# Cities and the Sea Level*

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

January 25, 2024

**Abstract**

Construction on low elevation coastal zones is risky for both residents and taxpayers who bail them out. To investigate this construction, we analyze spatially disaggregated data covering the entire US Atlantic and Gulf coasts. We find that the 1990 housing stock reflects historical avoidance of locations prone to sea level rise (SLR) and flooding, but net new construction from 1990-2010 was similar in SLR-prone locations and safer ones; and within densely built coastal areas, net new construction was higher in SLR-prone locations. These findings are difficult to rationalize as mere products of moral hazard or imperfect information, suggesting that people build on risky locations to benefit from nearby urban agglomerations. To explain our findings, we develop a simple model of a monocentric coastal city, which we use to explore the consequences of sea level rise. This model helps explain cities' role in expanding flood risks, and how future sea level rise may reshape coastal cities, creating significant challenges for policymakers.

KEYWORDS: Cities, Climate Change, Sea Level Rise.

JEL CLASSIFICATION: R11, Q54, R14

*Corresponding author: Michaels (g.michaels@lse.ac.uk). We thank Juan Alvarez-Vilanova and Tiernan Evans for excellent research assistance. We thank Emek Basker, Tim Besley, David Castells-Quintana, Peter Christensen, Vernon Henderson, Alan Manning, Henry Overman, Steve Pischke, Eric Strobl, and John Van Reenen, and seminar/conference participants at Bank of Spain, CERIS Workshop Galway 2023, CESifo Area Conference on Global Economy, ERSA 2018, Federal Reserve Bank of Philadelphia, IFAU, IZA, LSE, NUI Galway, Queen Mary University of London, UEA North America Meetings 2022, University of Kent, UAB, UCSB, University of Stockholm, Universitat Pompeu Fabra, 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

Low Elevation Coastal Zones (LECZ) are attractive places to live. In 2000 they were home to around 10% of the world's population, a figure expected to grow significantly by 2050.[1] But this growing attraction of LECZ poses challenges, as some of their terrain is susceptible to floods. The problem of flooding is expected to worsen as glaciers melt and the oceans warm and expand due to climate change. 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. Sea level rise (SLR) is affecting some LECZ more than others, and the US Atlantic and Gulf coasts suffer from some of the fastest rates of local SLR in the world (Dahl et al. 2017). Along with SLR, climate change may also increase the severity of tropical storms (Berardelli 2019), whose impact is already acutely felt in the US. Since 2005, the US has suffered 173 major weather and climate disasters, most of which were caused by tropical storms, other severe storms, and flooding (NOAA 2021a).

The severity of major flooding events, such as tropical storms, creates an important role for the government, which deals with catastrophic events that private markets do not fully insure (e.g., CBO 2017, Pralle 2018, and Bakkensen and Barrage 2021). Recent government estimates suggest that roughly a third of the annual costs imposed by tropical storms in the US is accounted for by public funds (US Congressional Budget Office 2019), with taxpayers' burden further increased by the costs of social insurance responses to storms (Deryugina 2017). And the costs of flooding are rising over time. For example, overall payouts on the National Flood Insurance Program (NFIP) have increased at a rate of roughly 4% per year in real terms since the 1970s, while a series of official reports have highlighted a widening gap between claims and premia.[2]

Since both residents and taxpayers are exposed to rising flood risk, our paper asks: does new coastal construction avoid risky locations? To study whether this is the case, we use up-to-date high-resolution maps (NOAA 2021b), which identify SLR-proneness consistently, using a fine vertical and spatial breakdown.[3] This lets us pinpoint locations that will be under water at high tide if sea levels rise by 1 foot, or approximately 30.5 cm, and are even today highly prone to flooding (Dahl et al. 2017 suggest that many areas on US Atlantic and Gulf coasts could experience 1ft SLR as early as 2045). To measure outcomes, we use data on housing units from the Census and American Community Survey from 1990-2010; specifically, we use the finest level of census data available - census blocks. By virtue of their small size, these blocks allow us to identify housing that is at risk of SLR in ways that are impossible using coarser data, such as for counties.

These data reveal two new stylized facts. First, while the density of coastal construction up to 1990 was negatively correlated with SLR (and flood) risk, construction from 1990-2010 was, on average, uncorrelated with such risk. Quantitatively, in 1990, 12% of the coastal housing stock was in blocks prone to one-foot sea level rise. But for (net) new housing built from 1990-2010, this share was

[1] These estimates are from Neumann et al. (2015), who 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.

[2] Estimates of NFIP cost growth are by the authors, described in detail in Appendix B. The widening gap between claims and premia are discussed, for example, in CBO (2017 and 2019) and Bakkensen and Barrage (2021).

[3] In contrast, commonly used data from Federal Emergency Management Agency were not consistently updated across areas. They also identify locations as prone to "1 in T years flood", which is harder to interpret, especially with climate change.

26%, bringing the share of the stock in 2010 to 14%. Second, in areas that were sparsely built in 1990, net new construction from 1990-2010 still avoided SLR-prone locations; but in areas that were densely built in 1990, net new construction was positively correlated with SLR risk. We note however, that even within the densest census tracts, new construction focused on medium-risk SLR-prone areas, still avoiding the riskiest ones.

To paint a richer picture of coastal construction, we use cross-sectional data from around 1990 to show four auxiliary findings. 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. Fourth, 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.[4] Our empirical findings are robust to controlling for Core-Based Statistical Area (CBSA) fixed effects, which address potential concerns about variation in local sea level rise and zoning regulations. They are also robust to controlling for distance to the coast (except where that distance is the main variable of interest) and to excluding census blocks, which were mostly shielded from private residential construction, because they are either protected areas, military bases, or parks. We find very similar patterns using satellite-derived data on built area, which cover all types of construction rather than just housing.

Our empirical findings are difficult to reconcile using only the two common explanations for building on flood-prone areas: moral hazard due to government subsidy, or imperfect information (or understanding). These factors may be important, but they do not explain why SLR-prone areas are typically avoided and built on only when land is scarce. Our findings suggest that a third and previously unexplored mechanism is at work: people build on SLR-prone coastal locations to benefit from nearby urban agglomerations.

To account for our empirical findings, we develop a simple 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.[5] The city founder chooses a location that trades off these two factors – close to the coast, but not right at it. This location becomes the city's focal point – the Central Business District (CBD).[6] 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 building on increasingly flood-prone land.

After explaining how our empirical observations are accounted for in the model, we extend the model in several ways. Our first extension allows for sea level rise. The second allows for a limited set of high-elevation locations near the coast, which are safe from flooding and command high

[4] In a related study, Wing et al. (2018) use a flood hazard model to study changing flood risk at a disaggregated spatial level.

[5] 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).

[6] The dynamics we study, where a city forms around a historically determined location, which remains a focal point even when fundamentals change, echoes the findings of Bleakley and Lin (2012) on path-dependent city locations.

prices. Third, we allow for costly and irreversible conversion of land to housing from alternative uses, which makes the developers' decisions dynamic rather than static. Fourth, we examine government subsidies to flood-prone areas. Fifth, we consider the possibility that employment is spread across the city, rather than being concentrated solely in the CBD. Finally, we consider the impact of land use regulations.

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 imposes additional costs beyond increased flood risk, by further distorting the shape of LECZ cities, which significantly lengthens the time costs of commuting to work. Finally, these cities face a potential crisis if their CBD comes under threat of being permanently submerged.

We view our model primarily as a qualitative tool for evaluating coastal development. Our model considers a relatively benign case, where people correctly anticipate events and have time to adapt the city gradually. Further, the model does not consider idiosyncratic conditions, which specific cities may face. With these caveats in mind, our welfare analysis sheds light on the costs of sea level rise and of government subsidies to SLR-prone areas, both of which concentrate in the 1km closest to the coast. At the same time, we find that path dependence in the precise location of the CBD entails relatively lower costs, at least as long as the CBD is not submerged.

The main contributions of our paper are threefold. 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 study the distribution of the coastal housing stock and show that new construction in recent decades focused in SLR-prone areas, especially near densely populated areas. Our findings are novel and policy relevant. Finally, we develop a simple model, which provides a parsimonious explanation for our findings. We extend this model and use it to study how SLR may reshape coastal cities and consider the costs of SLR and subsidies for construction in SLR-prone coastal areas.

The literature on flood risk and housing markets has tended to concentrate on estimating price effects for properties exposed to flood risk (for a review see Beltran et al. 2018). Relatively less attention has been paid to quantities, and most previous empirical studies of exposure to SLR risk have tended to be coarser in their spatial resolution (e.g., Burby 2001, Brody et al. 2007, and Dinan 2017). In recent work, Barrage and Furst (2019) analyze the relationship between new housing additions and exposure to sea level rise for US coastal counties. Our empirical analysis takes a much more fine-grained

view, enabling us to present a more nuanced picture of how exposure to flood risk is evolving along the US Atlantic and Gulf coasts. A related literature focuses on policy aspects of managing flood risk, including the relationship between land-use, regulation, and damages from flooding (e.g., Kousky et al. 2013; Taylor and Druckenmiller 2022), as well as institutional aspects of flood insurance in the US (e.g., Kriesel and Landry 2004; Michel-Kerjan 2010; Kousky and Michel-Kerjan 2017).

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 defenses. 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).[7] Our paper posits 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 (in other contexts, different environmental hazards, such as areas prone to wildfires, may play a similar role). 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. 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 the 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.

A recent literature quantifies the economic cost of climate change using structural models. For example, Balboni (2020) studies exposure of Vietnam roads to SLR and finds that infrastructure investments that ignore future SLR risks might lead to inefficient persistence in coastal cities, and Desmet et al. (2021) use a spatially disaggregated, dynamic model of the world economy to quantify the roles of migration and local agglomeration for population dynamics and SLR cost. While our model is more stylized, it opens a window into the previously unexplored internal structure of coastal cities and their adjustment to climate change, offering a parsimonious explanation for our novel findings.

Finally, our paper is also related to the literature on 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.

[7] Ortega and Taṣpınar (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.

# 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 our focus on flood-prone and SLR-prone LECZ and the analysis of fine spatial units. 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 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. 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. We also make use of more aggregated geographic units, including census-designated places (administrative cities or towns, which may make up part of a metropolitan area), 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).[8]

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.[9] A map of the area that we study is shown in Appendix Figure A1.[10] 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 Housing and land cover data

Our housing data come from the US Census and American Community Survey, observed in 1990 and 2010, which cover all housing unit types.[11] We harmonize all our census data to the geographical

[8] To characterize the shape of coastal construction, we also obtain data on census-designated places (Manson et al. 2019). These data are useful because they show not only the outline of places, but also (unlike metropolitan areas) their historical CBDs. We use these to calculate place asymmetry, as explained in the Data Appendix.

[9] 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.

[10] 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.

[11] What we refer to as "2010" is more precisely data for 2006 - 2010 from the American Community Survey.

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 544,071 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).[12] We complement the housing quantity data with house price data from the census, as we discuss in the appendix.

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 2021c). 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.[13] In our analysis, we focus on the four developed categories, which represent different extents of constructed surfaces (including buildings, roads, and parking lots).[14] We aggregate the Landsat data to 150m x 150m cells (the 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.[15]

### 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).[16] The maps we use 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 (e.g., the SLR maps assume that New Orleans is safe even from 6-foot of SLR.).

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)

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

[13] 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.

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

[15] 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.

[16] While these maps reflect the state of understanding about SLR towards the end of our study period, is consistent with our goal of explaining construction in recent decades on SLR-prone area assuming that people are individually rational and forward looking. Imperfect knowledge and understanding may further exacerbate the costs of SLR in the real world.

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 of 1ft SLR for the entire sample of blocks is 0.046, and the area-weighted mean share is around 0.19.

### 2.4 Building restrictions data

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. In the Data Appendix we discuss three different types of regulations: restricted areas, where housing development may be particularly constrained, and other zoning data; state "setback lines" close to the coast, beyond which construction may be more regulated; and local government regulations on building density.

### 2.5 Government subsidies and additional data

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.[17] 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. While claims made under the NFIP do not reflect 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 - and all cost data - to 2020 US dollars using a GDP price deflator (Federal Reserve Bank of St. Louis 2020). We then aggregate the damages data across the entire period available (1973-2019). To obtain a measure of damage per housing unit, we divide the total damages by the number of

[17] In 2012 the Biggert-Waters Flood Insurance Reform Act was introduced in an attempt to phase out subsidies on the NFIP and bring the program towards fiscal solvency. These reforms proved controversial, particularly with respect to the impact on homeowners facing large increases in flood risk premiums. The 2014 Homeowners Flood Insurance Affordability Act partially repealed and modified the Reform Act (Bakkensen and Barrage 2021).

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.

Government subsidies to flood-prone areas also come in the form of constructing and maintaining flood defenses. The SLR data we use account for existing flood defenses (major federal leveed areas) and assume that these remain protected under any SLR scenario. Initiatives to build major new flood defenses raise concerns about costs to taxpayers, time to build, effectiveness as sea levels rise, and potential environmental damage (See for example a discussion of possible flood defense schemes for New York City: https://www.nytimes.com/2020/01/17/nyregion/the-119-billion-sea-wall-that-could-defend-new-york-or-not.html). Additional data sources used for our model simulation are detailed in Appendix Table A8, and we discuss the parameter estimates themselves in Section 4.5.1.

# 3 Empirical findings

This section documents two novel stylized facts about changes over time in coastal housing risk, and four auxiliary findings about the cross-section of coastal housing; we focus on the area which lies within 10 km of the US Atlantic and Gulf coasts, as discussed in Section 2. First, we present the two stylized facts, which are the focus of our study. Second, to set these in context, we describe three auxiliary findings on the location of coastal housing in the cross-section. Third, we show one additional finding, pertaining to the mechanisms that underlie coastal housing construction. Finally, we return to the two stylized facts and discuss their robustness.

## 3.1 Two stylized facts on changes in the risk of coastal housing

The first stylized fact is that **while historic construction near the coast avoided SLR-prone locations, more recent construction has not**. This is shown in Table 1. In 1990, areas with medium or high SLR risk accounted for about 12% of the housing units in our area of study. But from 1990-2010, 26% of net new construction took place in medium or high-risk blocks. Consequently, the fraction of the 2010 housing stock on SLR-prone locations was 14%.

The second stylized fact tells us where the risky new developments took place: **Recent construction in SLR-prone areas took place in dense locations, but not in sparse ones**. Table 2 reports regression estimates using the specification:

$$\Delta hunits_i = \beta_{11} + \beta_{12} SLR1ft_i + \epsilon_{1i}, \tag{1}$$

where $\Delta hunits_i$ is the change in the number of housing units in census block $i$, $SLR1ft_i$ is 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, and $\epsilon_{1i}$ is an error term, which is clustered by Core-based statistical area (CBSA) here

and in all the spatial regressions we report below.[18] 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. The regressions in Table 2 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. All this suggests that where there is plenty of space to build, SLR-prone locations are avoided, in line with the evidence discussed above; SLR-prone locations are, however, built on in dense areas, presumably because no other local alternatives exist.

We highlight these two stylized facts because of their importance for understanding the growing risk from SLR and flooding in coastal areas from 1990-2010. Should these trends in coastal construction continue, the costs of flooding will rise rapidly in the coming decades.

To put these stylized facts into context, however, it is useful to look back at construction in coastal areas looked in 1990. With that objective in mind, we now proceed to characterize four auxiliary findings on the cross-section of coastal housing development.

### 3.2 Auxiliary findings on the cross-section of coastal housing

The first auxiliary finding 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.[19] 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.[20] Specifically, housing density plummets close to the coast and declines more slowly on the inland side. 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

[18] CBSAs include metropolitan statistical areas and micropolitan areas, and we add one cluster for all non-CBSA observations in the sample. In earlier versions of the paper, we obtained similar standard error estimates when clustering by state, an approach used by Donaldson and Hornbeck (2016) and others. We also 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.

[19] 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.

[20] The housing density is similar in its peak and in nearby distance bins, and it displays some geographic variation. For example, in the US South, housing density peaks closer to the coast, consistent with a higher amenity value of the beach. But in each subsample that we examined, density falls steeply very close to the coast and more gradually further away from it.

from estimating the regression:

$$\ln(hdensity_i) = \beta_{21} + \beta_{22}\mathbf{Bin}_i + \epsilon_{2i}, \tag{2}$$

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 $\epsilon_{12}$ is an error term. 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 Data Section).[21] The results are largely unchanged.[22] 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 we use 50-meter bins, and the decline in density near the coast is even steeper. 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 live throughout most of the year. Specifically, only around 1 percent of the housing units in our sample are second homes.[23] 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.[24]

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

$$asymmetry_j = \beta_{31} + \beta_{32}\mathbf{Bin}_j + \epsilon_{3j}. \tag{3}$$

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

[21] 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.

[22] 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).

[23] 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.

[24] 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.

from the coast; and $\epsilon_{3j}$ is an error term.[25] As Table 3 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 in the Greater Boston area: those close to the coast are asymmetric, while those further away are more symmetric.

The third auxiliary finding 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:

$$asymmetry_j = \beta_{41} + \beta_{42} size_j + \beta_{43} \ln\left(dist\_coast_j\right) + \beta_{44}\left[size_j \cdot \ln\left(dist\_coast_j\right)\right] + \epsilon_{4j}, \tag{4}$$

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; $\textbf{dist\_coast}_j$ is the mean distance from each place's blocks to the coast; and $\epsilon_{4i}$ is an error term. The estimates in columns (1) and (2) of Table 4, which add the restriction $\beta_{43} = \beta_{44} = 0$, 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.

### 3.3 Auxiliary findings on mechanisms that shape the coastal housing distribution

Whereas the three auxiliary findings above tell us how economic activity concentrates near but not right at the coast, here we present evidence about why this is the case. Our fourth auxiliary finding is that **blocks prone to SLR are much more sparsely built on, and among the SLR-prone blocks, those further from the coast are even more sparsely built.** To examine how much SLR-prone blocks 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. 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 risk group, especially the for the two riskier groups, housing density tends to increase as we approach the coast. In other words, people seem to understand SLR risk and avoid it, even as they value proximity to the coast. This is consistent with the observation that as we approach the coast, SLR-proneness rises steeply. As Panel (a) of Figure 3 shows, the proportion of low-risk blocks is consistently over 90% in the 3-10km area from the coast. However, as we get closer to the coast within the three km range, this proportion declines rapidly to less than 20%, while the proportion of medium-risk and high-risk blocks increases significantly. 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%

[25] 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.

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, is offset by a convex disamenity due to flood risk as we near the coast.

To see why this matters, we demonstrate the steep rise in damages from flooding as approach the coastal areas in Figure 4, which shows the point estimates and 95% confidence intervals from the regression:

$$\ln\left(damage_k\right) = \beta_{51} + \beta_{52}\mathbf{Bin}_k + \epsilon_{5k}, \tag{5}$$

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 $\epsilon_{5k}$ is an error term.[26] 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.[27]

Having characterized the four auxiliary findings, 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 below in the extensions, by noting that while locations in blocks close to the coast are generally flood-prone and therefore in low demand, there may be small higher-elevation areas within these blocks, where flooding is much less of a problem, and where prices are high.

[26] 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.

[27] As we discuss in the Data Appendix, NFIP costs cover only a fraction of total damages from flooding. The low uptake of NFIP flood insurance, as well as the presence of demand frictions and selection biases, may introduce measurement discrepancies. Another concern is the SFHA designation, particularly for inland areas. Despite these limitations, we utilize NFIP data in this study because it provides spatially disaggregated information at the census tract level. Excluding restricted areas at this level of analysis is unnecessary due to the larger scale of census tracts, which primarily focus on built areas.

Returning to our first four auxiliary findings, we note that Auxiliary Finding 4 helps explain Auxiliary Findings 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.

### 3.4 Robustness of the stylized facts on changes in coastal housing risk

Whereas the four auxiliary findings describe coastal area housing at a point in time, mostly around 1990, our two (main) stylized facts describe how they changed from 1990-2010. Appendix Table A1 reports regression estimates from specification (1) for the full sample, using as outcomes 1990 housing units and the change in housing units from 1990-2010. Columns (1) and (2) of Panel A show that 1990 housing units are strongly negatively correlated with $SLR1ft$, but for 1990-2010 this correlation is close to zero. Subsequent columns show that these relationships are robust to controlling for CBSA fixed effects and restricting the sample to only urban (CBSA-located) blocks, and Panel B shows estimates controlling for distance bins to the coast, as well as (again) CBSA fixed effects. The final columns of Panel B, with a full set of controls show a strongly negative correlation of construction and SLR risk in 1990, and a positive though imprecise relationship for changes from 1990-2010.

Next, we revisit the baseline estimates of Stylized Fact 2 as reported in Table 2. We now add CBSA fixed effects and distance bins to the coast, and restrict the sample to CBSAs, and report combinations thereof. Appendix Table A2 shows that the results are robust: in sparse census tracts, areas prone to SLR were avoided from 1990-2010, while the opposite was true in dense census tracts. In Appendix Table A3 we return to the specifications estimated in Table 2, but this time excluding restricted areas, and the pattern is again robust. Finally, in Appendix Table A4 we repeat the analysis 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 sparse "neighborhoods" new construction was strongly negatively correlated with SLR risk, while in the densest "neighborhoods" the correlation is positive though imprecisely estimated. This imprecision may arise because it is harder to detect the densest locations using the satellite imagery data, which only measure whether each pixel is built and not how densely it is built.

To further investigate the role of amenity in shaping Stylized Fact 2, we use winter weather as a demand shifter (Rappaport 2007) and examine whether in areas that grew faster due to good weather there was a difference in construction on SLR-prone locations between dense and sparse census tracts. Specifically, we estimate regressions of the type:

$$\Delta hunits_i = \beta_{61} + \beta_{62} SLR1ft_i + \beta_{63} Mildwinter_i + \beta_{64} Mildwinter_i * SLR1ft_i + \epsilon_{6i}, \quad (6)$$

where $Mildwinter_i$ measures higher January temperatures or an index of mild winters more generally (see details in Appendix Table A5) and $\epsilon_{6i}$ is an error term. Columns (1) and (2) of Appendix Table A5 confirm that locations with mild winters experienced faster growth in the number of housing units from 1990-2010, consistent with Rappaport (2007). More importantly, the rest of the table shows that this growth was uneven: in sparse census tracts, the interaction of mild weather and

$SLR1ft$ is weakly negative, while in dense census tracts it is strongly positive. This suggests that growing demand for housing translates into risky construction locations only where housing is already densely built, and few alternatives remain. To illustrate the construction locations in dense, SLR-prone areas from 1990-2010, Appendix Figure A7 shows four case studies of developments 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. Construction in the two areas in Florida is particularly pronounced, as we can expect from the findings in Appendix Table A2.

Finally, we show in an extension to Stylized Fact 2 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 (2), 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.[28] 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 A6 shows that these results are robust to excluding restricted areas. Finally, Appendix Table A7 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

In this section, we introduce a model of coastal development that reconciles the stylized facts and auxiliary findings discussed earlier. The model is parsimonious and designed primarily to build intuition, provide qualitative insights, and explore counterfactual scenarios. We begin this section by outlining the model's assumptions. We then characterize the equilibrium and relate it to our empirical findings. Finally, we extend the model to consider several counterfactual scenarios, and cautiously explore potential welfare implications.

## 4.1 Baseline assumptions

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. 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.[29] 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.

[28] In line with Stylized Fact 1, we note that in dense locations, blocks with low shares of SLR saw larger increases in housing unit growth than those with no SLR (captured by the omitted category).

[29] Later, when we explore SLR, we relax this assumption by allowing the initial location of the coast to shift inland over time.

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

$$U^F(x) = -\theta_1 x - \theta_2 x^{-\sigma}. \tag{7}$$

As we discuss below, to rationalize our empirical findings, 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.[30] For the same reason, we also assume $\theta_2 > 0$ and $\sigma > 0$, reflecting a convex disamenity (higher risk of flooding).[31] As discussed later, the pattern of housing density increasing and then decreasing as we move inland from the coast aligns 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, $x_0$.[32]

There is a continuum $[0, \overline{x}]$ of competitive and forward-looking developers, each of whom owns a plot of land of measure 1 in location $x$; we assume that $\overline{x} > 0$ is sufficiently high not to constrain the land side development. 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 $p_A$.[33] 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, we assume that there is a continuum of perfectly mobile residents with sufficient mass to populate the city. In every period $t = 1, .., T$, each resident may live in the city or outside it. If a resident lives 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:

$$U(c_t(x), h_t(x), x) = c_t(x)^{\alpha} h_t(x)^{1-\alpha} - \theta_1 x - \theta_2 x^{-\sigma} \tag{8}$$

where $c_t(x)$ and $h_t(x)$ denote private consumption goods and housing in period $t$ and location $x$, and $\alpha \in (0,1)$ is the consumption share of income. 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:

$$p_t(x) h_t(x) + c_t(x) = w_t - |x - x_0| \tag{9}$$

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

[30] 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.

[31] To keep the model simple and allow for future extensions, we omit risk aversion, although our specification can be seen as a reduced-form way of capturing risk aversion. Moreover, the disamenity included in our model accounts for indirect effects of flooding on soil suitability for housing and the impact of wind gusts.

[32] 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.

[33] We follow the literature by labelling non-housing use as agriculture. 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).

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

$$x_0 = \left(\frac{\sigma\theta_2}{\theta_1}\right)^{\frac{1}{\sigma+1}}. \tag{10}$$

We note that while our formulation of the model emphasizes a tradeoff between the risk and reward of locating near the coast, other historical factors relating to proximity to an agricultural hinterland and access to national and international markets may have also played a role in determining the historical locations of CBDs. Our analysis focuses on cities for which this balance meant that the CBD location is close to the coast.

**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 build in all locations such that

$$p_t(x) \geq p_A. \tag{11}$$

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

$$p_t(x_{Rt}) = p_t(x_{Lt}) = p_A. \tag{12}$$

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.

To complete the description of the equilibrium, we note that in each period, supply and demand for housing determine the price of land in each location, $p_t(x)$ and the city's population, $pop_t = \int_{x_{Lt}}^{x_{Rt}} \frac{1}{h_t(\hat{x})} d\hat{x}$.

### 4.3 Relating the model to the empirical findings

We now discuss how the model accounts for the two stylized facts and four auxiliary findings that we documented. We begin with Auxiliary finding 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, Auxiliary findings 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 Auxiliary finding 1, 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:*

$$\textit{For each period } t = 1, ..., T\textit{: if } x < x_0 \textit{ then } \frac{\partial \ln\left(dens_t(x)\right)}{\partial x} > 0\textit{; if } x > x_0 \textit{ then } \frac{\partial \ln\left(dens_t(x)\right)}{\partial x} < 0. \quad (13)$$

**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". Proximity to the CBD grants individuals smaller, high-value land plots, while locations closer to the city's outskirts offer larger, more affordable land. And indeed, as our first auxiliary finding 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. Nonetheless, we still need to address the issue of higher prices observed in sparsely populated areas within 0-1 km of the coast. We revisit this below where we consider a limited number of elevated coastal locations.

Another characteristic of coastal areas, as highlighted in Auxiliary finding 2, is their asymmetry, with the CBD situated closer to the coast-side edge than the inland edge. We show that this is the case in the model.

**Lemma 1** *The city develops asymmetrically around the CBD:* $|x_{Rt} - x_0| > |x_0 - x_{Lt}|$.

**Proof.** See appendix. ■

Our final static empirical result, Auxiliary finding 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. This can be seen if we consider the minimal wage required to sustain the city, $\widetilde{w}$. As the wage falls to this minimum level and the city becomes very small, it is no longer asymmetric.

**Lemma 2** *Vanishingly small cities are symmetric:* $\lim_{w_t \searrow \widetilde{w}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$.

**Proof.** See appendix. ■

Turning to the first of our stylized facts, we assume that the 1ft SLR area is closest to the coast, in a range $[0, D]$, where $D < x_0$. Historically, wages were low and cities we small, so the SLR-prone area is avoided. But as a cities grew, construction began to encroach on the SLR-prone area.

**Lemma 3** *If $w_t$ is sufficiently low, the city avoids SLR-prone areas ($x_{L0} > D$). But if $w_t$ increases sufficiently, the city eventually expands into 1ft SLR areas ($x_{L0} < D$).*

**Proof.** See appendix. ■

This result leads to a different perspective on the geographic constraints of the city than Saiz (2010) and Harari (2020), in whose work 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.

Finally, we show that SLR-prone areas are developed in densely built locations, but not in sparse ones.

**Lemma 4** *As long as the wage is low, the city is small and sparsely populated, SLR-prone locations are not developed; but as the wage increases sufficiently, population density rises, and SLR-prone locations are developed.*

**Proof.** Follows immediately from Lemma 3. ■

This provides an intuitive explanation to Stylized Fact 2, where expansion into flood-prone areas occurs in densely built locations.[34]

### 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; government subsidies to partly offset the disamenity of proximity to the coast; multiple employment centers within the city; and land use regulations.

#### 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

$$U(c_t(x), h_t(x), x) = c_t(x)^{\alpha} h_t(x)^{1-\alpha} - \theta_1 (x - x_{ct}) - \theta_2 (x - x_{ct})^{-\sigma}. \tag{14}$$

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 that once a location becomes submerged due to sea level rise, it becomes uninhabitable. Sea level rise also affects the attractiveness of non-submerged areas. Locations on the coast-side of the CBD become less appealing because the costs of the disamenity rise at a faster rate than the benefits from the amenity, as observed in the city founder's problem. On the other hand, locations further inland may experience temporary benefits from sea level rise, as the amenity value of proximity to the coast outweighs the increased disamenity from flooding. Assuming a fixed CBD, this may lead to even greater asymmetry in coastal cities and result in higher average commuting costs. The issue of "misshapen cities" (Harari 2020) may be further aggravated by SLR.[35]

[34] This result does not prove that the share built on SLR-prone locations increases monotonically.

[35] Considering what happens to the city when its CBD is under water is beyond the scope of this paper.

#### 4.4.2 Irreversible housing construction

The analysis thus far assumes a simplified scenario where developers can switch land use between agriculture and housing at no cost, making each period independent. To relax this assumption, we introduce a time cost associated with switching, creating a "dynamic" version of the model in contrast to the static baseline. To maintain tractability, we assume that converting land from agriculture to housing incurs a cost equal to one period's price and is irreversible (while conversion from housing to agriculture is infinitely costly). Developers solve their problem by comparing the present discounted value of prices across all time periods until $T$ (or until the plot becomes submerged with sea level rise). So the developer's problem in city location $x$ and period $t = 1, ..., T$ is:

$$\underset{t=1,...,T}{Max} \left\{ \sum_{s=t}^{T} \delta^s p_A, \underset{\widehat{s}=t,...,T}{Max} \left[ \sum_{t \leq s < \widehat{s}} \delta^s p_A + \sum_{T \geq s > \widehat{s}} \delta^s p_s(x) \right] \right\}. \tag{15}$$

This condition replaces expression (12). 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 hypothesize that barriers to increasing housing density as the city expands contribute to the expansion of the extensive margin towards the coast and amplify the distortion caused by sea level rise.

#### 4.4.3 Government subsidies

While the model we use is deterministic (for analytical simplicity), uncertainty plays a role in the lives of coastal dwellers. While in principle this leaves room for private insurance markets, in practice governments are usually heavily involved in providing flood insurance (including re-insurance), since the shocks that households suffer are not idiosyncratic, but correlated over large areas. In providing such insurance, governments often end up subsidizing coastal residents. This may happen, for example, because of outdated flood maps, which underestimate rising flood risks, as well as ex-post bailouts. Indeed, as we discuss in the data section, there is evidence that the US government subsidizes coastal development.

Here we explore allocating government subsidy in proportion to the losses we document in Auxiliary finding 6 (in proportion to the fitted values from equation (19) below). This allows us to explore how subsidies affect city development, and how they interact with SLR.

#### 4.4.4 Land use regulations

We consider two types of land use regulation: the first is aimed at reducing construction on flood-prone areas, while the second restricts land use within safer areas of the city (See for example Gyourko et al. 2019 for a discussion of local land regulations). To model regulations aimed at restricting

construction in flood-prone areas, we consider them as a tax on housing, which decreases in the distance to the coast. The simplest case is a partial or full offset to any government subsidy (see above), which requires no further elaboration. Next, we consider the case of regulations that restrict housing supply in safer parts of the city. A typical policy is a green belt placed at the inland edge of the city. In an open city model where there are no migration frictions across cities, the attractiveness of building near the coast, will not be affected by the green belt. But in a closed-city version (e.g., Duranton and Puga 2014), a green belt increases demand for housing in the sections of the city where construction is allowed, resulting in development further towards the coast. To illustrate this point, we revise our model by assuming exogenous population growth over time in a closed city model. Both wages and rents adjust to ensure that the equilibrium population, which is the aggregation of housing density over all developed land parcels, grows exogenously at a rate of 5% or 10% per period, similar to the growth rates in the open city model. The value of agricultural land is still pre-determined and identical to that of the baseline model.[36]

#### 4.4.5 Additional extensions

In the appendix we consider two further extensions. One involves a finite number of elevated locations near the coast, which are safe and hence highly desirable; the other considers a case where employment is dispersed throughout the city, instead of being concentrated in a CBD.

### 4.5 Simulations

#### 4.5.1 Parameter estimation

Below 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. In the appendix we explain the choice of parameters we use to estimate the model under different scenarios. These parameters are reported in Appendix Table A8.

#### 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 empirical findings that we discuss above: in

[36] In recent and related work Ospital (2023) studies the effects of excessive regulation on wildfire risk in California.

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 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.[37] 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.[38] 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), at 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.

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

[38] 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.

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 next two panels of Figure A10 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 land 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.

The bottom two panels of Figure A10, (e) and (f), consider the case with an alternative commuting elasticity of 0.0729 as per Duranton and Puga (2019). This lower elasticity accounts for the presence of multiple employment locations within the city, reflecting the idea that people may find jobs in various areas beyond the CBD and do not always have to commute to the center. Here the city expands more than in the baseline, although the cost of this expansion is lower since people's commutes are shorter.

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 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.[39]

Finally, Figure A11 considers the case of a closed city with restrictive regulation. In Panel (a) we simulate decadal population growth of 5%, without a greenbelt restriction, and in panel (b) we add a green belt extending inland from the 1990 inland edge of the city. Panels (c) and (d) repeat the analysis with a decadal population growth of 10%. In both cases, the green belt pushes population further into the flood-prone area near the coast.

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.

[39] 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.

#### 4.5.3 Welfare analysis

The main rationale for the model is to provide a parsimonious explanation for the stylized facts and auxiliary findings, and qualitatively explore different scenarios, as we do above. Our model considers a relatively benign case, where people correctly anticipate events, and the city adapts gradually. Further, the model does not consider idiosyncratic conditions that specific cities may face. We also note that time discounting crucially affects all our estimates, and in many cases the losses increase over time. With these caveats in mind, we cautiously proceed to explore some of the model's welfare implications. We note that in the model residents' utility is fixed, so we measure welfare using the value of landowners' land, net of any transfers.

In Columns (1) and (2) of Appendix Table A9, we compute the percentage changes in the present discounted value of land, between the scenario without SLR and each of the two SLR scenarios (baseline and fast SLR). In the city as a whole, the losses from SLR are around 1.5-2.4 percentage points, with higher losses with fast SLR and in the dynamic case. Land within 1km from the coast, however, bears the brunt of these losses, falling in value by around 20-30 percent due to SLR. The intuition for this result is that SLR submerges land near the coast and changes the payoffs for inland locations, so those near the coast lose out, while those further inland (where the marginal coastal amenity is higher than its marginal disamenity) may gain, at least as long as they are not submerged.

We next consider the economic loss from government subsidies, which encourage the city's (over) expansion, especially the coastal side. Column (3) of Table A9 shows that even without SLR, the net loss from subsidies in the city as a whole is around 1 percent in the static and 2 percent in the dynamic case. The loss from the subsidy is again unevenly distributed, concentrating mostly in the 1km near the coast, where it is around 6-7.5 percentage points. Columns (4) and (5) show a similar loss from the subsidy with SLR, since the city adjusts by expanding inland. We note, however, that this loss comes on top of the loss from SLR itself, as discussed above. As discussed above, the model does not account for risk aversion and the role of flood insurance that may be difficult to obtain privately. At the same time, we note that the ratio of the subsidy to (land values without subsidy) is higher than the welfare loss, amounting to 5.3-6.4 percent for the city as a whole and 15-20 percent within 1km of the coast, most of which is capitalized into land values.

Finally, we examine the role of an immobile (path dependent) CBD, which affects the city's shape and the commuting cost. In Columns (6) and (7) of Table A9, we report the percentage change in cumulative land value in the model where the CBD is fixed in the first period relative to a model where the CBD adjusts costlessly each period. The loss here (0.16-0.25 percentage points) is small compared to the previous cases. An important caveat to this finding is that it considers only cases where the CBD is not inundated by SLR.

#### 4.5.4 Policy implications

Governments could enact various policies to mitigate the problems discussed above. First, to limit taxpayer exposure, governments could consider taxing new developments in flood-prone areas, if there are viable alternative uses to the land, which are not taxed. The difference between the dynamic scenarios (where extensive margin adjustments are costly) and the static scenarios suggest that with SLR, raising the costs of extensive margin development restricts it to some extent.

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 (see https://www.floodre.co.uk/can-flood-re-help-me/eligibility-criteria/). 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.

Third, if withdrawing subsidies is unfeasible, 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 (e.g., https://www.ft.com/content/f95aa4e2-b3e6-11e7-aa26-bb002965bce8). Or governments could restrict the number of times a given property is bailed out, or offer other incentives to move instead of rebuilding, as Canada has recently done (e.g., https://www.nytimes.com/2019/09/10/climate/canada-flood-homes-buyout.html). 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 (e.g., https://www.ft.com/content/5a463614-c7e4-11e9-af46-b09e8bfe60c0)

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 two stylized facts and four auxiliary findings. 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. 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


Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, Nikolaus Wolf. 2015. "The Economics of Density: Evidence from the Berlin Wall". Econometrica 83(6): 2127-2189.

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).

Bakkensen, Laura A. and Lint Barrage, Going Underwater? Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics, *Review of Financial Studies*, 2021

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

Barrage, Lint and Jacob Furst. 2019. "Housing investment, sea level rise, and climate change beliefs," *Economics Letters*, 177: 105-108.

Beltran, A., D. Maddison, and R. Elliott (2018) "Is flood risk capitalised into property values?" *Ecological Economics*, 146:668-685.

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/

Bester, C. Alan, Timothy G. Conley and Christian B. Hansen. 2011. "Inference with Dependent Data Using Cluster Covariance Estimators". *Journal of Econometrics*, 165(2): 137-151.

Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence" *Quarterly Journal of Economics* 127(2): 587–644

Brody, S.D., Zahran, S., Maghelal, P., Grover, H. and Highfield, W.E., 2007. "The rising costs of floods: Examining the impact of planning and development decisions on property damage in Florida." *Journal of the American Planning Association*, 73(3):330-345.

Burby, Raymond J. 2001. "Flood Insurance and Floodplain Management: The US Experience." Environmental Hazards, 3(3): 111–22.

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. 1999. "GMM estimation with cross sectional dependence". *Journal of Econometrics*, 92(1): 1-45.

Couture, Victor, Gilles Duranton and Matthew A. Turner. 2018. "Speed". *Review of Economics and Statistics*. 100(4): 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. 2011. "Household Expenditures, Wages, Rents". *Review of Economic Dynamics*. 14(2): 248-261.

Dinan, Terry. 2017. "Projected Increases in Hurricane Damage in the United States: The Role of Climate Change and Coastal Development," *Ecological Economics*, 138: 186-198,

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* 131(2): 799–858

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

Duranton, Gilles and Diego Puga. 2019. "Urban Growth and its Aggregate Implications." NBER Working Paper 26591.

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.

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." National Bureau of Economic Research wp 26573.

Hallegatte, Stephane, Colin Green, Robert J. Nicholls and Jan Corfee-Morlot. 2013. "Future flood losses in major coastal cities," *Nature Climate Change*. 3(9): 802-806.

Harari, Mariaflavia. 2020. "Cities in Bad Shape: Urban Geometry in India". American Economic Review 110(8): 2377-2421.

Henderson, J Vernon, Tanner Regan and Anthony J. Venable. 2020. "Building the City: from Slums to a Modern Metropolis". *Review of Economic Studies* 88(3): 1157–1192.

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*, 107, no. 6 (2017): 1365-1398.

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

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

Kousky, Carolyn, Sheila M. Olmstead, Margaret A. Walls, and Molly Macauley. 2013. "Strategically Placing Green Infrastructure: Cost-Effective Land Conservation in the Floodplain." *Environmental Science and Technology* 47 (8): 3563–70.

Kydland, Finn E. and Edward C. Prescott. 1977. "Rules rather than discretion: The inconsistency of optimal plans". *Journal of Political Economy*, 85 (3): 473-491.

Magontier, Pierre, Albert Solé-Ollé and Elisabet Viladecans-Marsal. 2019. "The political economy of coastal destruction". Mimeo.

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

Marcy, Doug, 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. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area". *American Economic Review Papers and Proceedings* 57(2): 197-210.

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

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. 2015. "Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment". PloS one, 10(3).

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

Ospital A. 2023. "Urban Policy and Spatial Exposure to Environmental Risk", Mimeo.

Pralle, S. "Drawing lines: FEMA and the politics of mapping flood zones." *Climatic Change* 152, 227–237 (2019).

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. 2019. "IPCC Special Report on the Ocean and Cryosphere in a Changing Climate". Geneva, Switzerland: United National Inter-Governmental Panel on Climate Change.

Pryce, Gwilym, Yu Chen & George Galster. 2011. "The Impact of Floods on House Prices: An Imperfect Information Approach with Myopia and Amnesia", Housing Studies, 26(2): 259-279.

Rappaport, J. 2007. "Moving to Nice Weather" *Regional Science and Urban Economics*

Rappaport, J. and Jeffrey D. Sachs. 2003. "The United States as a Coastal Nation". Journal of Economic Growth, 8(1): 5-46.

Saiz, Albert. "The Geographic Determinants of Housing Supply". 2010. *Quarterly Journal of Economics*, 125(3): 1253–1296.

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).

Wing, O.E., Bates, P.D., Smith, A.M., Sampson, C.C., Johnson, K.A., Fargione, J. and Morefield, P., 2018. "Estimates of present and future flood risk in the conterminous United States." *Environmental Research Letters*, 13(3).

Table 1: (Stylized fact 1) Much construction near the coast took place in areas with SLR risk

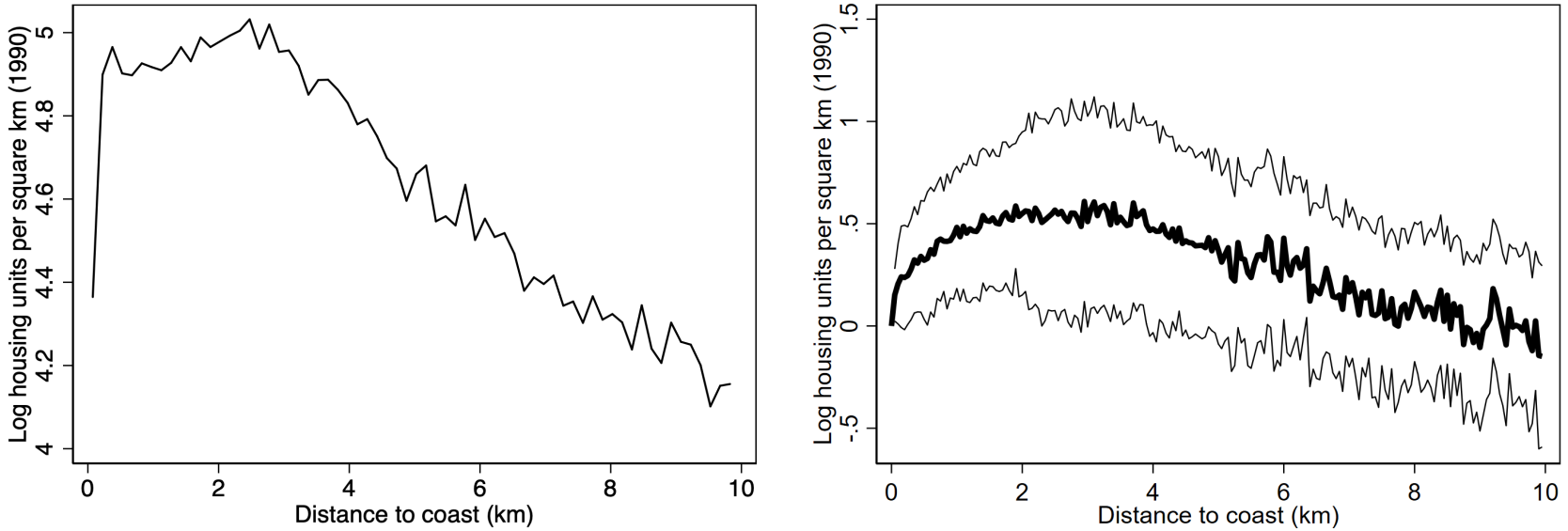
| | (1) Housing units in risky blocks (millions) | (2) Housing units in all blocks (millions) | (3) Fraction of housing units in risky blocks (%) |
|---|---|---|---|
| 1990 | 1.77 | 14.87 | 12% |
| 2010 | 2.62 | 18.11 | 14% |
| Change (1990-2010) | 0.85 | 3.24 | 26% |

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 2: (Stylized fact 2) SLR-prone areas were developed in dense tracts but not in sparse ones

| Housing units per sq km (1990) | (1) $\leq 10$ | (2) $(10, 100]$ | (3) $(100, 1000]$ | (4) $>1000$ |
|---|---|---|---|---|
| Share 1ft SLR | -3.16*** | -3.18*** | 2.26*** | 6.76*** |
| | (1.07) | (0.82) | (0.80) | (2.03) |
| Constant | 4.16*** | 7.59*** | 5.72*** | 4.62*** |
| | (0.32) | (0.20) | (0.12) | (0.17) |
| Observations | 24,927 | 149,461 | 283,208 | 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 CBSA, with non-CBSA blocks grouped into a single cluster. For reference, urban population density is defined as at least 386 people (not housing units) per square km. Results robust to controlling for log distance to the coast. Significance: *p<0.1; **p<0.05; ***p<0.01.




(a) (b)

Figure 1: (Auxiliary Finding 1) Housing units concentrate 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. Standard errors in parentheses are clustered by CBSA, with non-CBSA blocks grouped into a single cluster. Housing units per square km in Panel (b) is calculated using housing units and land area from census data at the block level. The issue in relation to calculating land area noted above is less relevant here, since the regression restricts the sample to blocks with non-zero housing units (422,311 blocks).

Table 3: Places near the coast are asymmetric, with CBD closer to the coast

| | (1)<br>Asymmetry |
|---|---|
| Distance to coast bins (km) ∈ [0,1) | 0.14*** |
| | (0.03) |
| Distance to coast bins (km) ∈ [1,2) | 0.14*** |
| | (0.03) |
| Distance to coast bins (km) ∈ [2,3) | 0.08*** |
| | (0.02) |
| Distance to coast bins (km) ∈ [3,4) | 0.07** |
| | (0.03) |
| Distance to coast bins (km) ∈ [4,5) | -0.01 |
| | (0.03) |
| Distance to coast bins (km) ∈ [5,6) | -0.00 |
| | (0.04) |
| Distance to coast bins (km) ∈ [6,7) | -0.04 |
| | (0.03) |
| Distance to coast bins (km) ∈ [7,8) | -0.02 |
| | (0.03) |
| Distance to coast bins (km) ∈ [8,9) | -0.05 |
| | (0.03) |
| Constant | 0.54*** |
| | (0.03) |
| Observations | 1,583 |

Notes: (Auxiliary finding 2) Census-designated places near the coast are asymmetric, with CBD closer to the coast. 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 CBSA, with non-CBSA places grouped into a single cluster. 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) km from the coast. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 4: Bigger places are more asymmetric, conditional on being near the coast

| | (1) Asymmetry | (2) Asymmetry | (3) Asymmetry | (4) Asymmetry |
|---|---|---|---|---|
| Ln($\|X_R - X_L\|$) | 0.015* | | 0.332*** | |
| | (0.008) | | (0.053) | |
| Ln(area) | | 0.028*** | | 0.178*** |
| | | (0.005) | | (0.032) |
| Ln(Distance to coast) | | | 0.114* | 0.225*** |
| | | | (0.067) | (0.069) |
| Ln($\|X_R - X_L\|$)*Ln(Distance to coast) | | | -0.030*** | |
| | | | (0.008) | |
| Ln(area)*Ln(Distance to coast) | | | | -0.018*** |
| | | | | (0.004) |
| Constant | 0.497*** | 0.166** | -1.102** | -1.749*** |
| | (0.067) | (0.081) | (0.451) | (0.525) |
| Observations | 1,583 | 1,583 | 1,583 | 1,583 |

Notes: (Auxiliary finding 3) Bigger census-designated places are more asymmetric, conditional on being near the coast. 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 CBSA, with non-CBSA units grouped into a single cluster. Significance: *p<0.1; **p<0.05; ***p<0.01.

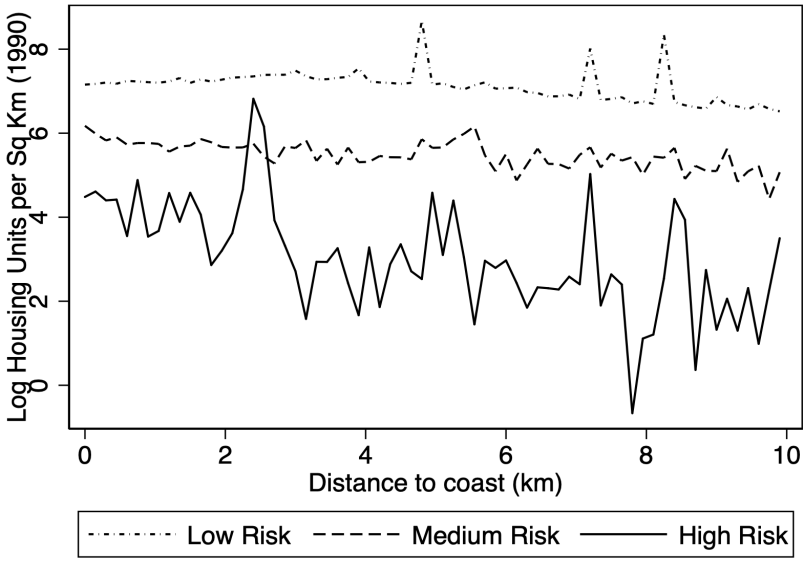
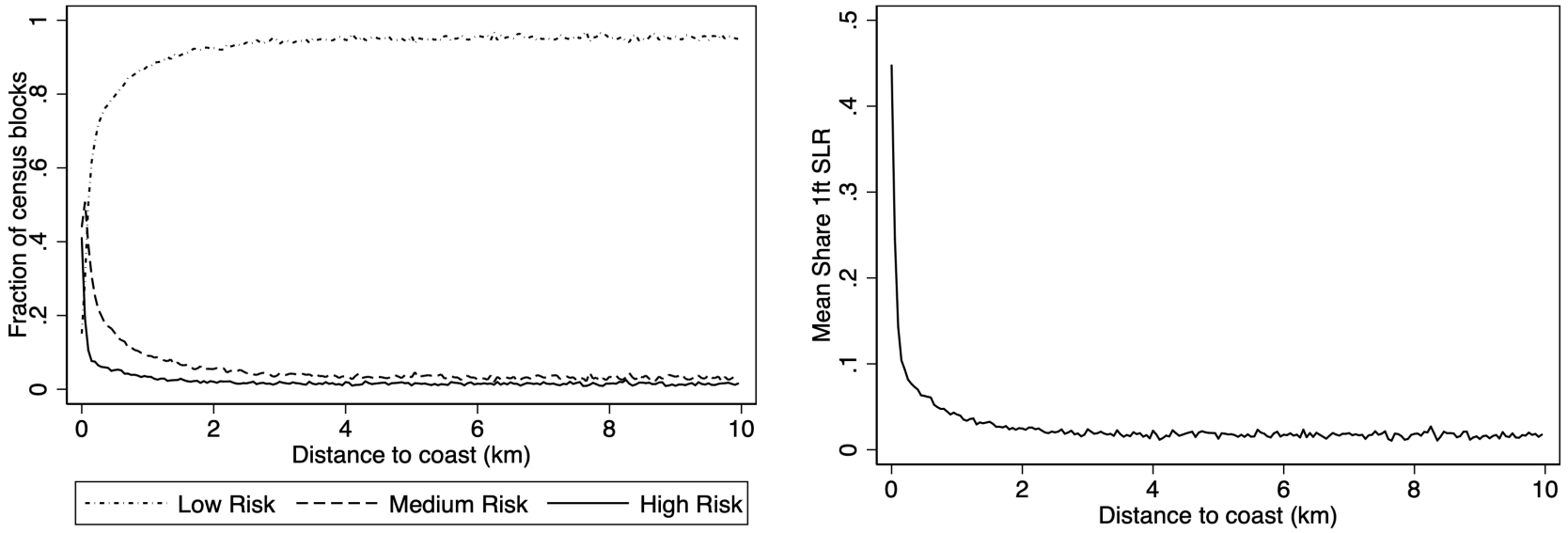



Figure 2: (Auxiliary Finding 4) Areas that are highly prone to sea level rise (SLR) are less built. The figure shows the log of average block level density (housing units per square kilometer in 1990, based on census data) by distance to the coast in 150m bins, and share of area under water with 1 foot 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: 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.

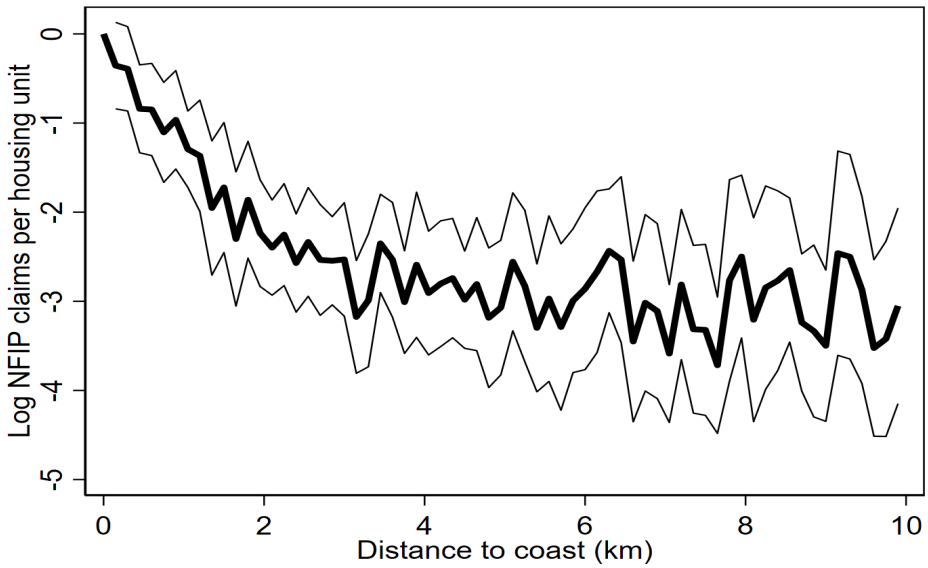



Figure 4: Damages from flooding decline rapidly with distance from the coast. The figure shows the estimated coefficients and 95% confidence interval from a regressions of log NFIP claims per housing unit, on 150m distance bins from the coast. Standard errors in parentheses are clustered by CBSA, with non-CBSA units grouped into a single cluster. The NFIP claims data are for the years 1973-2019, observed at the census tract level, and have been 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 5: (Extension of fact 2) SLR-prone areas developed in dense tracts were those with least SLR

| | (1) | (2) |
|---|---|---|
| Share 1ft SLR ∈ (0.0,0.1) | 13.53 | |
| | (8.66) | |
| Share 1ft SLR ∈ [0.1,0.2) | 15.91*** | |
| | (4.76) | |
| Share 1ft SLR ∈ [0.2,0.3) | 8.26*** | |
| | (2.49) | |
| Share 1ft SLR ∈ [0.3,0.4) | 16.53** | |
| | (6.92) | |
| Share 1ft SLR ∈ [0.4,0.5) | 10.40 | |
| | (8.59) | |
| Share 1ft SLR ∈ [0.5,0.6) | 4.54 | |
| | (4.18) | |
| Share 1ft SLR ∈ [0.6,0.7) | -0.33 | |
| | (3.66) | |
| Share 1ft SLR ∈ [0.7,0.8) | -2.60 | |
| | (4.18) | |
| Share 1ft SLR ∈ [0.8,0.9) | -2.85 | |
| | (1.97) | |
| Share 1ft SLR ∈ [0.9,1.0) | -2.79 | |
| | (2.10) | |
| Share 1ft SLR ∈ 1 | -3.88*** | |
| | (1.21) | |
| Some SLR | | 15.83** |
| | | (6.65) |
| Share 1ft SLR | | -19.40* |
| | | (9.93) |
| Constant | 4.24*** | 4.24*** |
| | (1.41) | (1.41) |
| Observations | 86,471 | 86,471 |

Notes: This table reports coefficients from two regressions where the outcome is the change in housing units at the block level from 1990-2010. The sample here is restricted to blocks where tract housing density, excluding the own block, exceeds 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 CBSA, with non-CBSA units grouped into a single cluster. In column (2), the variable $SomeSLR$ is an indicator for blocks that have share 1ft SLR >0, while $Share1ftSLR$ is a continuous measure of the share of each block prone to 1ft SLR. Results robust to controlling for log distance to the coast. Significance: *p<0.1; **p<0.05; ***p<0.01.

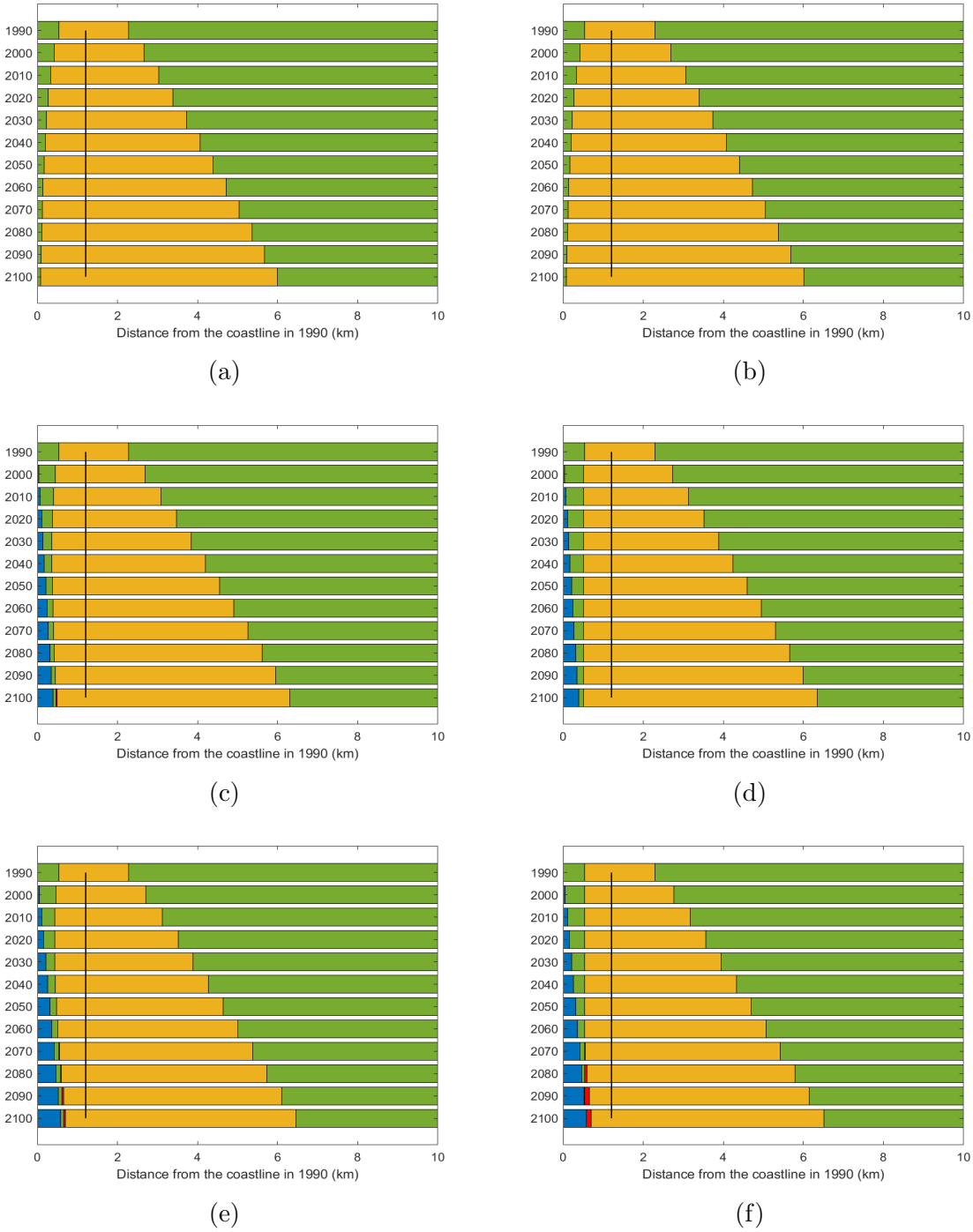



Figure 5: The figure shows results of simulations, as described in the main text, and based on parameter values detailed in Table A8. Panel (a) shows results from the static model with no SLR. Panel (b) is the same but for the dynamic version of the model. Panel (c) shows results from the static model with baseline SLR (0.0577m per decade, CBD at 2m elevation). 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 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

### Proofs

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

$$\textit{For each period } t = 1, ..., T\textit{: if } x < x_0 \textit{ then } \frac{\partial \ln(dens_t(x))}{\partial x} > 0\textit{; if } x > x_0 \textit{ then } \frac{\partial \ln(dens_t(x))}{\partial x} < 0. \quad (16)$$

**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.[40] 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

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

Due to the preferences $h_t(x) = (1-\alpha)\frac{I_t(x)}{p_t(x)}$, and using (17) we get

$$\ln(dens_t(x)) = \frac{\alpha}{1-\alpha}\ln(\alpha) + \frac{\alpha}{1-\alpha}\ln(I_t(x)) - \frac{1}{1-\alpha}\ln\left(\bar{U} + \theta_1 x + \theta_2 x^{-\sigma}\right). \quad (18)$$

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

**Lemma 1** *The city develops asymmetrically around the CBD:* $|x_{Rt} - x_0| > |x_0 - x_{Lt}|$.

**Proof.** Define the indirect utility function from housing and consumption $U(c_t(x), h_t(x), x)$ as $v(p(x), x)$, and define the location specific part of utility $U_{Loc}(x) \equiv -\theta_1 x - \theta_2 x^{-\sigma}$

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

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

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

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

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 Lemma 5 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 2** *Vanishingly small cities are symmetric:* $\lim_{w \searrow \widetilde{w}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$.

**Proof.** As $w \searrow \widetilde{w}$, the city shrinks, so $x_{Lt} \to x_0$, and 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 $\lim_{w \searrow \widetilde{w}} \frac{|x_{Rt} - x_0|}{|x_{Rt} - x_{Lt}|} = 0.5$. ■

**Lemma 3** *If $w_t$ is sufficiently low, the city avoids SLR-prone areas ($x_{L0} > D$). But if $w_t$ increases sufficiently, the city eventually expands into 1ft SLR areas ($x_{L0} < D$).*

**Proof.** We first note that $\lim_{w_t \searrow \tilde{w}} x_{Rt} = x_0$, so if $w_t$ is low the city avoids the SLR-prone area. Next, since the outside options of renters and developers are fixed, an increase $w_t$ incentivizes the development of the city on both sides. For any given location $x$ there is a sufficiently high $w_t(x)$ such that $x$ is developed. Therefore, on the coast side, the city eventually expands to the point where $x_{L0} < D$. In this process, the fraction of housing on 1ft SLR areas in the city grows over time. ■

**Lemma 5** *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 $x > x_0$. ■

## Extensions

### 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, $\theta_2 = 0$.[41] 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.[42] But as a result, housing prices near the coast (measured in the locations that have housing) may be higher than elsewhere.[43]

### Dispersion of employment throughout the city

A growing literature in urban economics (e.g., Ahlfeldt et al. 2015 and Balboni 2020) studies the dispersion of employment throughout the city. To address this in our model without losing tractability, we follow Duranton and Puga (2019) by assuming that the monetized commuting costs of a worker who lives at a distance $x$ from the CBD are $\tau(x) = \tau x^{\gamma}$, where the elasticity of commuting costs with respect to distance, $\gamma$, is less than 1.

[41] 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.

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

[43] 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 model does not account for heterogeneity in income.

### Parameter estimation

Here we explain our choice of parameters for the simulated model, where we study the synthetic 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.[44]

As we discuss in detail in Appendix Table A8, 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, $\alpha$ (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 ($\delta$), as discussed in Appendix Table A8.[45]

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 $\sigma$, is more challenging. We cannot measure utility directly, so instead we use our data on economic losses, as discussed in Auxiliary finding 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:

$$damage_k = \psi x_k^{-\sigma} + \varepsilon_{k16,} \tag{19}$$

to estimate $\sigma$.

Next, we use (18) and the data from Panel (a) of Figure 1 to estimate $\theta_1$, $\theta_2$, and $\bar{U}$ using the regression:

$$\ln\left(dens_t\left(x_k\right)\right) = \frac{\alpha}{1-\alpha}\ln\left(\alpha\right) + \frac{\alpha}{1-\alpha}\ln\left(I_t\left(x_k\right)\right) - \frac{1}{1-\alpha}\ln\left(\bar{U} + \theta_1 x_k + \theta_2 x_k^{-\sigma}\right) + \varepsilon_{k17}. \tag{20}$$

We report the regression estimates in Appendix Table A8, and in Appendix Figure A8 we compare the fit of our estimates to the data. As the figure shows, the parsimonious specification captures well

[44] 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.

[45] The CBD location in the model is closer to the coast (around 1.3 km) than the peak of the empirical density distribution (around 2.5 km). This may be in part because the density distribution peak is not very steep.

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. 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 $\sum_{T\geq s>1}\delta^s p_s(x) \geq \sum_{T\geq s\geq 1}\delta^s p_A$.[46]

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 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.[47] 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.

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).[48] 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 annual growth of NFIP expenditure from 1978-2019. We then allocate the costs by distance to the coast, using the estimates from regression (19).

## Online Appendix B: Data

*Defining the Sample*

[46] 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.

[47] 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.

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

In our analysis we focus on the 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.[49] 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).[50] 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 Sea Level Rise*

Detailed maps of areas anticipated to be inundated for various future sea level rise scenarios were obtained from NOAA's digital coast platform.[51] 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.[52] This surface is then subtracted from a digital elevation model (based on lidar data).[53]

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").[54] 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 (For example, New Orleans is assumed not to be inundated in any of the SLR

[49] 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 A3), which shows the CBD of the city of Boston and several census designated places whose own CBD is within 25km of Boston's.

[50] 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).

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

[52] The base elevation is NOAA VDATUM MHHW.

[53] Given the way the SLR maps are constructed, our measure of areas that will be inundated with 1ft of SLR includes areas that are currently under water at high tide.

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

scenarios).

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 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., 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 the SLR area for the relevant blocks.[55]

*Census-designated places*

Data on census-designated places are taken from the NHGIS archive (Manson et al. 2019) based

[55]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. One block (in FL) with an anomalously large share 1ft SLR value was dropped from the blocks dataset, as was one block (in NC) with a negative area value. For the remaining blocks in our data, the share 1ft SLR variable was winsorized to 1 (this affected 1,891 blocks, 99% of which had share 1ft SLR of less than 1.01).

on census data and definitions.[56] 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.[57] 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. 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 $x_R$ 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.[58]

*Core-Based Statistical Areas*

Throughout our analysis, we cluster the standard errors on the Core-Based Statistical Areas (CBSAs), pooling all the locations that do not fall into any CBSA into one cluster and in some cases dropping them altogether. Blocks and tracts are embedded in counties, and counties are matched to CBSAs using the 2010 Census crosswalk. In regressions with places, we match places to CBSA in 2010 using the Geocorr 2018 Geographic Correspondence Engine. For places matched to multiple CBSAs, we set the cluster variable to the CBSA with the largest geographical overlaps with the place. We match each cell to a CBSA if its centroid falls within boundaries of the CBSA.

*Housing Data*

Our housing data for both prices and quantities, including a measure of which housing units are primary residences, 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 American Community Survey. The data are obtained from the National Historical Geographic Information System (NHGIS) data archive (Manson et al. 2019).[59]

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 backwards 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).

[56] Examples of census-designated places in the Boston area include Boston, Cambridge, and Somerville. The definition of places used by the Census Bureau is detailed here:
https://www.census.gov/content/dam/Census/data/developers/understandingplace.pdf.

[57] 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.

[58] 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.

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

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 $p_{areaij}$ :

$$p_{areaij} \equiv \frac{(B_i^s \cap B_j^t)}{(B_i^s)} \tag{21}$$

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 harmonize 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. Harmonizing backwards from 2010 to 1990 therefore requires that the intersect value contained within the crosswalk files, $p_{areaij}$ be amended as follows:

$$p_{areaji} = \frac{(B_i^{90} \cap B_j^{10})}{(B_j^{10})} = \frac{(B_i^{90} \cap B_j^{10})}{(B_i^{90})} * \frac{(B_i^{90})}{(B_j^{10})} \tag{22}$$

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:

$$H_{ji}^{90,10} = p_{ji} * H_j^{10} \tag{23}$$

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[60] data from NOAA's Coastal Change Analysis Program (C-CAP).[61] Through the C-CAP program, NOAA produces nationally standardized land cover and land change data for the coastal regions of the U.S.[62] 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

[60] 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.

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

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

(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).[63]

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 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".[64] It is therefore unclear 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 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 five-year period.[65]

For use in our analysis, we take land cover observations in 1996 and 2010 (and calculate changes

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

[64] 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.

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

from 1996-2010). We aggregate the 30m x 30m pixels in the C-CAP data to 150 x 150-meter cells.[66] 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 and Zoning*

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.

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.[67] 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.[68] 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.[69] Finally, information on the location of US military bases along the US coastline is taken from the US Department of Transport.[70]

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

[66] 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.

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

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

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

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

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 remain qualitatively and quantitatively unchanged by this restriction on the data.

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.[71] 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.[72] 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.

*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.

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.

[71] 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).

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

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.[73] 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.[74] 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.[75] 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 up and standardized to 2020 US dollars.[76]

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.[77]

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

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

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

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

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

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

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 infrastructure. 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


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 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." National Bureau of Economic Research wp 26573.

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". Washing-

ton 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/

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).

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

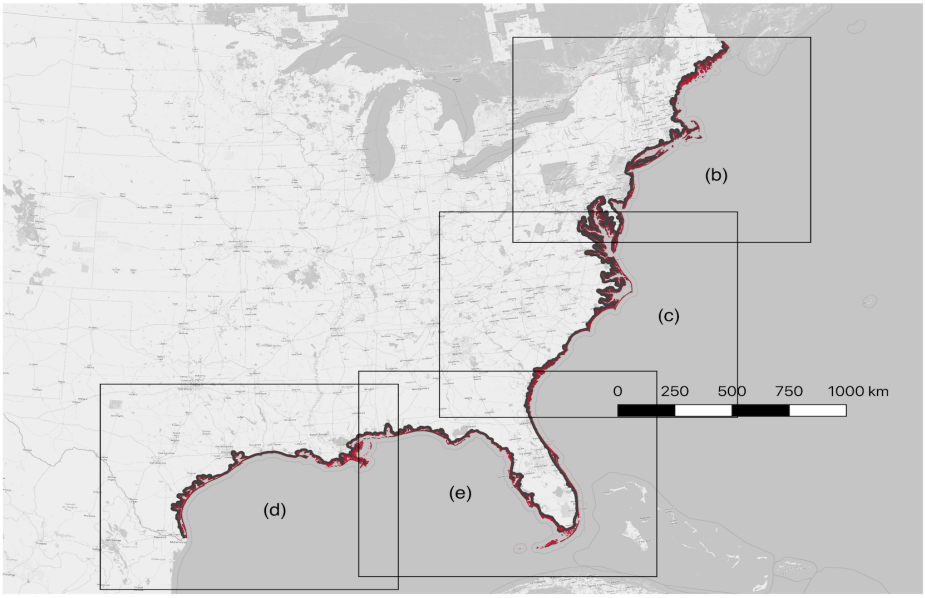
US Congressional Budget Office (CBO). "Expected Costs of Damage from Hurricane Winds and Storm Related Flooding". Washington DC: Congressional Budget Office (2019). Accessed July 2020 at:

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

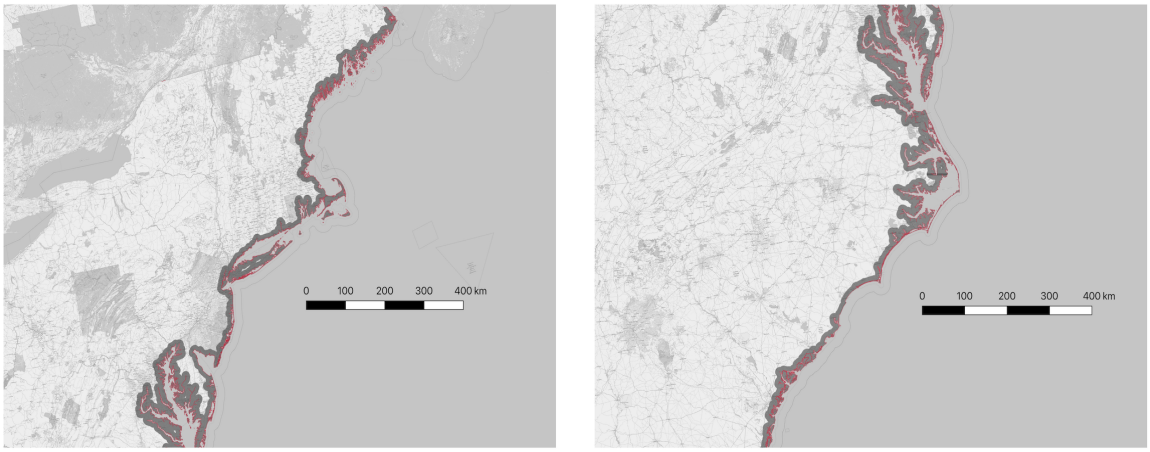
US Department of Transport (USDOT) – Bureau of Transportation Statistics (BTS). "Military Bases Dataset". Washington DC: USDOT. Accessed June 2020 at

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

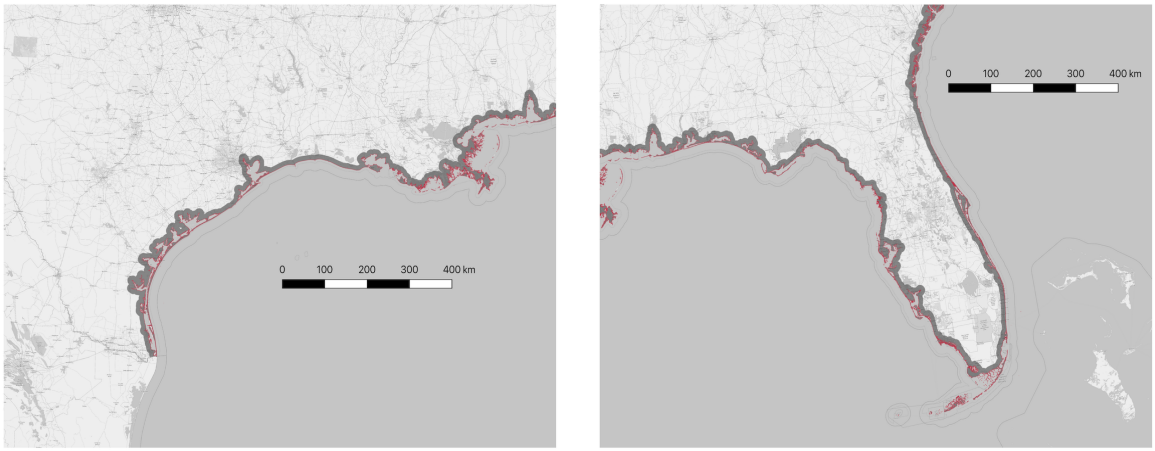
US Geological Survey (USGS) – Science Base. "Protected Areas Database of the United States (PAD-US)". Reston, VA: USGS (2017). Accessed June 2020 at
https://www.sciencebase.gov/catalog/item/5963ea3fe4b0d1f9f059d955.




(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.

Table A1: Robustness to stylized fact 1

| Housing units per sq km (1990) | (1) Stock (1990) | (2) Change(1990-2010) | (3) Stock (1990) | (4) Change (1990-2010) | (5) Stock (1990) | (6) Change (1990-2010) |
|---|---|---|---|---|---|---|
| Panel A | | | | | | |
| Share 1ft SLR | -24.63*** | -0.37 | -13.20*** | -0.60 | -15.67*** | -0.14 |
| | (6.71) | (0.95) | (2.65) | (0.76) | (2.63) | (0.93) |
| Constant | 28.47*** | 5.97*** | 27.94*** | 5.98*** | 30.18*** | 6.05*** |
| | (4.95) | (0.59) | (0.12) | (0.04) | (0.10) | (0.04) |
| Observations | 544,067 | 544,067 | 544,067 | 544,067 | 479,625 | 479,625 |
| CBSA FE | no | no | yes | yes | yes | yes |
| Sample | full | full | full | full | urban | urban |

| Housing units per sq km (1990) | (1) Stock (1990) | (2) Change (1990-2010) | (3) Stock (1990) | (4) Change (1990-2010) | (5) Stock (1990) | (6) Change (1990-2010) |
|---|---|---|---|---|---|---|
| Panel B | | | | | | |
| Share 1ft SLR | -22.59*** | 1.11 | -13.73*** | 0.99 | -16.07*** | 1.74* |
| | (6.11) | (1.07) | (2.87) | (0.92) | (3.18) | (0.98) |
| Constant | 24.45*** | 7.88*** | 23.91*** | 8.50*** | 26.18*** | 9.13*** |
| | (2.50) | (1.33) | (2.55) | (1.34) | (2.87) | (1.43) |
| Observations | 544,067 | 544,067 | 544,067 | 544,067 | 479,625 | 479,625 |
| CBSA FE | no | no | yes | yes | yes | yes |
| Sample | full | full | full | full | urban | urban |
| Distance Controls | yes | yes | yes | yes | yes | yes |

Notes: This table reports regression estimates from specification (1) for the stock of housing units in 1990 (in Columns 1, 3 and 5) and the change in housing units from 1990 to 2010 (in columns 2, 4 and 6). Columns (1)-(4) use the full sample and columns (5)-(6) use only the urban sample (locations within CBSAs). Columns (3)-(6) include CBSA fixed effects, and in columns (3)-(4), we have all non-CBSA locations in one group. Panel B repeats the same specifications as Panel A, with the addition of 1km distance bins to the coast. Standard errors in parentheses are clustered by CBSA, with non-CBSA blocks grouped into a single cluster in columns (1)-(4). Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A2: Robustness to stylized fact 2

| Housing units per sq km (1990) | (1) ≤10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Panel A | | | | |
| Share 1ft SLR | -2.08*** | -2.94*** | 1.29 | 7.75*** |
| | (0.61) | (0.61) | (1.19) | (1.89) |
| Observations | 24,927 | 149,461 | 283,208 | 86,471 |
| CBSA FE | yes | yes | yes | yes |
| Sample | full | full | full | full |

| Housing units per sq km (1990) | (1) ≤10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Panel B | | | | |
| Share 1ft SLR | -1.91 | -3.12*** | 0.90 | 7.87*** |
| | (1.51) | (1.01) | (1.20) | (1.90) |
| Observations | 9,099 | 106,810 | 277,590 | 86,126 |
| CBSA FE | yes | yes | yes | yes |
| Sample | urban | urban | urban | urban |

| Housing units per sq km (1990) | (1) ≤10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Panel C | | | | |
| Share 1ft SLR | -2.78** | -0.72 | 3.99*** | 6.02*** |
| | (1.18) | (0.85) | (1.22) | (2.03) |
| Observations | 24,927 | 149,461 | 283,208 | 86,471 |
| CBSA FE | no | no | no | no |
| Sample | full | full | full | full |
| Distance Controls | yes | yes | yes | yes |

| Housing units per sq km (1990) | (1) ≤10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Panel D | | | | |
| Share 1ft SLR | -2.03*** | -0.34 | 3.40*** | 5.81*** |
| | (0.70) | (0.78) | (1.16) | (1.72) |
| Observations | 24,927 | 149,461 | 283,208 | 86,471 |
| CBSA FE | yes | yes | yes | yes |
| Sample | full | full | full | full |
| Distance Controls | yes | yes | yes | yes |

| Housing units per sq km (1990) | (1) ≤10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Panel E | | | | |
| Share 1ft SLR | -1.75 | 0.58 | 3.05** | 5.90*** |
| | (1.89) | (0.95) | (1.18) | (1.74) |
| Observations | 9,099 | 106,810 | 277,590 | 86,126 |
| CBSA FE | yes | yes | yes | yes |
| Sample | urban | urban | urban | urban |
| Distance Controls | yes | yes | yes | yes |

Notes: This table replicates findings reported in Table 2 in the main paper, adding CBSA effects in a full sample of blocks, with all non-CBSA units grouped into one SE (Panel A), CBSA fixed effects, dropping non-CBSA blocks (Panel B), distance controls (Panel C), and combinations of these (Panels D and E). 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 CBSA, with non-CBSA blocks grouped into a single cluster. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A3: Robustness to stylized fact 2 excluding restricted areas

| Housing units per sq km (1990) | (1) <=10 | (2) (10, 100] | (3) (100, 1000] | (4) >1000 |
|---|---|---|---|---|
| Share 1ft SLR | -2.96 | -3.42*** | 1.86* | 7.69*** |
| | (1.84) | (1.00) | (1.10) | (2.36) |
| Constant | 3.97*** | 7.80*** | 5.77*** | 4.53*** |
| | (1.07) | (0.92) | (0.67) | (1.39) |
| Observations | 15,814 | 130,511 | 286,486 | 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 CBSA, with non-CBSA blocks grouped into a single cluster. Urban population density: at least 386 people (not housing units) per square km. Results robust to controlling for log distance to the coast. In this table, data exclude blocks where 50% or more of block area is restricted from 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. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A4: Robustness to stylized fact 2 using the Landsat data

| Avg. % developed (1996) | (1) <6.25 | (2) (6.25,12.5] | (3) (12.5,25] | (4) (25,50] | (5) >50 |
|---|---|---|---|---|---|
| Share 1ft SLR | -5.16*** | -26.61*** | -26.93*** | -15.43*** | 2.44 |
| | (1.26) | (2.82) | (2.78) | (2.85) | (4.95) |
| Constant | 4.89*** | 24.65*** | 26.73*** | 20.01*** | 8.87*** |
| | (1.17) | (2.16) | (2.22) | (2.09) | (2.03) |
| 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 CBSA, with non-urban blocks grouped into a single cluster. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A5: Robustness to stylized fact 2 - weather shifters

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Change in Housing Units (1990-2010) | | | | | | | |
| Housing units per sq km (1990) | All | All | <= 10 | (10, 100] | (100, 1000] | > 1000 | <= 10 | (10, 100] | (100, 1000] | > 1000 |
| JanTemp | 1.55*** | | 2.69* | 2.78*** | 1.32* | -0.13 | | | | |
| | (0.58) | | (1.45) | (0.68) | (0.66) | (1.04) | | | | |
| JanTemp#Share 1ft SLR | | | -2.91 | -0.96 | 2.10* | 5.09*** | | | | |
| | | | (1.79) | (1.08) | (1.10) | (1.10) | | | | |
| Share 1ft SLR | | | -3.36*** | -3.87*** | 1.46 | 6.79*** | -3.33*** | -4.07*** | 1.41 | 6.87*** |
| | | | (1.10) | (0.71) | (1.08) | (1.46) | (1.02) | (0.66) | (1.05) | (1.42) |
| MildWinter | | 1.66*** | | | | | 1.46* | 2.55*** | 1.48** | 0.12 |
| | | (0.48) | | | | | (0.85) | (0.56) | (0.59) | (0.90) |
| MildWinter#Share 1ft SLR | | | | | | | -1.44 | -0.45 | 2.27** | 5.10*** |
| | | | | | | | (1.27) | (0.81) | (1.01) | (1.23) |
| Constant | 5.96*** | 5.96*** | 4.32*** | 7.52*** | 5.63*** | 4.58*** | 4.24*** | 7.48*** | 5.65*** | 4.65*** |
| | (0.60) | (0.58) | (0.84) | (0.58) | (0.67) | (1.25) | (0.91) | (0.59) | (0.66) | (1.35) |
| Observations | 544,061 | 544,061 | 24,923 | 149,460 | 283,207 | 86,126 | 24,923 | 149,460 | 283,207 | 86,471 |

Notes: This table reports estimates of specification (5), for the full sample of all blocks (in Columns 1 and 2), and dividing the sample according to the level of housing units per square km in census tracts (1990), excluding own block (Columns 3-10). The outcome in each case is the change in housing units at the block level from 1990-2010. Standard errors in parentheses are clustered by CBSA, with non-urban blocks grouped into a single cluster. The variable JanTemp is a standardized version of the Daily Minimum January Temperature from Rappaport (2007). The variable MildWinter is an index calculated as $(JanTemp - Snowfall - HDD - Colddays - Snowdays)/5$, where JanTemp is as defined previously, and Snowfall, HDD, Colddays and Snowdays are standardized versions of: Annual Snowfall, Mean Annual Heating Degree Days, Cold Temperature Days, and Annual Snowdays, respectively (all from Rappaport 2007). Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A6: Robustness to Table 5 excluding restricted areas

| | (1) | (2) |
|---|---|---|
| Share 1ft SLR ∈ (0.0,0.1) | 13.99 | |
| | (8.85) | |
| Share 1ft SLR ∈ [0.1,0.2) | 14.97*** | |
| | (4.35) | |
| Share 1ft SLR ∈ [0.2,0.3) | 8.45*** | |
| | (2.18) | |
| Share 1ft SLR ∈ [0.3,0.4) | 15.14** | |
| | (6.50) | |
| Share 1ft SLR ∈ [0.4,0.5) | 9.92 | |
| | (8.04) | |
| Share 1ft SLR ∈ [0.5,0.6) | 7.35** | |
| | (3.08) | |
| Share 1ft SLR ∈ [0.6,0.7) | -0.32 | |
| | (3.27) | |
| Share 1ft SLR ∈ [0.7,0.8) | 0.99 | |
| | (3.57) | |
| Share 1ft SLR ∈ [0.8,0.9) | -3.10 | |
| | (1.90) | |
| Share 1ft SLR ∈ [0.9,1.0) | -2.72 | |
| | (2.06) | |
| Share 1ft SLR ∈ 1 | -3.83*** | |
| | (1.25) | |
| Some SLR | | 15.77** |
| | | (6.46) |
| Share 1ft SLR | | -18.72* |
| | | (9.51) |
| Constant | 4.14*** | 4.14*** |
| | (1.41) | (1.41) |
| 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 CBSA, with non-CBSA blocks grouped into a single cluster. 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 have been further restricted to exclude blocks where 50% or more of block area is restricted from development, as designated by the Protected Areas Database of the US (PAD-US), or identified as military land or as a local (city) park. By removing blocks that have a large proportion of their land restricted for development, which are typically low density blocks, we end up with more census tracts that exceed the density threshold of at least 1000 housing units per square km and therefore a slightly larger sample in the regressions reported here than in those reported in Table 5 in the main paper. Exact definitions and data sources for restricted areas are detailed further in the Data Appendix. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A7: Robustness to Table 5 using the Landsat data

| | (1) | (2) |
|---|---|---|
| Share 1ft SLR ∈ (0.0,0.1) | 8.74*** | |
| | (2.83) | |
| Share 1ft SLR ∈ [0.1,0.2) | 3.32* | |
| | (1.93) | |
| Share 1ft SLR ∈ [0.2,0.3) | 2.79 | |
| | (2.59) | |
| Share 1ft SLR ∈ [0.3,0.4) | 4.58 | |
| | (3.42) | |
| Share 1ft SLR ∈ [0.4,0.5) | 0.93 | |
| | (3.47) | |
| Share 1ft SLR ∈ [0.5,0.6) | 3.11 | |
| | (4.71) | |
| Share 1ft SLR ∈ [0.6,0.7) | 4.45 | |
| | (5.06) | |
| Share 1ft SLR ∈ [0.7,0.8) | 1.09 | |
| | (3.86) | |
| Share 1ft SLR ∈ [0.8,0.9) | 9.66*** | |
| | (3.55) | |
| Share 1ft SLR ∈ [0.9,1.0) | -0.97 | |
| | (3.31) | |
| Share 1ft SLR ∈ 1 | -4.26* | |
| | (2.20) | |
| Some SLR | | 6.98*** |
| | | (2.60) |
| Share 1ft SLR | | -7.84*** |
| | | (2.79) |
| Constant | 8.59*** | 8.59*** |
| | (2.09) | (2.09) |
| 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 CBSA, with non-urban blocks grouped into a single cluster. 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. Significance: *p<0.1; **p<0.05; ***p<0.01.

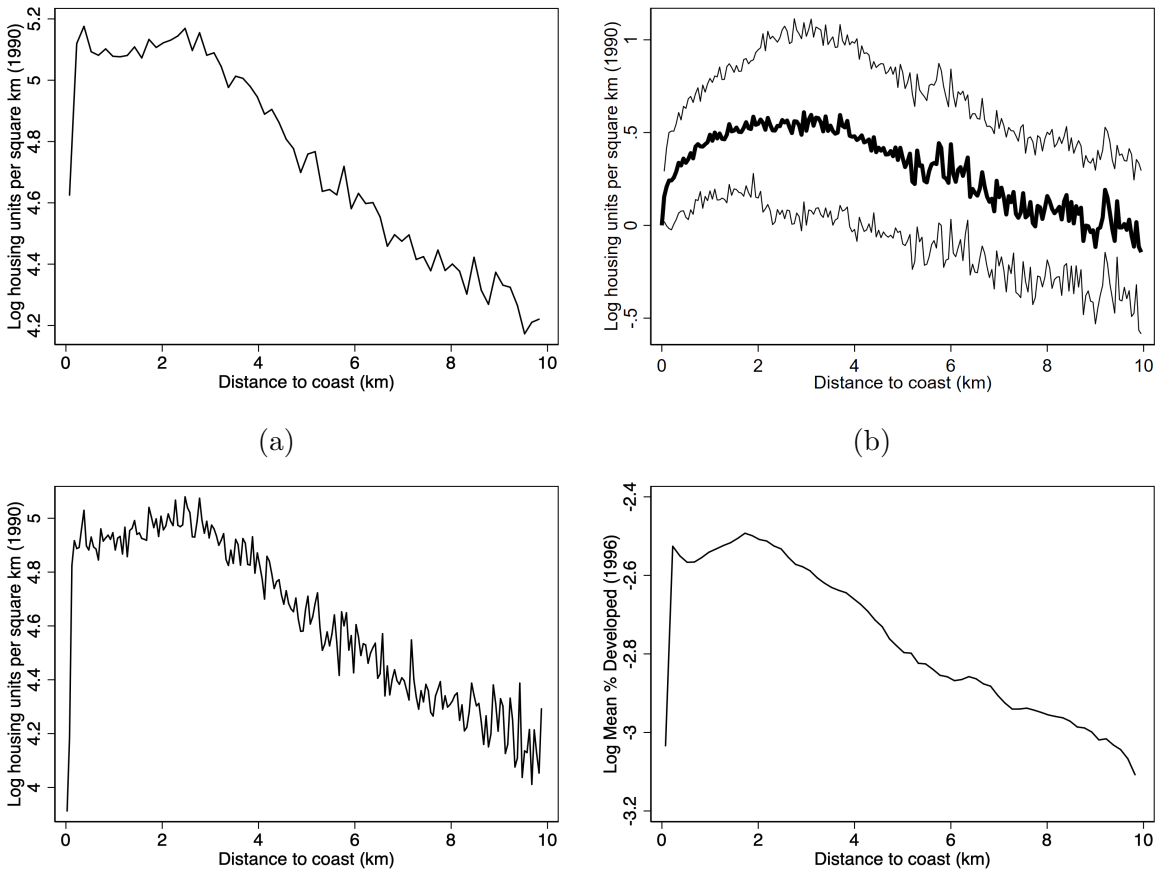



Figure A2: (Robustness of Auxiliary Finding 1) Housing concentrates near – but not right at – the coast. Panel (a) shows log housing density (average number of housing units in 1990 per square kilometer) by distance to the coast for 150m bins, based on census data at the block level. Panel (b) shows 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 are restricted to exclude blocks (and gridcells) where 50% or more of block (or gridcell) area is restricted from 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 Panel (a) of Figure 1 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.

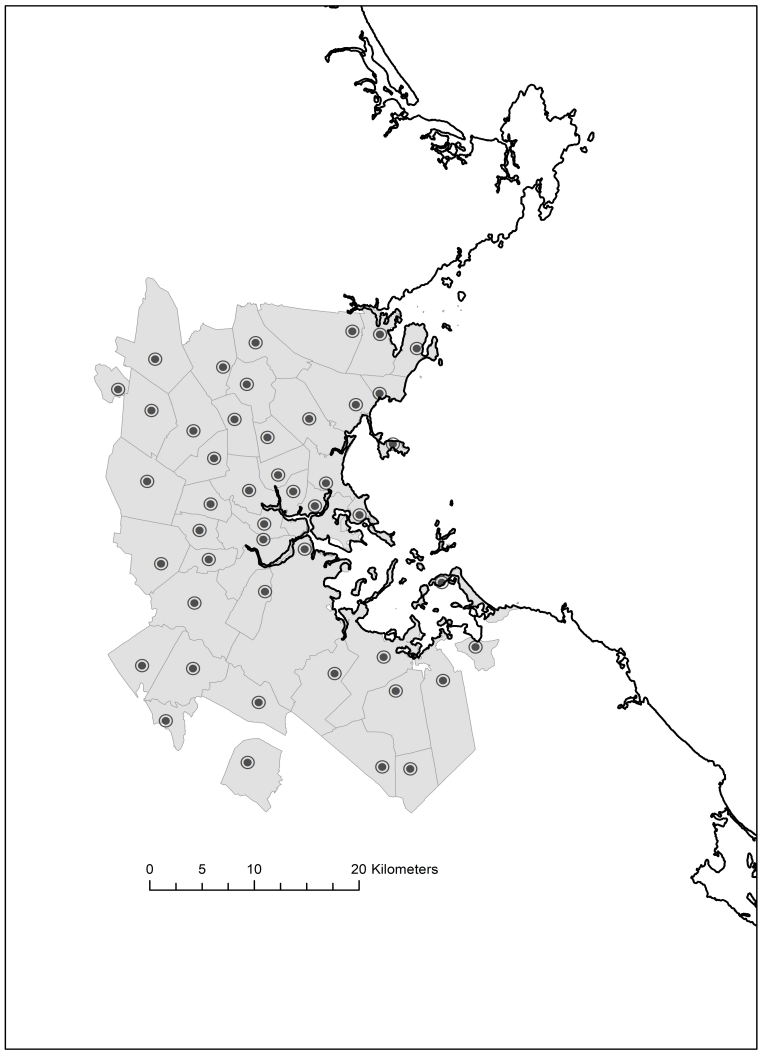
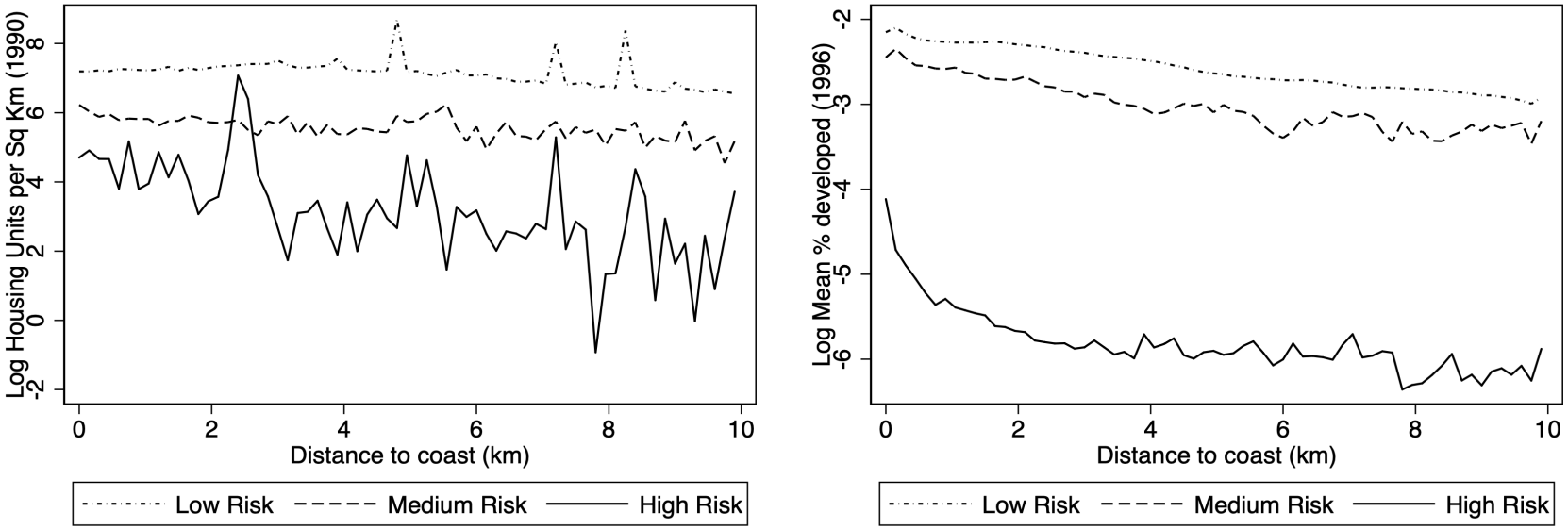
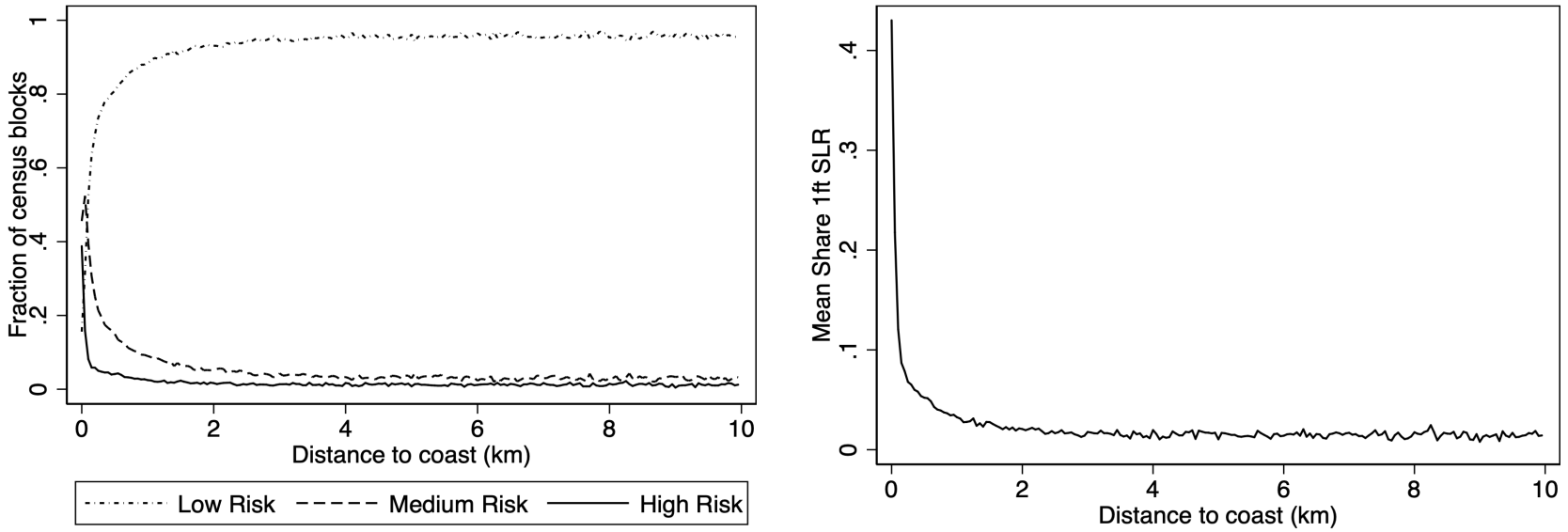



Figure A3: (Example of Auxiliary Finding 2) Places near the coast are asymmetric. The figure shows the CBD of the city of Boston plus a number of Census Designated Places (CDPs) whose own CBD is within 25km of Boston's. The polygons of each CDP, as well as the points denoting the location of their CBD, 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.



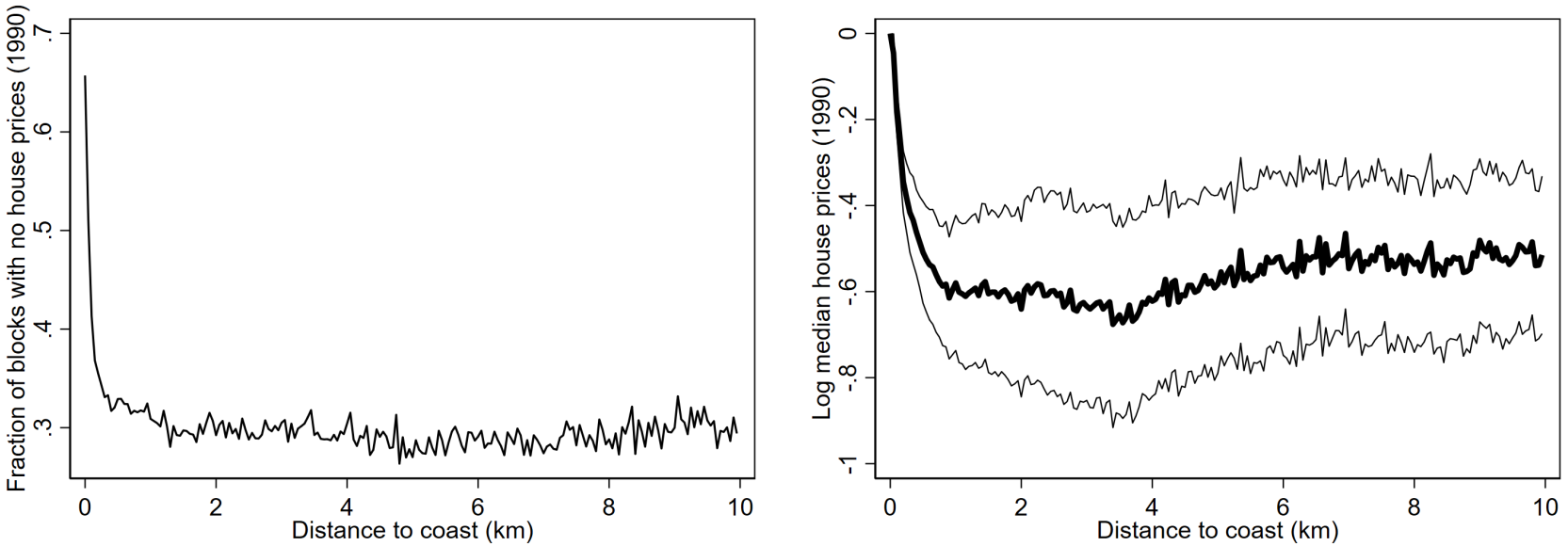

(a) (b)

Figure A4: (Robustness to Auxiliary Finding 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 have been restricted to exclude blocks where 50% or more of block area is restricted from 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).



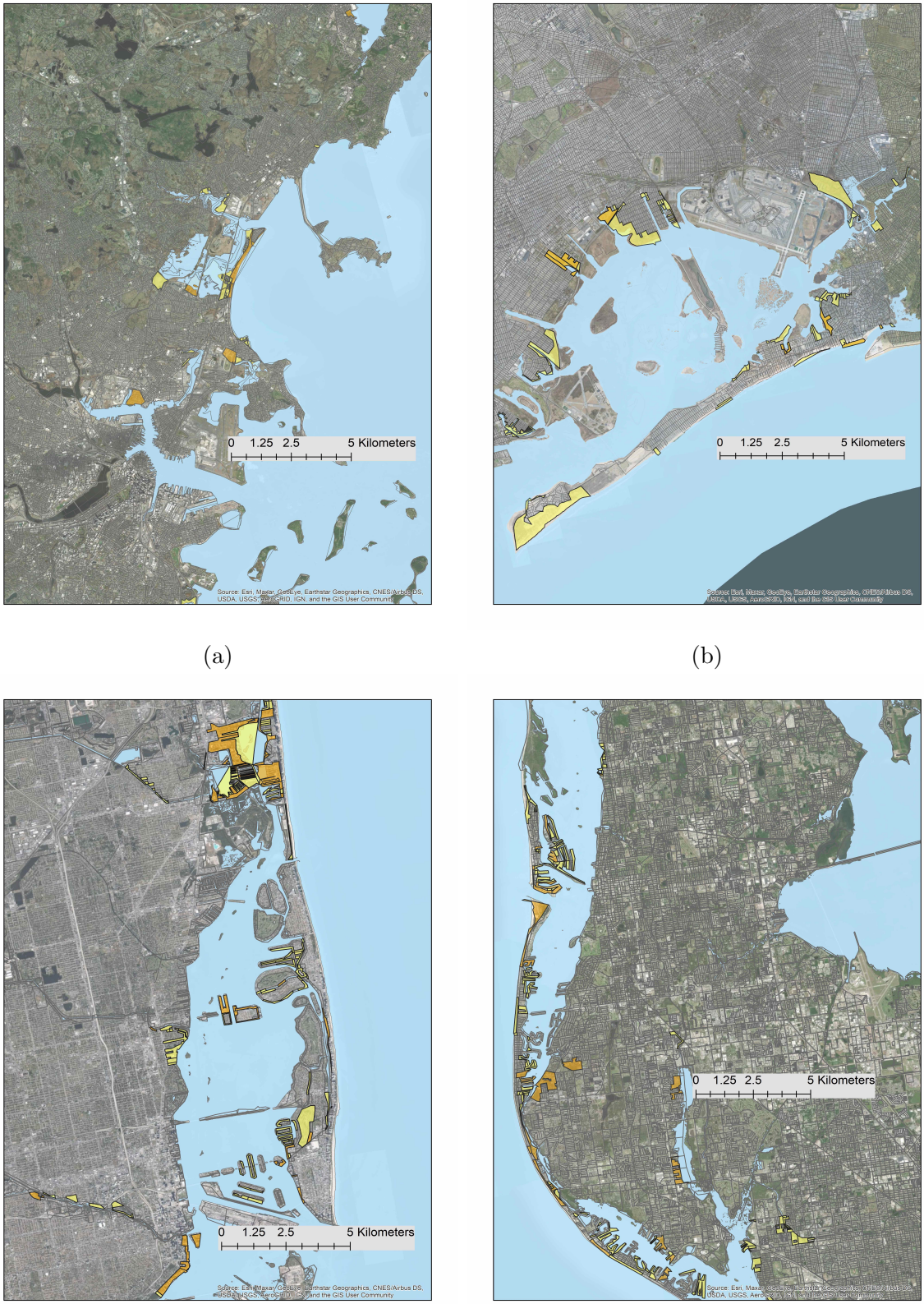

(a) (b)

Figure A5: (Robustness to Figure 3 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 have been restricted to exclude blocks where 50% or more of block area is restricted from 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. Standard errors in parentheses are clustered by CBSA, with non-CBSA units grouped into a single cluster.




(a) (b)

(c) (d)

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).

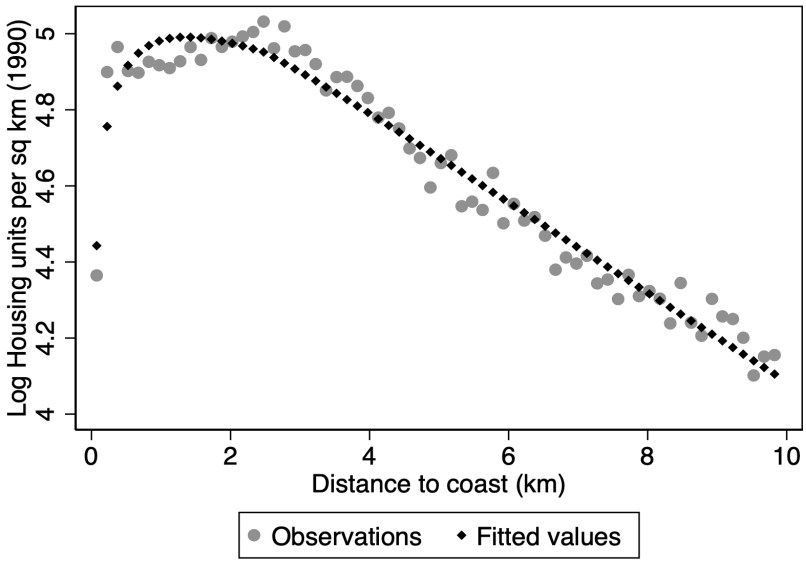



Figure A8: Estimating the thetas. The figure shows observations and fitted values for log housing density by distance to the coast (as per Figure 1a). The fitted values are based on estimating equation (17), which we use to estimate $\theta_1$, $\theta_2$ and U.

Table A8: Parameter estimates

| Parameter | Value | Source | Notes |
|---|---|---|---|
| $\alpha$ | 0.75 | Davis & Ortalo-Magne (2011) | Consumption share of income |
| $\sigma$ | 0.3454 (0.0637) | Estimated | Estimated by equation (17) using data from Fig4 (Robust standard error) |
| $\theta_1$ | 0.1419 (0.0049) | Estimated | Estimated by equation (18), using data from Fig1a (Robust standard error) |
| $\theta_2$ | 5763.27 (672.73) | Estimated | Estimated by equation (18), using data from Fig1a (Robust standard error) |
| U | 3892.92 (61.10) | Estimated | Estimated by equation (18), using data from Fig1a (Robust standard error) |
| w | 536580 | Observed | 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 | Observed | Average hours worked per year in 1990 |
| commuting distance | 500 | see notes | Total distance per km of commute (in km per year) |
| $x_0$ | 2.475 | Observed | Location of peak of housing density (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 above sea level) |
| 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 |
| $\delta$ | 0.82 | Treasury | Discount rate per decade, based on 30-year interest rate on last day of 2020 |
| $p_A$ | see notes | Estimated | Agricultural rent, equated to housing rent at $x_L$ in 1990 (static model) |
| $\tilde{s}$ | see notes | CBO (2019). FEMA (2021) | Public subsidy to flood damages |
| $\gamma$ | 0.0729 | Duranton & Puga (2019) | Elasticity of commuting cost with respect to distance from the CBD |

Notes: The estimate for $\sigma$ is from a regression of NFIP damages per housing unit on distance to the coast, as per equation (15) and Figure 4. Estimates for $\theta_1$, $\theta_2$ and U are based on estimating equation (16), 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 synthetic CBD in km from the coast. It is the peak of the observed density of housing units per sq km (as in Figure 1a). $x_L$ 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 $\delta$ 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 USD19.4bn 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.

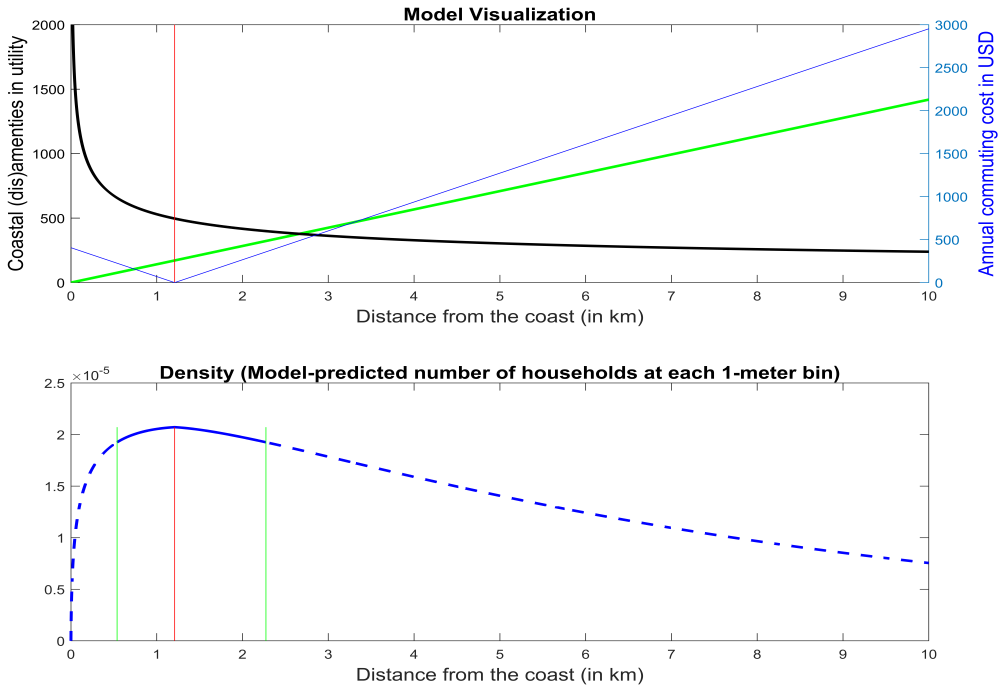



Figure A9: This figure summarizes aspects of the simulated model for the year 1990. The top panel of this figure is a visualization of the simulation from the residents' perspective. It shows the linear decline in coastal amenity (in green) and the convex decline in flooding disamenity (in black) as 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.

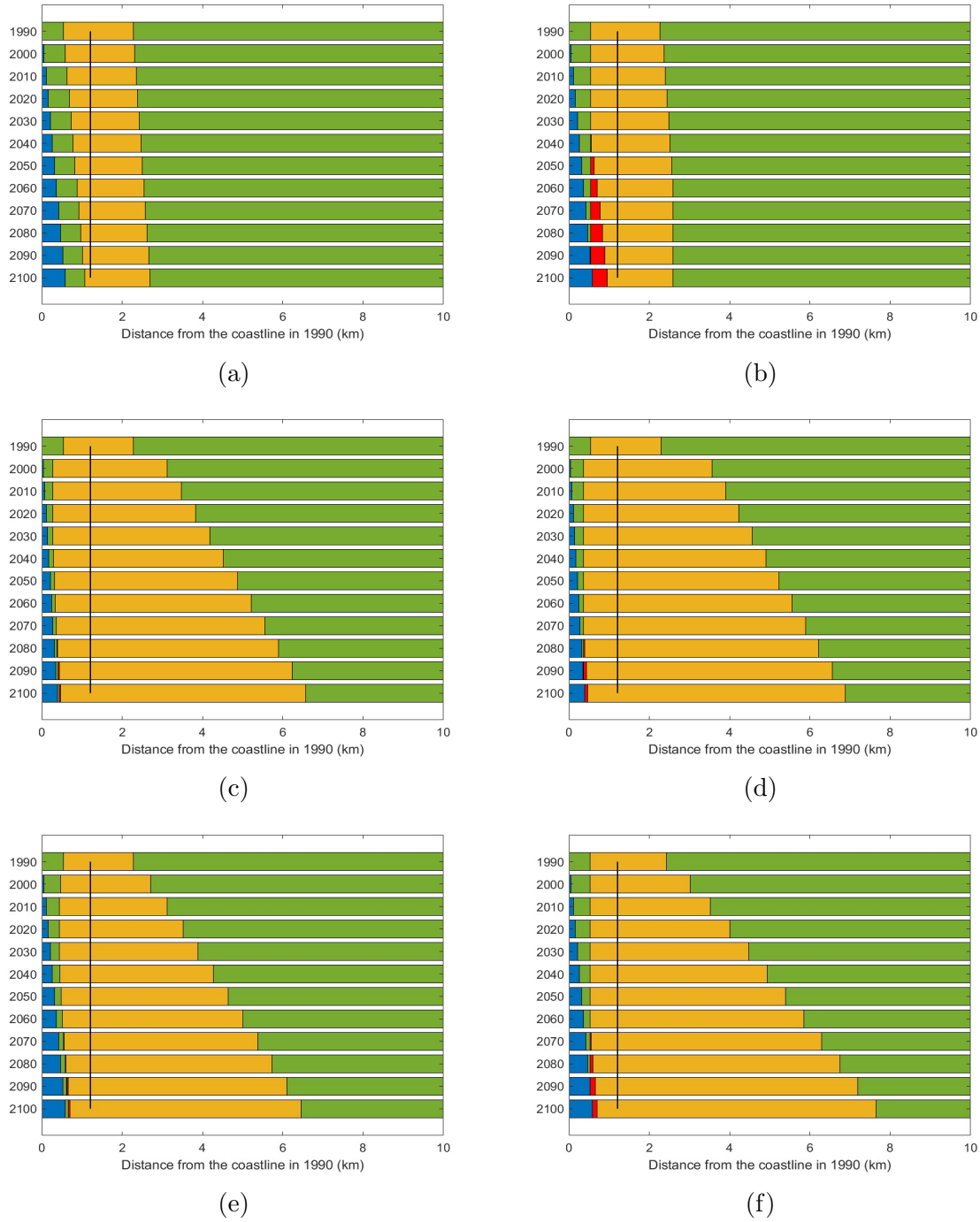



Figure A10: (Additional simulation results) Panel (a) shows results from the static model with fast SLR and without wage growth. Panel (b) is the same as (a) but for the dynamic version of the model. Panel (c) shows results from the static model with 1% wage growth per decade, baseline SLR and adding a subsidy of 33%. Panel (d) is the same as (c) but for the dynamic version of the model. Panel (e) is the same as Panel (e) in Figure 5 with a commuting elasticity of 0.0729 as per Duranton and Puga (2019). Panel (f) is the same as Panel (e) in Figure 5 with a commuting elasticity of 0.0729. 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.

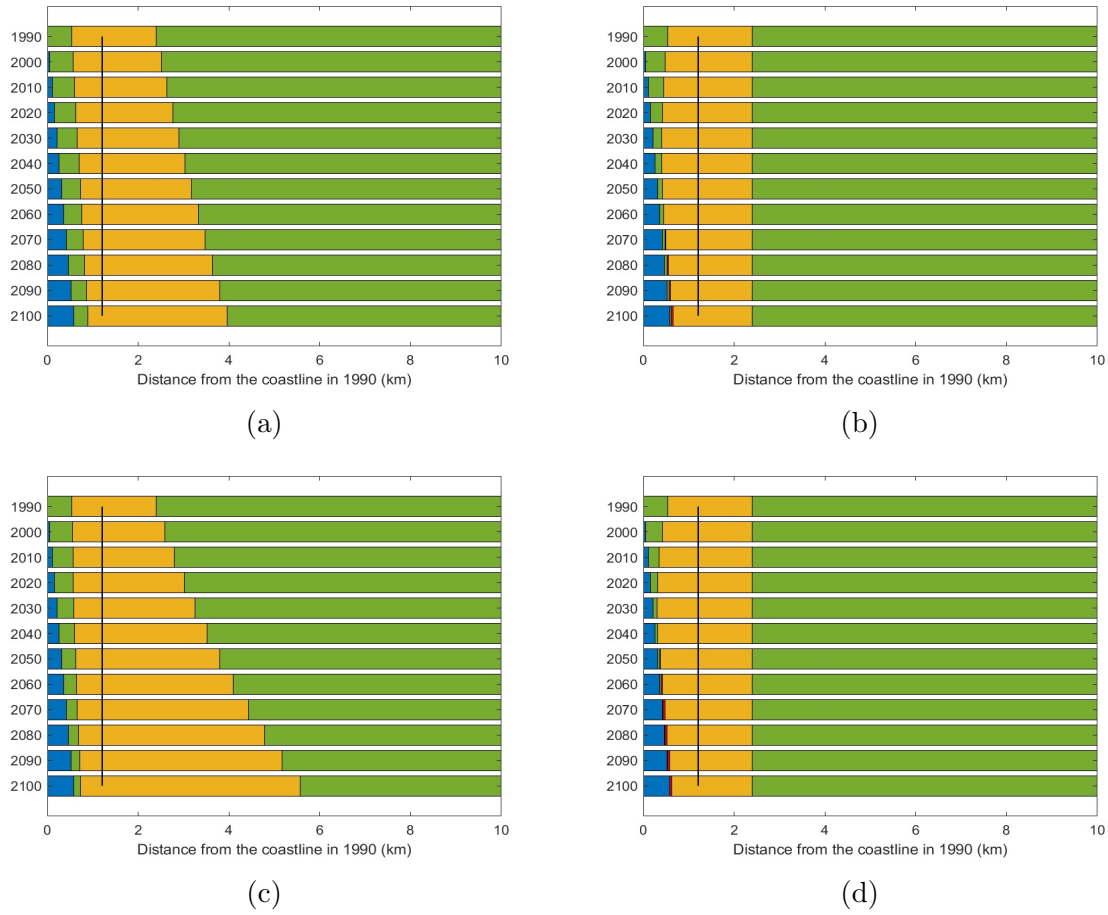



Figure A11: (Additional simulation results) Panel (a) shows results from the static model with population growth of 5% per decade. Panel (b) is the same as (a) but with the restriction that the city cannot expand on the right hand edge. Panel (c) shows results from the static model with 10% population growth per decade. Panel (d) is the same as (c) but with the restriction that the city cannot expand on the right hand edge. All the simulations presented in this figure are generated for the scenario with fast SLR. 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.

Table A9: Welfare Implications

| | Percentage change in land value due to SLR | | Percentage change in net land value due to subsidy | | | Percentage change in land value due to fixed CBD | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | Baseline SLR | Fast SLR | No SLR | Baseline SLR | Fast SLR | Baseline SLR | Fast SLR |
| Static model | | | | | | | |
| 0-1km | -20.58% | -29.59% | -3.49% | -3.94% | -4.18% | 0.23% | 0.25% |
| 0-10km | -1.53% | -2.22% | -0.90% | -0.96% | -0.99% | -0.15% | -0.23% |
| Dynamic model | | | | | | | |
| 0-1km | -21.31% | -30.82% | -7.50% | -5.95% | -5.94% | 0.25% | 0.26% |
| 0-10km | -1.59% | -2.36% | -1.99% | -1.98% | -2.06% | -0.16% | -0.25% |

Notes: This table presents results of the welfare implications of our model. In the model, residents' utility is fixed, so welfare is measured here using the value of landowners' land, net of any transfers. Columns (1) and (2) show the percentage changes in the present discounted value of land, between the scenario without SLR and each of the two SLR scenarios (baseline and fast SLR). Columns (3), (4) and (5) consider the additional economic loss from government subsidies, which encourage the city's (over) expansion, especially the coastal side. Columns (6) and (7) report the percentage change in cumulative land value in the model where the CBD is fixed in the first period relative to a model where the CBD adjusts costlessly each period.