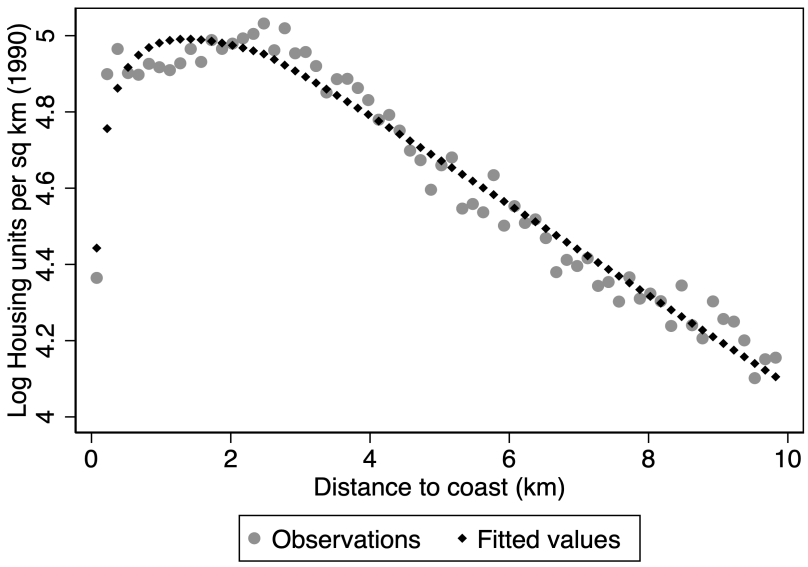


Figure A8: Estimating the thetas. The figure shows observations and fitted values for log housing density by distance to the coast (as per Figure 1a). The fitted values are based on estimating equation (17), which we use to estimate θ_1, θ_2 and U, as discussed in the main text, and detailed in Table A5.

Table A5: Parameter estimates

Parameter	Value	Source	Notes
α	0.75	Davis & Ortalo-Magne (2011)	Consumption share of income
σ	0.3454 (0.0637)	Estimated	Estimated by equation (16), using data from Fig4 (Robust standard error)
θ_1	0.1419 (0.0049)	Estimated	Estimated by equation (17), using data from Fig1a (Robust standard error)
θ_2	5763.27 (672.73)	Estimated	Estimated by equation (17), using data from Fig1a (Robust standard error)
U	3892.92 (61.10)	Estimated	Estimated by equation (17), using data from Fig1a (Robust standard error)
w	536580	US Bureau of the Census (1991)	Median income in 1990 (in 2020 USD per decade)
commuting speed	43.1	Couture et al. (2018)	Average commuting speed in 1995 (in km per hour)
hours worked	1872	OECD (2021)	Average hours worked per year in 1990
commuting distance	500	see notes	Total distance per km of commute (in km per year)
x_0	1.2	Model-generated	Location of the model-generated CBD (in km from the coast)
x_{Lt}	0.538	Observed	Location of coast-side edge of the synthetic city (in km from the coast)
Elevation of CBD	2	Hallegatte et al. (2013)	Population-weighted median elevation of Miami (in m)
SLR rate	0.0577	Portner et al. (2019)	Average of high and low emissions scenarios to 2100 (in m per decade)
wage growth	1%	By assumption	Wage growth per decade
δ	0.82	Treasury (2021)	Discount rate per decade, based on 30-year interest rate on last day of 2020
p_A	see notes	Estimated	Agricultural prices, equated to housing prices at x_L in 1990 (static model)
$\tilde{s}$	see notes	CBO (2019), FEMA (2021)	Public subsidy to flood damages

Notes: The estimate for σ is from a regression of NFIP damages per housing unit on distance to the coast, as per equation (16) and Figure 4. Estimates for θ_1, θ_2 and U are based on estimating equation (17), using data on housing density (as per Figure 1a). The estimation is joint for the three parameters, and only determines two of them relative to the third one. Commuting distance is the distance in km travelled per year per km of commute, and is obtained by assuming 250 working days per year, with return commmutes. x_0 is the location of the model-generated CBD in km from the coast. x_L is the mean distance from the coast to the coast-side edge of places that are 2-3km from the coast. The baseline rate of sea level rise that we assume is 0.0577m per decade, which is halfway between the two main IPCC scenarios of 43cm and 84cm by 2100 (Portner et al. 2019), assuming a linear increase over time. For a CBD at 2m elevation, this implies a rate of advance of the coastline of 34.9m per decade. We also consider a faster SLR scenario, corresponding either to faster local SLR, or to a city whose CBD elevation is 1.33m above sea level. The discount rate δ is based on the 30-year interest rate on the last day of 2020, which was 1.65 percent (Treasury 2021). We convert this to a decennial version, which we round to 18%. This implies a discount rate per decade of 0.82. The agricultural price p_A is equated to housing price at x_L in 1990 for the static model. In the dynamic model we invert out agricultural price as the constant price that makes developers indifferent between converting land from agricultural to residential use and not converting at x_L in 1990, as detailed in the paper. To calculate the subsidy, we start with the figure of \$19.4bn in public spending on storms and floods per year (CBO 2019). We allocate 36% of this value to our study area, based on NFIP share (FEMA 2021). We adjust back to 1990 levels based on historical growth of 3.8% per year in NFIP claims in our study area, and allocate value per housing unit based on the estimates presented in Figure 4.

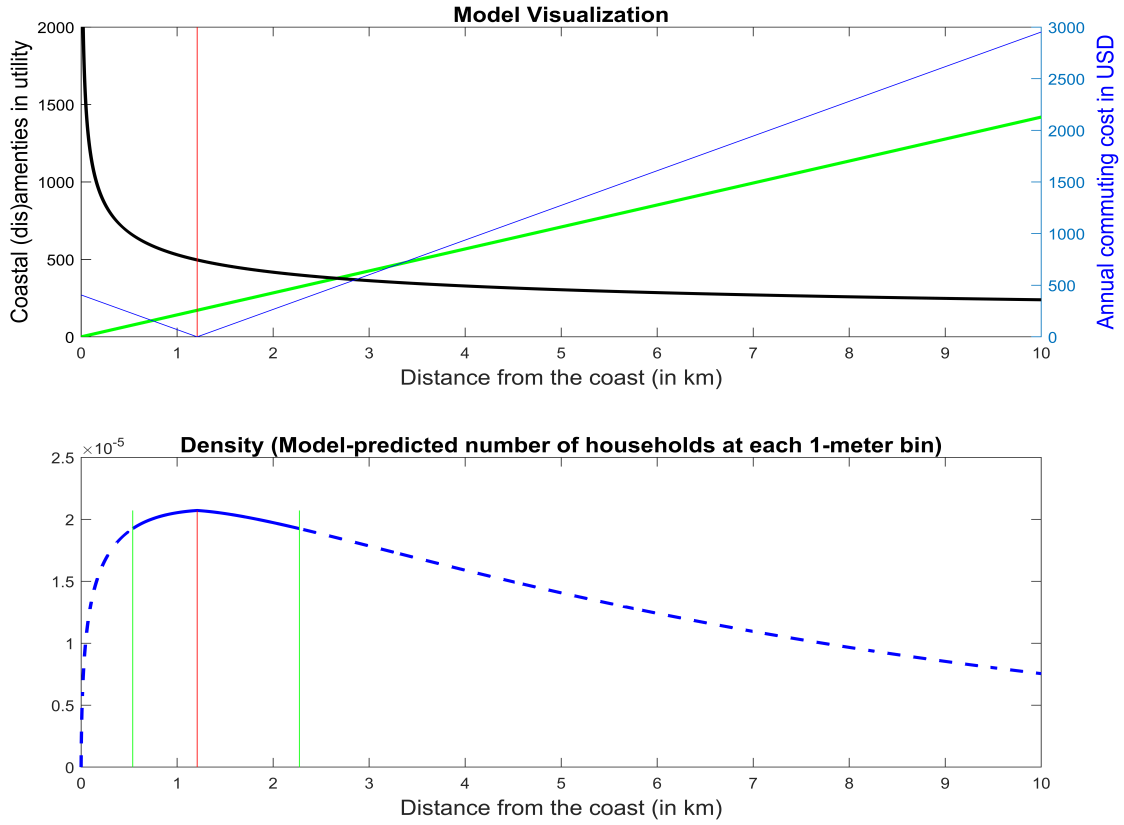


Figure A9: This figure summarizes aspects of the simulated model for the year 1990. The top panel of this figure is a visualization of the simulation from the residents' perspective. It shows the linear decline in coastal amenity (in green) and the convex decline in flooding disamenity (in black) as we move away from the coast, both on the left-hand vertical axis, with the marginal effect of both equating at the CBD. It also shows the commuting costs rising in distance to the CBD (in blue, on the right hand vertical axis). The bottom panel shows the housing density, with the solid blue line denoting actual density and the dashed blue line showing what density would have been, were it not for the city boundaries. The city boundaries themselves are denoted by two vertical green lines.

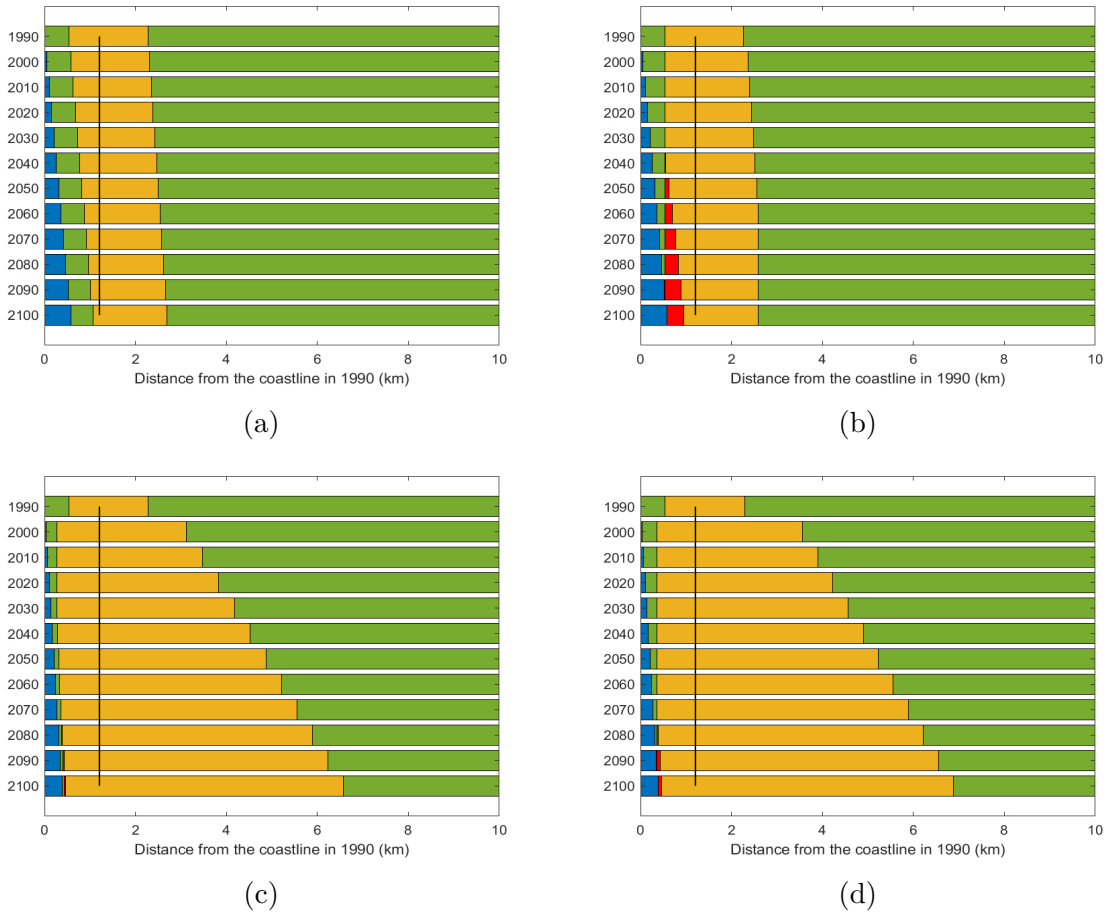


Figure A10: (Additional simulation results) The figure shows results of simulations, as described in the main text, and based on parameter values detailed in Table A5. Panel (a) shows results from the static model with fast SLR and without wage growth. Panel (b) is the same as (a) but for the dynamic version of the model. Panel (c) shows results from the static model with 1% wage growth per decade, baseline SLR and adding a subsidy on flood damages. Panel (d) is the same as (c) but for the dynamic version of the model. The value of the subsidy is detailed in Table A5. In each panel of the figure yellow denotes the city, green denotes agricultural land, blue denotes the sea, and red denotes areas whose housing density declined at least 10 percent from the maximum level. The time period is on the vertical axis in 10 year time-steps, and the horizontal axis shows distance to the coast in km.